Philosophical Foundations of Information Fusion

Towards a Social-Process Paradigm

Submitted by

Andrew Samuel Dawson
BA, MA.

A thesis submitted in total fulfilment of the requirements for the degree of
Doctor of Philosophy

School of Humanities, Communications and Social Sciences
Faculty of Arts

Monash University
Churchill, Victoria 3842
Australia.

March 2011

Table of Contents

Summary

In the past decade the focus of research into artificial intelligence (AI) has moved from individual agents to multi-agent systems (MAS). This shift has been reflected in the increasing use of social metaphors to describe MAS-related design problems and solutions. The social interpretation of MAS has also been embraced by social scientists interested in developing MAS-based simulations of society. The design of some MAS applications, such as high-level information fusion, involves additional social complexities that arise from their embedded role in complex human social organizations. One response to these developments has been the plea for 'socionic' collaboration between computer scientists and sociologists akin to the inter-disciplinary collaboration that produced bionics. This suggestion has produced some useful results; however the socionic traffic of ideas has been hindered by the paradigm gulf that separates sociology and computer science. Dawson's work is a step toward bridging the philosophical component of that gulf. Dawson's goal is to develop a coherent system of concepts that encompasses both MAS-based and human forms of social interaction and will serve to frame and guide the development of ontology based multi-agent systems. To create this unifying vision Dawson draws upon the philosophy of A. N. Whitehead to develop a socially oriented, Whiteheadian philosophical interpretation of Dale Lambert's Post-Classical MAS.

While Whitehead's work is relatively well known, Lambert's unique process philosophy inspired approach to AI has attracted little attention. Drawing upon Lambert's unpublished doctoral dissertation, *Engineering Machines with Commonsense* (1996), Dawson lays out the broad epistemic issues, metaphysical presuppositions, and heuristic strategies that have motivated Lambert's turn from Classical AI to his process inspired alternative: 'Post-Classical AI'. Lambert's primary objective in his dissertation was to develop a Post-Classical blue-print for a MAS-based mind. To that end Lambert developed Post-Classical philosophies of process, representation and belief, and showed how each of these elements could be formally implemented to provide a Post-Classical artificial mind with its cognitive capability. In order to reveal the deeper implications of Lambert's tersely expressed debt to process philosophy, Dawson provides a commentary which explains the key formal elements of Lambert's work in terms of Whitehead's metaphysical system. Through this interpretation it is possible to view the computational operations of Lambert's MAS as a partial model of the complex creative processes that lie at the heart of Whitehead's metaphysical vision. This interpretation of Lambert's Post-Classical paradigm signals the possibility of a novel 'Whiteheadian AI'.

A key aspect of Dawson's analysis is his identification of the 'sociological turn' latent in Lambert's Post-Classical paradigm. Lambert's approach to grounding agent belief is analogous in form to Whitehead's explanation of the relationship between human ideas and the actualities that they denote. In both cases agents shape and are shaped by social processes that selectively fine tune and prioritize the social customs that mediate interaction between agents and their environments. The sociological dimension of Lambert's Post-Classical paradigm is further illustrated with reference to the ontology-based multi-agent architecture at the core of Lambert's design of a high-level information fusion system. For nearly two decades a process vision has inspired and guided Lambert's internationally recognized contribution to the development of high-level information fusion systems. Interpreted in the light of Whitehead's philosophy, the striking originality of Lambert's Post-Classical paradigm and subsequent work in information fusion is revealed. Dawson concludes that Lambert's Post-Classical paradigm represents a promising nexus of philosophical vision, formal method, and practical technology and outlines his own socionic proposal for its further development.

Statement of Authorship

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis by which I have qualified for or been awarded another degree or diploma.

No other person's work has been used without due acknowledgement in the main text of the thesis.

This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Andrew S. Dawson
Melbourne,
March, 2011.

Notices

Notice 1

Notice 2

Acknowledgements

This work has had a long gestation, and many people have played a part in its realization. La Trobe University played a crucial part in preparing the ground for this project. Nearly twenty years ago I was comfortably engaged in well-trodden line of research into the philosophical foundations of social theory, when I received a call from Brendan Rogers. I relate the story in the introduction, but my thanks to Brendan for shaking me out of my slumbers with respect to the field of artificial intelligence. Without this encounter I would not have developed an interest in the philosophical implications of developments in formal logic and mathematics. This chain of inquiry prepared the ground for my encounter with Dale Lambert's work and stimulated my interest in multi-agent based simulations of society.

I owe a very significant intellectual debt to the late Edmund Husserl and Alfred North Whitehead; in their work I found the encompassing systems of ideas that both prepared the way for the present day computational revolution and shaped the form of modern social theory. Whatever contribution I make to the computational turn in social theory will be due to the scaffold of ideas that these thinkers have provided. I would also like to thank Johann Arnason, of La Trobe University, who supervised and guided my sociological and philosophical studies through the nineties, and encouraged my inquiries into the theories of civilizations implicit in the philosophies of Whitehead and Husserl. His belief in the value of asking the big questions, combined with rigorous standards of scholarship, were in my case not readily reconciled, as witnessed by our unsuccessful ten year long endeavor to bring my doctoral studies to a conclusion. Nevertheless I owe a great intellectual debt to Johann.

My more immediate intellectual obligation in this dissertation is to Dale Lambert; in Lambert I found a worthy heir to the Whitehead and Husserl spirit of combining mathematical rigor with penetrating philosophical analysis. Lambert's graciousness in responding to my dilettantish efforts to understand his work has been much appreciated. I would also like to thank a number of Whiteheadian scholars, including Arran Gare, Michel Weber and Douglas Bendall; at different times their support has played a significant role in the development of this project. Thanks also to Albert Lewis and Stephen Blake for their patience in responding to my questions about Whitehead's mathematical background, and the staff and fellow participants at ANU's logic summer schools of 2005 and 2007 for tolerating a sociologist in their midst.

To my present supervisors, Peter Murphy and Paul Atkinson, of Monash University, many thanks for enjoyable conversations, valuable feedback and wise counsel in steering this work to a conclusion. Joe's Café will occupy a permanent place in my memory as a site of intellectual and gastronomic refreshment. I would also like to express my gratitude for the institutional and financial support provided by Monash University.

Finally, I come to those personal debts that extend beyond this world. I dedicate this work to those that have been most important to me over the years; to my parents, Sam and Dorothy, to my brothers David and Peter, my sisters Gillian and Margaret; and most especially, to my wife, Maria Teresa. Her friendship and support has been an undeserved gift and any thing of value that I create is a paltry offering beside the trust and love that she has bestowed upon me. Lastly, thanks to my children, Thomas and Claire; may there be time for a bit more fun now that this milestone is behind us.

List of Abbreviations

Lambert

EMWC	*Engineering Machines with Commonsense*

Whitehead

UA	*A Treatise on Universal Algebra with Applications*
SMW	*Science and the Modern World*
SME	*Symbolism: Its Meaning and Effect*
PR	*Process and Reality: An Essay in Cosmology*
AI	*Adventures of Ideas*
MT	*Modes of Thought*

Introduction

In this work I develop a Whiteheadian interpretation of Dale Lambert's Post-Classical approach to artificial intelligence. I then attempt to show that this Whiteheadian perspective provides philosophical foundations from which it is possible to obtain a deeper insight into Lambert's work in the field of high-level information fusion. My perspective on Lambert's work is motivated in part by sociological aspirations; from a sociologically informed Whiteheadian perspective, Lambert's approach to high-level information fusion has significant implications for our understanding of and modes of participation in the processes that shape society.

What is high-level information fusion and why should its successful development be of such significance? Information fusion has been defined as "…the process of utilising one or more data sources over time to assemble a representation of aspects of interest in an environment" (Lambert 2001b; 2001a). It had its origins in military systems that were designed to automatically integrate data from multiple sensors to provide a real-time representation of unit-deployments in battle fields, however in the last decade systems have begun to undergo a transformation; instead of simply digitally replicating the 'coloured pins on a map' style of representation, information fusion systems are evolving into ontology-based multi-agent systems that pro-actively initiate situation analysis and plan appropriate responses. The trend since the end of the Cold War, away from conventional warfare towards complex counter-insurgency operations in which economic and socio-cultural factors have a crucial strategic significance, has lent further impetus to the drive to develop 'high-level', symbolically mediated information fusion systems dealing with all aspects of the human situation.

The information fusion system analysed in this work is symptomatic of the shift toward a "semantic web" and semantic machines that communicate on the basis of meanings. The Semantic Web promises to transform the Internet from a domain of human interaction within a passive environment of web pages, into a domain of human and machine-agent interaction, within dynamic, reactive environments. Tim Berners-Lee, inventor of the World Wide Web, describes his vision of the Sematic Web as follows:

> I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers. A 'Semantic Web', which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The 'intelligent agents' people have touted for ages will finally materialize. (Berners-Lee & Fischetti, 1999)

The World Wide Web has already created significant new social spaces and new forms of social interaction and the Semantic Web promises to accelerate these changes. Berners-Lee's vision of the Semantic Web is based upon the development of a new generation of software 'agents' that communicate with each other and with humans on the basis of intended meanings, rather than at the level of syntax. Although machines do not "understand" in the human sense, they can be taught to associate meanings with the information they hold by “constraining possible interpretations through formal logics” (Lambert & Scholz, 2007). The information fusion project that this work analyses is dedicated to achieving this semantic capability but it rests upon a profound paradigm shift; a shift inspired by the vision first outlined by Lambert in his pioneering 1996 dissertation, *Engineering Machines with Commonsense* (*EMWC*).

The notion of ‘engineering commonsense’ is rightly regarded with suspicion by computer scientists and social scientists. Social scientists are very familiar with the dynamic, context dependent nature of ‘commonsense’. Different indigenous societies provide obvious illustrations of the way in which socio-cultural contexts define different forms of ‘commonsense’; but this applies equally to the specialized intellectual communities within a modern society. What is commonsense for a computer programmer is likely to appear a dark art to a sociologist and what appears commonsense for a sociologist may appear to be artsy mumbo-jumbo to a computer engineer. How can machines possibly deal with the complexities that define the context of communications between people?

This was one of the questions in my mind some twenty years ago, when I was first approached by Brendan Rogers, a computer engineer who wanted to discuss my work on the philosophy of A. N. Whitehead (Rogers 1991; Dawson 1991). I had just completed a dissertation which attempted to show that the social systems theory developed by the most influential social theorist of the twentieth century, Talcott Parsons, was grounded at a meta-theoretical level in the conceptual presuppositions of Whitehead’s metaphysical system. Apart from the examiners, a handful of people may have glanced at my thesis, and I was incredulous that a computer science engineer wanted to borrow a copy. His aim, he said, was to extract from Whitehead’s philosophy the conceptual framework of a new approach to engineering artificial intelligence. “What kind of reductionist nonsense is this!” was my first response, followed in equal measure by feelings of curiosity. Later, when he asked if I would be interested in providing regular feedback on his project my feelings of intellectual inadequacy came to the fore: “But I know nothing about computer science!” Rogers was not put off. “Don’t worry”, he said; “the challenge is to clarify the philosophical concepts, the mathematics and logic comes later”.

This reassured me, and a loose collaboration began which was gradually to open my mind to the possibility of a new type of machine ‘intelligence’: one grounded in a process grounded, sociologically oriented vision of human communications. This new vision did not become manifest over-night. As Brendan plunged into the intricacies of Whitehead’s magnum opus, *Process and Reality*, I was pursuing the equally dense details of Edmund Husserl’s philosophy. Periodically, prior to our meetings I received drafts of Brendan’s latest take on Whitehead, and often found myself going back to work that I had thought I was familiar with, only to find new questions and possible interpretations (Rogers 1993a). As the decade advanced I became increasingly conscious of the mathematical and logical inquiries that had shaped the thinking of both Whitehead and Husserl.

Meanwhile, my supervisor Johann Arnason, a leading social theorist in the field of world civilizations, was encouraging me to consider the theories of civilization implicit in the later works of Whitehead and Husserl (Arnason 2003). I found myself in an extremely tangled web! However, by going back to the intellectual roots of

Husserl's and Whitehead's philosophies a striking intellectual panorama began to emerge; a panorama that involved a radical revision of the dominant scientific vision of the Cosmos, and implicitly, a different vision of the nature and role of Reason in the development of human civilizations. It was a vision that opened the way to a new concept of 'machine reason'.

Alfred North Whitehead and Edmund Husserl had both began their careers as professional mathematicians, one in Cambridge and the other in Berlin, working under the tutelage of leading mathematicians of their era. Both made important contributions to clarifying the logical foundations of mathematics and, on the basis of their early work, developed philosophical systems that were intended to provide a new foundation for scientific thinking. By the 1920s, an interest in the broader social and cultural implications of science led both thinkers to speculate on the nature of Reason and its role in Western civilization. Both men had lost a son in the carnage of the First World War, and found themselves trying to come to terms with a world in which pre-war optimism concerning a rational, and enlightened European civilization had been drowned in a sea of mud and blood. While many post-war intellectuals responded to these catastrophic events by arguing that Science and Reason were false Gods, Husserl and Whitehead argued that we had simply failed to understand and unlock their true potential.

Hence in the post-war era, Husserl and Whitehead developed profound critiques of the dominant modes of scientific thought; in both cases the key to their reform of the sciences was a new experientially grounded metaphysical vision of the Cosmos (Husserl 1970c; Whitehead 1926). Both Husserl and Whitehead envisaged a new era of intellectual discovery, rooted in a new metaphysical horizon; this horizon was no dogmatic imposition, but a phenomenologically inspired system of concepts open to continual reformulation in the light of a radical empiricist epistemology and the conceptual constraints suggested by the increasingly powerful tools of formal logic and formal ontology. In addition to founding significant schools of philosophical thought, both thinkers have some claim to being regarded as 'grandfathers' of the present day science of artificial intelligence; equally, it could be argued with some justification that both thinkers are grandfathers of present day social theory. Yet rarely is this intellectual legacy treated as a whole.

One of the scientific disciplines most receptive to the ideas developed by Whitehead and Husserl was sociology. Having developed under the shadow of economics and psychology, sociology in the 1920s and 30s was an infant social science struggling to define its subject area in a way that avoided reifying dualistic dichotomies such as individualism and social holism, and positivism and idealism. Husserl and Whitehead, found a way to bridge these dualisms that proved to be enormously stimulating for social theorists such as Talcott Parsons and Alfred Schutz (Parsons 1937; Schutz 1967 [1932]). The work of these men has in turn played a significant role in shaping modern social theory. As a result of their work, sociology has embraced theories that emphasize the dynamic interdependencies between social environments and agent intentions and avoid the polarities associated with positivism and idealism, methodological individualism and social holism.

While Parsons and Schutz were energizing a new generation of sociologists, cognitive science in the 1950s was responding to the fortuitous intersection of developments in symbolic logic and electronic computers that was to signal a golden age of inquiry. Computational models of brain activity, some focusing on symbols and others on neural connections, were developed. The symbiotic relationship between the cognitive science discipline and engineers promoting the idea of artificial intelligence was reflected in the theories of both disciplines; it was imagined that symbolically programmed computers could simulate cognitive activity, and thereby emulate human intelligence. However, as the first decades of the computational era receded, the dream of artificial intelligence began to appear like the illusive pot of gold at the end of the rainbow.

By the end of the century the field of Artificial Intelligence had become well-defined, but it had shed its early dreams. The repeated failure of 'break-through' technologies to realize anything approximating human intelligence had called into question the ambitions of the founding fathers of AI. However the quest for artificial intelligence had played a role in nurturing significant 'spinoff' developments. Thus specialized branches of AI thrived under other names, providing practical solutions to diverse technical problems. 'Information fusion' is but one example of the proliferation of new fields of AI related research. Behind this manic activity few thinkers found time for philosophical reflection of the sought that Whitehead and Husserl had engaged in. Instead students of the computational sciences generally found that their efforts were consumed by competitive pressures associated with gaining credentials in rapidly developing, highly specialized fields of expertise; fields in which productivity is measured in terms of the output of five page conference papers, the dollar sums associated with successful grant applications and - the gold standard - research projects successfully nurtured via industry joint-partnerships into commercial applications. Unfortunately Rogers was a casualty of this system. By the mid-nineties he was forced to abandon his philosophical quest, and turn toward more commercially oriented projects.

Meanwhile, in Adelaide another Australian AI engineer was engaged in some serious thinking about the philosophical foundations of his discipline. Independently of Rogers, and following a different path, Dale Lambert had also concluded that progress in AI research was impeded by its conceptual foundations, and had concluded that process philosophy had a key role to play in the future of AI.

Lambert's initial research was carried out in the first half of the 1990s while he was working with the Australian Government's Defence Science and Technology Organization (DSTO); his conclusions concerning the philosophical presuppositions of the AI discipline were presented in the previously mentioned dissertation, *EMWC*. Lambert's goal in this work was to develop a new paradigm for engineering artificial intelligence. His subsequent work with DSTO has gone a long way toward validating the practical worth of that early vision and has carried him into domains that have traditionally been the preserve of social theorists. Information fusion, in his hands, has become an endeavour to engineer hybrid societies; societies made up of organized teams of human and machine agents. Thus Lambert's interests have taken him out of the realm of pure speculation into the interdisciplinary domain that wrestles with the psychological and sociological problems associated with deploying an automated information system. In his work I found the path along which to extend Whitehead's ideal of re-thinking our notion of society in terms of a process inspired ontology which could potentially be married with the conceptual frames of meaning developed by sociological theorists.

The work presented in the following chapters centres on Lambert's theoretical proposal for achieving a machine-based version of commonsense; that is, my focus is on the conceptual foundations and design of cognitive machines that can potentially fulfil Berners-Lee's vision of the Semantic Web. If Lambert's machine-agents fulfil their role they will institute a new era in the symbiotic relationship between people and machines, giving rise to a new form of hybrid society. This advance, if Lambert is correct, will build upon a vision of artificial intelligence that is grounded in a different set of philosophical presuppositions than those that shaped Classical AI.

Although process philosophy forms the background to Lambert's Post-Classical vision of AI, he does not describe in any detail the nature of the philosophical horizon from which he draws his process presuppositions. Instead he simply provides a formal definition of his process presuppositions and then builds the formal architecture of his process-based vision of mind on these process foundations. In the course of this reconstruction of the Classical tradition he advances a socially informed view of commonsense akin to that promoted by Whitehead; customs and routines adapted to particular environments rather than an idealized rationality are the basis of commonsense. However, again Lambert's focus is on the formal expression of concepts, not their philosophical justification. Hence, in this work I have attempted to provide a philosophical context for viewing 'commonsense' in terms of those process presuppositions.

Since my aim is to build a link to social theory, Whitehead's philosophy has a dual function; both to provide a broader horizon in which to situate Lambert's Post-Classical paradigm, and to provide a pathway into the conceptual systems developed by social theorists. As will become apparent, although Lambert has gone a long way toward adopting a social understanding of commonsense, he has not turned to sociological theory for inspiration. This is the 'socionic turn' that I believe is the natural fulfilment of his Post-Classical paradigm, and in order to show how this can be achieved my longer-term goal is to connect the language of Lambert's Post-Classical paradigm with the conceptual frameworks developed in social theory. In this work I decided to pursue this aim through Whitehead's metaphysical system. Although Husserl also aspired to create a process metaphysical system, Whitehead was able to bring his vision to a far more complete state and is generally recognized as the definitive process philosopher of the twentieth century. Thus in this work I shall present a Whitehead inspired reading of Lambert's ideas and lay the foundations for a deeper conceptual engagement between Lambert's Post-Classical paradigm and the social sciences.

Although this work is not a commentary on the Socionics research program, the latter has played a significant role in the development of this study.[1] The Socionic paradigm

[1] In the 1960s, Nicolas Rashevsky (1966), a pioneer of 'mathematical biophysics', coined the term 'Socionics' to describe a mathematical approach to the analysis of complex systems that looked 'upward' in the chain of complexity to sociology, rather than downward to physics, for inspiration. Rashevshy played a key role in founding the discipline of biophysics (Cull 2007) and an influential role in the founding of mathematical sociology (Fararo 1997). In the late nineties, independently of Rashevsky, a team of German sociologists and computer engineers began to use the same term to

is a useful complement to Lambert's Post-Classical view of AI, as it too shifts the focus away from the strategies and concerns of Classical AI toward socially defined concerns, but it goes a step further and explicitly engages with sociological theory in its quest for a conceptual framework and appropriate research methodologies. To date the advocates of a Socionic approach have not placed great stress on philosophical foundations, and I believe that this is a weakness which this work may help to address.

Computer scientists do not generally write commentaries on sociological theory and vice versa, sociologists tend to refrain from writing commentaries on developments within computer science. Until recently there was little motivation on either side to alter this situation. However, the Socionic turn was born out of a changed perception of where the future is leading both disciplines; Socionics is premised upon the possibility of a collaboration that crosses traditional disciplinary boundaries and encourages such forays. This impetus has come from a convergence of interests, stemming from engineering needs that arose in the context of research on distributed artificial intelligence (DAI), the subsequent development of socially inspired multi-agent systems (MAS), and the increasing number of sociologists attracted by the prospect of utilizing multi-agent based software to create simulations of various social situations. These developments took place against the back-drop of the plague-like expansion of the World Wide Web, the proliferation of on-line multi-player games, and the emergence of a new generation that takes for granted the steady progress toward more believable virtual worlds.

This is the historical context which has defined this Whiteheadian reading of Lambert's Post-Classical blue-print for information fusion. In the course of writing this work, I have had to make compromises. This work is only a beginning of the envisaged socionic conversation between Whitehead's philosophy, Lambert's Post-Classical paradigm, and Parsons' theory of social action. In my initial outline of this project, I had hoped to carry the conversation further and I developed drafts of chapters dealing with the metaphysical, social theoretical and computational dimensions of the socionic quest to frame the civilizing process; however, as I became more engaged in making specific links with Lambert's project it became obvious that time and space would impose constraints upon my ambitions. Thus this work only partially fulfils my initial aims. I have initiated a conversation between the philosophy of Whitehead and Lambert's computational paradigm; however I have yet to outline the connection between these domains and social theory. The computational model of the civilizing process must wait for another work, but as will be indicated in the chapter outline, some of the conceptual horizons for this model have been outlined.

---

describe their effort to develop multi-agent system (MAS) software, inspired by sociological theory. The goal of Socionics, as they conceived it, was to marry together computational and sociological research, and to apply the product of this research to the development of hybrid, human and machine agent-based societies. Thus the research oriented strand of Socionics was oriented to building sociologically inspired MAS software that could serve as the basis of MAS-based simulation of human society, which could then experimentally underpin an ongoing cycle of critical reflection upon and development of more adequate sociological theories, leading to further developments in socionic software, and so on. As an application spin-off, by building machine-agents with social capabilities analogous to those found in human society, socionic software could be the basis of functional, human-machine 'hybrid societies'. Socionics, so conceived, has the potential to frame research into the development of hybrid societies and their associated technologies (Malsch & Schulz-Schaeffer 2007).

Thus, in the chapters that follow I first begin by contrasting a narrow and broad notion of Reason; the narrow notion has been shaped by our understanding of the role played by logic in the workings of Reason, while the broad notion has been shaped by our understanding of the role of Reason in the formation of Western Civilization. The first chapter focuses on elucidating the narrow context defining our notion of Reason, by tracing the development of formal logic from its Aristotelian origins through to its machine implementation in computer programs.

In the second chapter I focus upon the philosophical presuppositions that framed the Classical AI attempt to replicate the problem-solving capabilities of the human mind. My focus is on Lambert's critique of the epistemological, heuristic and metaphysical presuppositions that framed and motivated the Classical quest. A key concern is with the linguistically mediated notion of representation that has guided Classical AI's vision of rational knowledge and commonsense judgments. Lambert's critique provides the justification for dispensing with key presuppositions of Classical AI and points the way toward an alternative position that adopts a situated orientation, retains a notion of representation, but dispenses with the linguistic, overly-rationalist notions of Classical AI.

The third chapter steps away from the narrow technical concerns that has motivated Lambert's critique of Classical AI and considers the broader philosophical context of Classical AI's mistaken notions of representation and rationality from the standpoint of the critique developed by Whitehead. This chapter then develops Whitehead's alternative, process-grounded vision of the nature of representation and commonsense, focusing on the societal basis of enduring objects, and the role that customs play in both defining these societies and in defining our representation of these societies. I conclude this chapter by making some important distinctions about the different foundations of human and machine intelligence; these distinctions have significant implications for the potentialities and limitations of machine 'intelligence'.

In the fourth chapter I begin my detailed examination of Lambert's Post-Classical proposal. Firstly I outline how Whitehead's philosophy can be used to provide a phenomenological and model-theoretic justification for Lambert's set-theoretical approach to creating a formal ontology of processes. Secondly, I briefly outline Lambert's ontology, noting some connections with Whitehead's formal ontology of extensive connections and arguing that both ontologies serve a similar foundational role in grounding subsequent notions of representation.

The subject of the fifth chapter is Lambert's theory of representation. I make a connection between Lambert's formal notion of representation and Whitehead's notion of societies, building upon technical linkages between Whitehead's notion of the 'propositions' that guide societal formation and the 'templates' that mediate identification of Lambert's self-replicating patterns of process. A focus is on showing how Lambert's combination of hardware sensors and software templates provides his agents with a situated connection to the world that is analogous to the connections at the core of Whitehead's Cosmological vision.

The sixth chapter turns from representation to belief. The focus of this chapter is on the semantic foundations of Lambert's system of representation. I show that

Lambert's semantics has essential structural parallels with the semantic ground of meaning that guides the process of event formation in Whitehead's system. In both cases perceiving subjects are responding to causal influences arising from the immediate physical environment but are interpreting those influences in terms of custom defined patterns of expectation. Ultimately a 'semantics of success' defines the world of meaning that guides agent behaviour. I show how Lambert builds upon his process grounded notion of representation to provide a mathematically based approach to simulating the types of selective processes that define purposeful activity in Whitehead's metaphysical system.

The seventh chapter provides a brief outline of Lambert's proposal for a Post-Classical mind. I outline his proposed architecture, showing how Lambert proposes translating the notions of process, representation and belief outlined in the previous chapters into an implementable strategy for engineering a multi-agent based, situated robot/mind. This chapter concludes with a review of strengths and weaknesses of Lambert's Post-Classical proposal, drawing upon the Whiteheadian horizon outlined in previous chapters to suggest that Lambert's paradigm needs to pay more attention to the ontological structure of the social environments that agents are interacting with. This is a theme that Lambert takes up in the post-*EMWC* phase in the development of his Post-Classical paradigm.

In the eighth chapter I summarize the progress made by Lambert and his team in implementing a Post-Classical inspired information fusion system. I outline the military context of this work and explain the intended role of an information fusion system. I then outline the architecture of his system and show that it revolves around a process ontology and multi-agent system that is broadly based upon principles elaborated in *EMWC*. My focus in this chapter is on the social context of information fusion and the important work that Lambert and his team have carried out in defining a multi-level ontology that incorporates explicit representation of social interaction and communication between human and machine agents. This chapter concludes with suggestions concerning the further Socionic development of the Post-Classical paradigm to reflect the hybrid-societal context of Lambert's information fusion system.

The conclusion briefly reviews the preceding chapters and concludes that the quest within computer science to interact with people in a manner that reflects intuitive human modes of understanding is leading AI towards Whitehead's 'broad', historically situated vision of Reason. Reason is part of a broader process of societal development that is framed by Whitehead in Cosmological terms, and by sociological theorists in terms of historical societal processes. These dynamic societal processes are the ultimate defining context which high-level information fusion systems must both represent and become participants in defining. Thus the flow of 'intent' that Lambert envisages as the defining characteristic of the societal level of information fusion becomes part of the flow of communications that, from the perspective of social theorists, define the various structured semantic environments of human and machine agent action. Engaging with these themes represents the next, Socionic phase of this project.

Lastly, an apology to those well versed in the sciences of artificial intelligence for this inter-disciplinary transgression. Feelings of intellectual inadequacy are inevitable when engaging in a cross-disciplinary study such as this. The time of polymaths is past, as there is simply too much material for any one person to master. Nevertheless, I believe that the need to bridge interdisciplinary divides provides at least some justification for venturing into mathematical and logical domains which I should respectfully leave to more qualified students. I know that I still have much to learn; but the technological impetus carrying the world into an era of hybrid societies has created a pressing need for an exchange of perspectives between the humanistic and engineering disciplines. To that end, my need to begin from the basics in the engineering domain may create a more accessible path into the field of information fusion for those like me, who have a background in the humanities and little exposure to the mathematical and logical foundations of the computer sciences. Having made my excuses, there is a need for the spirit of tolerance to run both ways. If the Socionic conversation that I envisage is to take root, then those in the humanistic sciences need to encourage and welcome the forays of engineers into their domain, even if at times their speculations on the nature of society seem a little ill-informed.

Chapter 1

The quest to make machines that 'think'

Chapter One – Section Headings

1.0 Introduction

The endeavor to engineer an information fusion system has a broad and narrow context; the broad context concerns the role of Reason in human civilization, while the narrow context concerns the attempt to formalize the Reasoning process. In this work my aim is to develop a process oriented philosophical approach to the computational sciences and the computational modeling of human society that acknowledges and harnesses the philosophical tensions that are common to both the narrow and broad perspectives on Reason. What are these tensions and how are they manifested in computer science and sociological theory?

Knowledge and Reason have always occupied an ambivalent place in human mythology, and the quest to build thinking machines awakens a similar ambivalence. In J. R. R. Tolkien's fantasy, *The Lord of the Rings*, knowledge and its applications have two faces: Mordor's machinery is a symbol of knowledge used to serve humanity's satanic urge for power while the Elf's mystical arts symbolize the use of knowledge to serve humanity's finer side. The same ambivalence is present in the classical notion of Reason and contemporary notions of rationality. Ancient philosophers, for example, identified Reason with the Divine Logos at the heart of the Cosmos and viewed the exercise of Reason as a path to liberation, while influential modern philosophers and social theorists have frequently portrayed scientific rationality as a dehumanizing, oppressive agency.

The founding fathers of the sociological tradition were greatly influenced by the Romantic cultural movement and tended to share Tolkien's pessimistic assessment of the impact of industrialization. For the great German sociologist, Max Weber, modernization was an 'Iron Cage' and scientific rationality was the main culprit in bringing about the disenchantment of European culture and society. Influenced by Weber and the work of Martin Heidegger, a number of European sociologists have subsequently explored the social and cultural impact of science and technology – generally reaching negative conclusions. In contrast to these intellectual tendencies, Alfred North Whitehead made significant contributions to the development of modern logic and developed a profound critique of the Western scientific tradition and of popular notions of scientific 'Reason'; yet despite this critique he retained a faith in Reason in a deeper philosophical sense, and remained an enthusiastic advocate of the positive potential contribution of logic, mathematics and science to human civilization.

In order to appreciate Whitehead's nuanced position regarding the role of Reason, logic and mathematics in our society, it is necessary to appreciate the remarkable developments that have taken place in these fields over the past couple of centuries. The historical quest to clarify and formalize our notion of Reason (culminating in the development of Computational Science and the endeavor to engineer intelligent machines) represents an epic chapter in the saga of human civilization – a chapter full of triumphs and pitfalls, and worthy of investigation for its own sake, but essential reading for an attempt to understand the development of a high-level information fusion system that has the capacity to model human societies and their constitutive civilizing processes.

The purpose of this chapter is to sketch out the narrow context of information fusion in terms that a reader with a humanities background and an aversion to formal logic and the mathematical sciences might appreciate. My approach is to provide a historical introduction to what has been one of the great intellectual adventure stories of the millennium. This will serve as a preface to a more detailed examination of the philosophical presuppositions of Classical AI, and an examination of Whitehead's metaphysical cosmology. Whitehead's philosophy provides a context that can bring together the narrow computational and broad sociological understandings of 'Reason', and potentially provide a framework within which the human sciences and the computational sciences can build a positive understanding of the potentialities of information fusion.

The prospect of creating automated information fusion systems is the culmination of a drama with many twists and turns. Automated information fusion systems are designed to integrate information from disparate sources and to extract from that information representations of situations that are of interest to the person using the system. They are, in effect, the electronic personal assistant that science fiction writers have long dreamed of, able to gather information from sources such as electronic sensors and public media, update their knowledge base, and respond to queries from human users or other machine agents. This sounds like the old dream of artificial intelligence (AI) and that is indeed so.[1]

---

[1] Due to exaggerated claims and unfulfilled promises the credibility of AI as a scientific discipline has suffered. As a result, many AI-inspired projects are marketed under names that are not tainted by the failed dreams of the past, such as "informatics, machine learning, knowledge-based systems, business

The technology underpinning information fusion systems draws upon advances made by engineers working in the field of AI, and Lambert's own project developed out of his critique of Classical AI (Lambert 1996). Thus, the background to information fusion is the development of modern symbolic logic, and related attempts to develop symbolic systems for classifying and representing phenomena of the real world. My brief tour through this field is designed to provide a non-technical introduction to the remarkable discoveries in logic, mathematics and computing that have transformed and are continuing to transform our modes of thinking, communicating, reasoning and making decisions. These discoveries form the essential background to the emergence of automated information fusion and the prospect of a new form of collective decision making process, based upon intelligent human and machine agents wedded together in hybrid virtual-societies.

1.1 The creation of modern Symbolic Logic and the birth of Computer Science

The present day computer sciences, including artificial intelligence (AI), have their origins in the endeavour to formalize deductive reasoning in an abstract symbolic language, and to use this formal language to represent and compute facts about our world. AI has focused on the question of implementing these processes artificially while an allied cognitive science has utilized this computational model of reasoning as a model of human cognition. The concepts of mind and reason that have inspired the quest to develop artificial intelligence have their roots in Greek philosophy, and the subsequent transformation of the Greek legacy by Rationalist and Empiricist philosophers during the Enlightenment. The Enlightenment transformation of the notions of mind and rationality and the resultant clarification of the nature of logical analysis deserves to be ranked alongside the Enlightenment transformation of Greek Physics as a great milestone in the shaping of our modern world. Just as the revolution that began with the work of Copernicus and culminated with the Newton's cosmological system transformed our view of the external universe, so also the work of modern logicians, from Boole through to Turing, has transformed our understanding of mind.

These revolutionary changes in our view of the universe, and in our understanding of the mind, carried a price. Unquestioning acceptance of the presuppositions of Newton's paradigm ultimately inhibited progress in understanding the physical world; so too the dominance of the computational model of mind have inhibited attempts to simulate human intelligence. However, the "failures" of the Newtonian paradigm need to be weighed against its tremendous achievements, and the modest success of the computationally inspired classical approach to AI needs to be viewed in the context of the enormous success of the computational revolution.

Both the strengths and weaknesses of the modern computational paradigm are rooted in the historical genesis of logical "computation". This genesis is a remarkable story that leads us to the heart of Western civilization. It began with the dream of revealing

---

rules management, cognitive systems, intelligent systems or computational intelligence". See http://en.wikipedia.org/wiki/AI_winter the essence of Reason, and culminated in the design of thinking machines.[2] Logic is the phantom presence that energizes the drama: it is the willow-the-wisp that has beguiled philosophical minds with its promise of a superior form of understanding; it is the subtle form that weaves empirical evidence into scientific explanation. Thus the historical saga begins in Classical times, with the complex philosophical legacy of the ancient Greeks.

1.11 Lull's Reasoning Machine: first steps towards AI

Modern science, including logic, has its roots in the work of Aristotle. For two thousand years, Aristotle's approach to logic defined the discipline. His work inspired the intellectual revival that took place in medieval Europe, and underpinned the first mechanical devices for reasoning, developed by thirteenth century Spanish philosopher, poet and missionary, Raymond Lull.

Lull's goal was to create an automatic system for answering questions (such as *What? When? Where? How? Why?)* about the most important issues in life (such as *What is man?* and *Why do we exist?*). His machine was to combine the two great sciences of his time, Logic and Theology, into a system that could provide the solution to any question, drawing upon a database that expressed sets of valid propositions concerning the nature of God and the Universe.[3] From the perspective of my efforts to build a computational model of the civilizing process, it is interesting to note that the spiritual motivations behind Lull's invention of the first AI: he was seeking to create a method for resolving the most profound questions of his time. Lull's machine can be viewed as part of a continuum of symbolic innovation that began with the first cave paintings, whereby people have sort to use external artefacts to represent, clarify and actively shape the nature and meaning of human existence.

Lull's idea of automated reasoning had its roots in Aristotle's notion of the syllogism. According to Aristotle, logic was the basis of all human knowledge, and his analysis of syllogisms was the first attempt to give formal expression to the way in which reason works. He defined knowledge in terms of true propositions and in his work on logic he investigated the way propositions are combined together in arguments so that the reader is persuaded, purely because of the formal pattern linking one proposition to the next, that a conclusion is true. Thus the familiar example of a syllogism:

All men are mortal (major premise).
Socrates is a man (minor premise).
Therefore, Socrates is mortal (conclusion).

Has the abstract form:

All S are P.
a is S.
Therefore, *a* is P.

Aristotle demonstrated that, provided that the premises are true, if an argument follows one of a number of valid formal patterns, then the conclusion must be true.

---

[2] Very few popular accounts do justice to this story. Notable exceptions are Heinz Pagels' classic text, *The Dreams of Reason* (1988), and Keith Devlin's *Goodbye Descartes* (1997), a very engaging and accessible account of the limitations of logic as a tool for building machines that think.

[3] Lull's machine consisted of nested circular disks, each inscribed with the attributes of a sphere of Being, and capable of being aligned to generate different combinations of attributes. See Sowa (2000:4-6).

Lull joined this mechanical technique of utilizing patterns of formal inference to arrive at true propositions, by building upon another of Aristotle's great achievements: his system of categories or 'ontology'.

Ontology is just as fundamental to modern AI as logic, since it underpins the creation of the databases that computational systems manipulate. Computing utilizes abstract representations of knowledge and it was Aristotle that developed the first systematic procedure for abstracting and classifying knowledge. Lull used a modified version of Aristotle's categories to represent all the possible natural things as sub-species of one supreme genus. He arranged propositions into various categories, and formulated these into tables; the hierarchical organization of these tables into genii, species and possible sub-species of propositions enabled the knowledge seeker to test whether particular combinations of propositions were logically valid. Lull had anticipated one of the most basic AI heuristics: "generate and test". Unfortunately, with multiple lists of species and sub-species, the number of possible combinations of propositions increased exponentially, a problem computer programmers still face when designing an effective system.

1.12 From Leibniz to Principia Mathematica

After Lull came a series of developments. Gottfried Leibniz, the seventeenth century German polymath, was intrigued by the possibility of resolving arguments automatically, by way of a procedure akin to a mathematical calculus. He suggested that all the ideas that mediate human thought are compounded from a relatively small "alphabet" of simple ideas. He envisaged the compounding process as a simple multiplication, analogous to the arithmetic operation. His crucial intellectual contribution to realizing this vision (he termed it his "wonderful idea") was recognition of the key role that abstract symbolic notation plays in arithmetic. He envisaged doing for human thought what algebra had done for mathematics. Thus, complex disputes could be solved by breaking down the argument into its simple elements, representing them with mathematical symbols, and resolving them by an elementary process of calculation. However he made minimal progress toward realizing this vision. The real break-through came in the nineteenth century, when the son of an English shoemaker, George Boole, turned his mathematical expertise to the subject of logic. This intellectual love affair brought forth his groundbreaking publication: *An investigation into the Laws of Thought, on Which are founded the Mathematical Theories of Logic and Probabilities* (1854).

The practical realization of Leibniz's dream came with Boole's simple algebraic reformulation of logic and Frege's development of predicate calculus. Boole developed a propositional calculus for which various axioms hold, such as the commutative and associative laws, based upon joining simple propositions, represented by symbols (such as *a, b, c*), using three operators, AND, OR, and NOT ($\wedge$, $\vee$ and $\neg$) and two additional elements (0,1) representing a false and a true proposition. However, Booles' logic had a limited capacity to capture aspects of human reasoning. For example, using Booles' system, the major and minor premises, and the conclusion of Aristotle's previously cited syllogism could be expressed as three propositions, *a*, *b*, and *c*, and the syllogistic relationship represented as:

- $a \wedge b \therefore c$

An expression of this sort is not very enlightening. What is needed is a method that captures the inner logic linking each proposition to the next. Frege's predicate calculus met that need.

Frege's development of the predicate calculus (also referred to as First Order Logic) was an extension of Boole's propositional logic. Frege used variables and quantifiers to produce a more fine-grained analysis of the logical structure of sentences. For instance, according the Frege, the proposition "All men are mortal" (which in Boolean logic might be represented by a single letter "*a*") has a complex formal logical structure, which could be captured by using a constant and free variable bound together as function-arguments; the precision of this language could be further enhanced by utilizing the quantifiers 'FOR ALL *x*' ($\forall x$), and 'THERE EXISTS SOME *x*' ($\exists x$) with which to further bind and qualify free variables. Using these developments, the Aristotelian syllogism previously cited would be expressed:

- $\forall x\ (Man(x) \rightarrow Mortal(x))$
- *Man(Socrates)*
- $\therefore$ *Mortal(Socrates)*

In plain English, this reads
- for all x, if x is a man then x is mortal
- Socrates is a man
- therefore Socrates is mortal.[4]

These developments showed that Aristotle's syllogistic system of reasoning could be encompassed within the predicate calculus, and further, it suggested that formal logic encompassed a far greater domain – a domain that could be extended with further refinements, to encompass an ever expanding family of logics that included within its scope the various branches of mathematics. The great Italian mathematician and logician, Giuseppe Peano, played a key role in showing that mathematics was founded upon systems of logical axioms, and developed a simple notation for translating the language of mathematics into logic. Frege's predicate calculus and Peano's simple notation paved the way for Whitehead and Russell's monumental effort to demonstrate that all the fields of mathematics could be logically derived from a simple system of self-evident axioms and inference rules.

In *Principia Mathematica* (1910-13) Whitehead and Russell utilized the new language of symbolic logic to show that the main branches of mathematics, including the various fields of analysis associated with numbers and geometrical forms, could be derived from a simple set of axioms.[5] Unlike the proofs traditionally utilized in arithmetic and geometry, intuition played no part in the formal validation process; Whitehead and Russell's demonstrations rested entirely upon mechanical deductions, based on systems of axioms and rules of inference. These developments signalled the

---

[4] Where "$\forall x$ " means "for all x", "→" means "implies", Man(Socrates) means "Socrates is a man", and Mortal(Socrates) means "Socrates is mortal".

[5] Although the intended fourth volume on geometry was not completed, in the three volumes that were published, they demonstrated that mathematics and formal logic were intimately connected, and inspired other mathematicians with the dream of demonstrating that mathematics was a logically complete and closed system of truths.

birth of modern Logic, and paved the way for the breakthrough into the age of modern electronic computers.

Whitehead and Russell's *Principia Mathematica* was crucial in another way. Not only did their work transform our vision of mathematics it also transformed our notion of language. That is, if logic was an instrument of mathematical rigor,[6] then this meant that language, to the extent that it was logical, was subject to formal analysis of the type developed in mathematics and could aspire to the same standards of logical rigour. Thus, building on the work of Russell and Whitehead, Wittgenstein argued that the new predicate calculus could be used for the logical analysis of language, and for the clarification of the nature of the relationship between language and reality. His work inspired an influential philosophical movement, dedicated to the logical purification of natural languages (especially by disclosing the supposedly fraudulent logical basis of metaphysical and ethical claims), and the clarification of linguistic expressions.

Philosophy, from the standpoint of linguistic analysis, was dedicated to clarifying what can be meaningfully stated; philosophers, once dedicated to a quest for wisdom, had become the sanitation engineers of the spoken word. Alfred Tarski, another great logician of the twentieth century, introduced the culminating element in this transformation of the Aristotelian legacy by creating a theory of reference that enabled truth to be deduced in a rigorous manner for any proposition generated by a formal language, provided that the formal language was itself specified in terms of a 'metalanguage'. Tarski's work enabled syntactically generated expressions to be given a semantic interpretation, enabling formal propositions to be linked to theories about the physical world in a precise, logical manner.[7] His semantic conception of truth provided the creators of formal languages with a way of avoiding the paradoxes that Russell had feared might undermine his attempt to place mathematics on logical foundations. These developments led to a convergence between logic and mathematics, liberating logic from natural language and mathematics from its arithmetical and geometrical roots. Formal logic, the child of this marriage between

---

[6] That is, capable of giving rigorous expression to the foundations of arithmetic, algebra, geometry and analysis, and equally rigorous proofs - purely syntactic in form - of the various theorems of mathematics relating to the study of quantity, structure, space, and change.

[7] Tarski's (1944) semantic theory of truth began as a successor to the correspondence theory of truth, whereby a proposition is true of it corresponds to a fact in the real world. Tarski proposed instead that truth in artificial formal languages can be defined by interpreting the propositions of a formal language in terms of a metalanguage. Take, for example, the ambiguity concerning the truth of the sentence "the square of any number is positive". Ambiguity can be removed by specifying the meta-language that the sentence refers to (in this case, the sentence is true for the language defined by the set of real numbers but false if it refers to a language that includes complex numbers). Tarski's approach led to the model theory for quantified predicate logic. Using this approach, the truth of a simple proposition such as "Socrates is mortal" can be specified by saying:

- If "Socrates" is a name and "is mortal" is a predicate, then "Socrates is mortal" expresses a true proposition if and only if there exists an object x such that "Socrates" refers to x and "is mortal" is satisfied by x.

In Tarski's formal language of predicate logic, this would be put more generally as:

- If "a" is a name and "Q" is a predicate, then "a is Q" expresses a true proposition if and only if there exists an object x such that "a" refers to x and "Q" is satisfied by x.

Donald Davidson and others have extended Tarski's approach to encompass natural languages (Devlin, 1997:86ff.; Dowden&Swartz, 2006).

two ancient disciplines, revealed new powers and horizons that went far beyond those anticipated by the parent disciplines. The computational era was about to begin.

1.13 Toward a New World: Gödel, Turing and the expanding Universe of Mathematics and Logic.

By the 1930s, the convergence of mathematics and logic had given rise to several relatively independent developments. First, the work of Zermelo in 1908 and Fraenkel in 1922 resulted in the canonical axiomatic set theory ZFC, which was thought to be free of the type of paradoxes first discovered by Russell. Formalized using first-order logic, the ZFC form of set theory became the foundational language of mathematics, and coupled with the axiomatic approach to mathematics developed by Peano, Russell, Whitehead, Hilbert and others, became the dominant medium of mathematical activity.[8] Dedicated to founding all of mathematics on set theory and striving for the maximum rigour and generality, the Bourbaki group became the embodiment of this new axiomatic spirit (to the despair of those students who enjoyed a more intuitive approach to mathematics). Second, the logical analysis of language gave rise to a form of philosophical activity that was to dominate twentieth century Anglo-American philosophy: linguistic analysis. The third development concerned the logical presuppositions of mathematics. Instead of transforming mathematics into an impregnable logical fortress, the attempt to ground mathematics in logic instead transformed our vision of the nature of mathematics.

Until the late nineteenth century, the mathematical edifice had rested upon three unassailable pillars. First, mathematics rested upon self-evident axioms (e.g. Euclid). Second, it was based on formal logic, as developed by Aristotle. Third, by combining self-evidently true axioms and Aristotelian logic, mathematicians could deduce theorems that were necessarily true. Up until Locke's analysis of the empirical method, it was believed that all the sciences could be erected upon these pillars. Locke broke with the Aristotelian notion that scientific Reason was essentially concerned with truthful Ideas, and the corresponding assumption that sense experience is not the judge of scientific propositions but merely a conveyance for

---

[8] Viewing the computer and mathematics as complementary instruments for exploring this new domain leads to a constructivist view of mathematics and challenges some of the assumptions that, since Cantor, have informed the set theoretical approach to mathematics. Set theory (the universal language of modern mathematics) incorporates in its axiomatic foundations some of the philosophical presuppositions that framed Hilbert's vision of mathematics. Some philosophically minded mathematicians have objected to using set theory as a foundation for mathematics (eg. Brouwer, Weyl and Gödel). Although set theory has proved to be a tremendously powerful tool, its application to real world phenomena does involve a counter-intuitive appeal to fundamental elements that have no phenomenal correlate. For this reason both Whitehead and Husserl developed mereological, rather than set-theoretical foundations for their formal ontologies (see Varzi, 1994; Smith, 1998; Cohn & Varzi 2003). Although Whitehead is sometimes mistakenly associated with Hilbert's formalist approach to the philosophy of mathematics, he has far more in common with the phenomenological approach of Husserl (ie. he strongly emphasized the role of subjective intuition and viewed mental construction as an essential aspect of the process of mathematical discovery, see Dawson, 2008). In their later philosophies, both Husserl and Whitehead, prepare the ground for more radical social constructivist approaches to mathematics and logic. One of the subtexts of this thesis is that mathematics is more than formal proofs about the properties of ideal structures. It is part of a creative process of human engagement with our cosmos. To reduce mathematics (or science) to a fragmented assortment of puzzles, and to strip those puzzles of a larger cosmological interpretive context, is to deprive mathematics (and science) of an important source of direction and motivation.

disclosing to the mind notions which it rationally intuits to be self-evident truths. As Fred Wilson, a pre-eminent philosopher of science has argued, Locke introduced a new notion of logic to science.[9]

Prior to Locke, logic was identified with syllogistic reasoning; after Locke the challenge was to articulate the principles of inference that apply to particular empirical sciences. That is, rather than posit a logical consistency that must necessarily hold between notions and propositions, Locke inaugurated a new scientific ideal, in which the goal of science was to articulate empirical methods that could be used in tandem with theoretical models to conceptualize and explain particular phenomena. Rather than a being based upon the certainty of innate ideas and demonstrable knowledge the empirical-experimental approach to science was based upon a practical appeal to experience and physical evidence. Henceforth the empirical sciences began to move away from the Aristotelian ideal of science, as exemplified in the mathematical model of axiomatic explanations, where explanation meant subsuming a domain of phenomena under a 'covering law'; instead they began to focus more on empirically and experimentally exploring the mechanisms that operate in nature, and developing explanations that consisted of representations of the dynamic interconnection of the parts that make up these functionally inter-dependent wholes.[10] Only in the realm of pure mathematics did the Aristotelian vision of a purely rational science retain its validity.

In the nineteenth century this pure mathematical edifice was shaken by several discoveries, but these discoveries, though challenging aspects of the Aristotelian edifice, initially held out the promise of re-affirming the notion that mathematics was Plato's ideal realm of pure reason. The first of these shocks was the discovery of non-Euclidian geometry; this led to the realization that new, consistent, and sometimes useful systems of mathematics could be developed by experimentally manipulating Euclid's axioms, postulating for example, that the seemingly self-evident 'parallel postulate' is false, and that parallel lines do meet. Second, beginning with the work of George Boole, Aristotelian logic was shown to be part of a broader logical scheme that could itself be treated as a branch of algebra.

Though disconcerting, Boole's discoveries, along with those of De Morgan, the Grassmann brothers, Hamilton, and others, laid the foundations for the discoveries of Peano, Frege, Whitehead and Russell, and the eventual realization that all the fields of mathematics, including the traditionally divided domains of arithmetic and geometry, could be unified and treated in a purely formal, axiomatic manner. The great achievement of Whitehead and Russell's *Principia Mathematica* was that of

---

[9] Aristotle assumed that the senses mediate intuited truths, and "they - not sense - constitute the judge of what is science. These intuited notions are unpacked by the rational faculties of the mind into judgements and syllogisms, at which point these notions constitute science as a body of self-evident and necessarily connected truths" (Wilson, 1985a: 71-2).

[10] While the nature of scientific explanation remains controversial, empirical studies of the practice of science have increasingly led researchers to question the validity of Carl Hempel's deductive-nomological or 'covering law' (1965) model of scientific explanation, with its ideal of deductive laws that subsume a domain of phenomena (via a formalized argument that resembles the Aristotelian ideal of scientific explanation). Historians and philosopher's of science are instead focusing on the pragmatic character of science, as reflected in the instrumental success of causal models as a basis for answering questions about the systems that are being studied. See, for example, Salmon (1984); Bechtel & Abrahamsen (2005); de Regt (2006).

demonstrating the unity of mathematics by showing that the various fields of mathematics could all be derived from a few simple logical axioms. For those traditionalists shaken by the discovery of non-Euclidian geometry, the notion that the realm of mathematical truths could be encompassed by a few simple axioms seemed to re-affirm the notion that people could enter into complete possession of a field of knowledge. Unfortunately the traditionalists were to be disappointed; the endeavour to place mathematics on firm logical foundations was to have unanticipated consequences.

In *Principia Mathematica,* Whitehead and Russell had demonstrated that mathematics could be treated as an exercise in formal logic, with various fields of mathematics, and related theorems arising as deductions from a logically rigorous foundation made up of a relatively small number of carefully defined axioms. There were, however, some foundational issues that their work failed to clarify. The great German mathematician, David Hilbert outlined the important outstanding tasks in a speech in 1928, when he challenged his fellow mathematicians to provide a number of proofs, including proofs that mathematics is complete (there are no mathematical propositions that cannot be proven to be true or false), consistent (a proposition cannot be shown to be both true and false), and decidable (there is a method, applicable to all mathematical proposition, that will demonstrate the truth of the two previous claims).

In response to Hilbert's challenge, Gödel developed his celebrated proof of the incompleteness of mathematics, showing that there will always be an infinite number of mathematical propositions that are both true and unprovable. He also showed that, although mathematics must be consistent, this consistency could never be proved. Finally, another celebrated mathematical genius, Alan Turing, showed that mathematics was not decidable in Hilbert's sense. Most mathematicians had little interest in the more philosophical dimensions of their subject and hence they ignored these strange results and continued to pursue practical mathematical goals in their specialist fields, using the new tools of set theory. However, in the light of subsequent progress, it is apparent that these refutations of Hilbert's conjectures marked the death-knell of the Aristotelian conception of science, and the opening up of a new and infinite mathematical horizon that can be likened to the Copernican revolution in physics.[11]

1.14 The New World of Computer-based Mathematical Exploration

Copernicus challenged the limited horizons of the medieval world by postulating that the Sun, rather than the Earth, was the centre of the universe; Galileo demonstrated that the telescope was the appropriate vehicle to explore this expanded universe. Together they inaugurated a cosmological revolution that was to vastly expand our sense of the cosmos and the relative significance of the human presence in that cosmos. Newton's empirical observations and mathematical formulations validated this revolutionary approach, and Locke generalized from Newton's methodology and created a new ideal for scientific investigation. Likewise, Gödel's work demonstrates that we can only discover a fraction of the mathematical universe, and moreover, subsequent discoveries show that unsolvable problems are more the norm than the exception, even when our definition of 'solution' is extended to encompass

[11] These linkages and their implications are explored in Robertson (1998: chapts. 1 to 4).

approximations. These limitations were revealed not just by Gödel's incompleteness theorem but also by Turing's resolution to the question of decidability of mathematical propositions. Turing was also responsible for providing the first theoretical proof that the computer was potentially the mathematical 'telescope' that would enable mathematicians to explore their vastly expanded realm.

Turing's work marked a watershed in the recent history of mathematics. His design of a general-purpose computing machine – while only a mind experiment - is considered the inspiration for the modern computer (von Neumann transformed Turing's vision into the practical architecture that underpins the design of nearly all present computers). Using his imaginary machine, Turing showed that a simple computer could perform any mathematical calculation. However, he then asked whether there were any mathematical calculations which, when fed into his machine, would result in a series of operations that would not come to a halt. Turing realized that his halting problem was equivalent to Hilbert's decidability question, and proved that no halting algorithm exists. One of the implications of this discovery was that any mathematical system as sophisticated as a Turing machine, which would include any axiomatic system of the sought that Whitehead and Russell built their edifice upon, could generate computations that cannot be solved.

The discoveries of Gödel and Turing marked the dawn of a new era, a new age of empirical discovery. Simultaneously they signalled an infinite mathematical universe, and at the same time, heralded the discovery of a new and powerful instrument for exploring this universe – Turing's universal computer. Though the philosophical implications of this change are still being debated, in practical terms Turing's work signalled the arrival of the computer as a tool – not just for calculation – but also for mathematical exploration. For mathematicians the computer serves as a telescope which peers into the infinite universe of computable entities, a machine for empirically exploring the realm of the possible; from this empirical standpoint, mathematics is the science which makes falsifiable predictions about the phenomena seen in that world.[12]

---

[12] The tremendous achievements that had inspired Hilbert's vision of a new science (especially Whitehead and Russell's unification of the former disciplines of arithmetic, geometry and algebra in an axiomatically grounded logical science of relations, and sets of relations), created the basis for a more profound revolution in which experimentation rather than rational deduction became the hallmark of mathematical inquiry. While this new era of computer-based exploration is well under way, mathematics is still largely conducted in the language (set theory) and based upon the assumptions (such as Hilbert's formalism) of an earlier era. However, there are rumblings within the hallowed halls of the mathematical discipline that suggest a deeper change of perspective. Gödel's and Turing's discoveries, and the emerging computational era that they signal, are beginning to reshape not just the practice of mathematics but the philosophies that frame that practice. Exemplifying this new era, Stephen Wolfram (2002), a pioneer in the investigation of artificial cellular automata, has laid out a vision of a new science that combines a computational view of logic, mathematics and the universe, and a corresponding research strategy based upon experimentation with cellular automata. Gregory Chaitin (2005), a prominent mathematician who has reflected upon the implications of the new computational era, has proposed a digital philosophy of mathematics that builds upon the work of Gödel and Turing. Chaitin, like Wolfram, advocates an experimental, semi-empirical approach to mathematics, and has argued that mathematics is more akin to physics, in its experimental, creative exploration of an open universe, than Hilbert's ideal of a closed mathematical universe, formulated in a static theory of everything. Indeed, for Chaitin (2007), the Physical Universe itself is the hardware of a vast information processing system, subject to the limitations that Gödel and Turing exposed in their pioneering work on mathematical systems.

Similar frustrations and unanticipated disclosures attended the efforts of those who sought to use the new techniques of symbolic logic to tame the untidy world of natural language. Inspired by *Principia Mathematica*, and sharing Hilbert's vision of a logically complete mathematical system, some linguistically oriented philosophers (following the lead of the young Wittgenstein) dreamed of creating a complete and perfectly rigorous variant of natural language. These logical positivists imagined that all the valid propositions that can be framed in natural language could be expressed in formal logic. However, despite some measure of success, this approach to realizing Lull's vision of automated reasoning was doomed to failure. Yet it was a failure that was instructive, as it has forced computer scientists and cognitive scientists to reflect more deeply upon the nature of language, representation, meaning and human and machine 'intelligence'.

As the next section will show, attempts to develop artificial intelligence were ultimately constrained by the philosophical vision that had mediated the naturalization of Reason, and transformed logic into a mechanical calculus. Just as the Newtonian paradigm was rooted in an atomism that clouded the vision of later generations of physicists, so too Symbolic Logic (and set theory)[13] was rooted in a vision of the mind that both constrained attempts to develop realistic models of human intelligence and inhibited our understanding of the abstract realms that both mind and machine were beginning to explore.

1.2 The impact of philosophical vision and engineering imperatives on the development of Artificial Intelligence

Two imperatives influenced the early development of AI. On the one hand, an engineering imperative, related to the difficulty of 'programming' machines to carry out tasks, inspired some researchers to attempt to 'engineer' human intelligence into the machine, so as to ease the task of communicating instructions to the machine; on the other hand, the philosophical horizon that had informed the development of modern symbolic logic and the first electronic computers led early researchers to view early computing machines as artificial 'minds'. Both impulses drew inspiration from notions of knowledge, representation and the operations of the mind and played a part in shaping the early development of AI. Hence, I shall begin by noting the role that philosophy played in framing these early developments, describe the role of the engineering imperative, and then comment on the conundrums that arose as engineers attempted to realize their vision of artificial intelligence.

The early history of Artificial Intelligence is dominated by two mechanical models of mind, with one model inspired by linguistic analysis and the other by biology. Both shared Boole's assumption that the "laws of thought" could be formulated in mathematical terms, and drew inspiration from the intellectual legacy of the previous century. The school inspired by linguistic analysis assumed that intelligent decisions flowed from the logical implications built into sets of beliefs about the properties of the world, values that the agent attaches to those properties and heuristic "rules of thumb" for formulating goals given a set of beliefs and values. The school inspired by behaviourist theories of human biology assumed that intelligent action was the result of conditioning processes mediated by the brain's neural architecture, whereby the

[13] See previous note [fn.8] on the limitations of set theory.

reinforcement of "intelligent" behaviours was attributed to the environmentally conditioned formation of neural connections. For convenience I will refer to these two models as "representationalism" and "connectionism". Both have their roots in scientific theories of the mind that can be traced back to the philosophy of Thomas Hobbes.

Hobbes was a seventeenth century English philosopher, and while best known for his political philosophy, he has also been referred to as the "grandfather of Artificial Intelligence" (Haugeland,1985:23). Ironically, Hobbes's goal was to underpin a naturalistic explanation of society (the "Leviathan"), not to build an intelligent machine, nevertheless in developing his materialistic explanation of the workings of the mind he did articulate a vision that incorporated many of the presuppositions of AI. He argued that the mind is the product of organized matter, and life is a function of this organized matter in motion. Mind could be interpreted as a process whereby the motion of matter in the external world was picked up by the physical sense organs, and that these organs mediated a corresponding organization of brain matter, by associating particular appearances or representations with particular motions of cerebral matter. These specialized "parcels" of matter were in turn combined and disbanded in a manner analogous to the operations of arithmetic. Hobbes argued that Reasoning, or "reckoning", is simply a reference to this process of physical computation, and that the subject experiences this manipulation of thought parcels as a logically necessary chain of thoughts.[14]

Hobbes's mechanistic model of mind was further developed in two distinct ways. On the one hand, his work can be viewed as an early expression of a tradition associated with the work of the founder of the associationist school of psychology, David Hartley. Hartley's speculations concerning the actual anatomical basis of associative learning culminated in the work of the father of modern neurological models of the brain, Alexander Bain. Bain's associationist explanation for the formation of neural links was to be embodied in the *connectionist* vision of Artificial Neural Networks (ANNs).[15] Hobbes can also be regarded as an ancestor of the r*epresentationalist* approach, in that his explanations can be viewed as a precursor of the *symbolic theory of mind* underlying representationalism. The father of modern computer science, Alan Turing, articulated a hypothetical model of a universal computational machine, capable of computing any algorithm, and Newel and Simon (1976) transformed this 'physical symbol system' into a formal definition of the attributes necessary for the creation of an intelligent machine. Their symbolic theory of mind echoes many of the themes outlined by Hobbes.[16]

---

[14] Hobbes' explanations can be found in his classic text *Leviathan,* published in 1651. For references to relevant passages, see Haugeland (1985:22-28) and Lambert (1996:12-14).
[15] A review of this history is presented in Medler (1998).
[16] The philosophical link between Hobbes and modern symbolic theories of mind is real but indirect, as it proceeded via a long philosophical controversy over unresolved tensions in Hobbes's philosophy. This controversy (empiricist disputes over the nature of the relationship between ideas and the external world) led via Hume to Kant's synthesis of the rationalist and empiricist traditions. The new Idealist schools of philosophy in turn inspired various empiricist reactions, culminating in logical positivism and the development of analytic philosophy. The analytic philosophers' combination of formalism and linguistic analysis is closely related to a linguistic solution to the problem of the relationship between ideas and "reality" (that thought should be identified with language, and notions of sense, reference and truth derived from analysis of logical patterns implicit in language usage). The relationship between

The more immediate philosophical inspiration for the representationalist approach to AI was the attempt by analytic philosophers to formalize essential elements of natural language usage, and their related assumption that natural languages correspond in certain essential ways with the real properties of the world. From this standpoint, sentences are symbolic vehicles for evoking propositions (the meaning content of the sentence), and true propositions are justifiable beliefs concerning the nature of the world. Accordingly, the mind can be viewed as a deductive machine, in which a store of valid propositions, together with laws covering how those propositions can be combined, enables various implications or conclusions to be drawn. Connectionist or neural network approaches to AI harboured similar visions of reducing natural language to a logical calculus. McCulloch and Pitts, whose (1943) paper "A Logical Calculus of the Ideas Immanent in Nervous Activity" initiated the modern era of neural networks, argued that neurons could be viewed as simple logic devices. "Logic", they argued, "is the proper discipline with which to understand the brain and mental activity".[17] Hence, connectionist approaches to AI drew upon the same Hobbesian vision of the mind as that informing the work of Newell and Simon, and their work provided the first detailed justification for Turing's thesis that the brain was a 'digital computing machine' that could be modelled using Boolean algebra.

1.21 Communicating with machines: from FORTRAN to AI

While AI can be viewed as an attempt to simulate the capacities of human intelligence, the historical context of the development of AI is much more closely aligned with the pragmatic goal of developing more useful computational tools. From the standpoint of this engineering imperative, AI can be viewed as a natural development of the desire to create user-friendly programming languages. That is, underlying the practical usefulness of the electronic, digital computer is the ease with which people can interact with them. The physical processes that mediate a modern computer are, in engineering terms, very complex. While the physical means for communicating our intentions to a bicycle are built into the basic form of the machine, no such opacity underpins the design of a computer.[18] Computers are engineered to take advantage of the binary properties of electronic switching devices. The central processing unit utilizes these devices to store instructions, memorize data, and execute instructions. In order for the computer to "understand" a program it must be written in a binary code that physically triggers the appropriate output. This task was accomplished in the

---

empiricist philosophy and analytic philosophy is discussed by Fred Wilson (1997) in his review of Dummett's *Origins of Analytic Philosophy* (1994).

[17] McCulloch and Pitt's (1943) theory employed Turing's mathematical notion of computation and envisaged "information flowing through ranks of neurons". Thus the brain "would embody a logical calculus like that of Whitehead and Russell's *Principia Mathematica*, which would account for how humans could perceive objects on the basis of sensory signals and how humans could do mathematics and abstract thinking" (Piccinini, 2004, p.179).

[18] That is, the human body's a priori dispositions are embedded in the design of the machine's hardware; when seated on a bike the body faces forward in the direction of travel, hands are coupled with handlebars and used for steering, while the legs natural propulsion movements are coupled with peddles and used for propelling the machine forward. Computer hardware lacks this functional synergy between a human intention and the physical means for achieving that intention. Software systems attempt to compensate for this lack of synergy by incorporating 'user-friendly' interfaces into their design; hardware peripherals facilitate input and output of information, with devices, such as the 'mouse' and touch computer screens designed to support the more intuitive feel of these software systems.

earliest machines by using cards with holes punched in them. However, there was no intuitive link between the machine instructions coded in these cards (or their successor technologies) and the symbols that we use to express our thoughts to each other. The challenge was to bridge the gulf between the human and machine languages.

The primary way in which engineers have attempted to facilitate communication between humans and machines has been to create programs that automatically translate (“compile”) more intuitively meaningful symbols into the assembly languages used to encode the binary machine language. The first such compiler program was FORTRAN (FORmula TRANslator), made in 1953, which (following the lead of Von Neumann) viewed computers as logical, rather than mechanical devices, and expressed the programming instructions in the symbols of formal logic. This marked the beginning of a progression toward higher-level programming languages endowed with graphical and speech interfacing. Extrapolating the principle of creating user-friendly machines, requiring no technical expertise to program, leads to an image of machines capable of responding to human instructions with the same ease as other humans. This vision of software that facilitates natural language communication between humans and machines converged with early efforts to engineer artificial intelligence.

Alan Newell (1982) has argued that the endeavour to bridge the communication gulf between human and machine has involved conceptualizing machines at three different levels. At the *Machine Level* a computer is understood in terms of the principles of a physical system. At this level, in accordance with the ideal of a universal Turing machine, a computer should have no specific function other than deterministically and automatically computing any computable algorithm. At the *Program Level* its operations are reinterpreted as those of a logical system. Rather than think in terms of physical events taking place in the system, the user thinks in symbolic languages analogous to formal logic. All modern programming languages, from Fortran to Python, mediate interaction at this level. And at the level of natural language communication – were this to be realized – interaction with the computer would be taking place at the *Knowledge Level*. This Knowledge Level, Newell contends, involves adding a new capacity to the machine – artificial intelligence (AI). AI is necessary if the Knowledge Level is to realize its goal of constructing machines that operate in ways analogous to human ‘knowing, believing and saying’. That is, at the Knowledge Level, the machine should be capable of understanding humans as agents with a set of beliefs, goals and actions, and of being understood by a human in a similar fashion. Thus, driven by the purely utilitarian demand of bridging the communication gulf between human and computer there has been a natural progression that led Newell to the vision of artificial "agents", endowed with a capacity to interact in an intelligent fashion with human agents.

1.22 Classical AI and early attempts to realize Newell's vision of intelligent agents

According to Allen Newell and Herbert Simon (1976), computers are physical symbol systems, functionally analogous to the human mind. The notion of a physical symbol system has its roots in the work of Turing and Von Neumann, and such a system can be viewed as a set of arbitrary symbols (physical tokens) that are manipulated on the basis of explicit rules. As the symbol system is a rule driven calculus, analogous to a mathematical system, output is determined syntactically, however by adopting the approach suggested by Tarski, syntax can be assigned a meaning, and is hence semantically interpretable (see above, fn.7).

Newel and Simon argued that the physical symbol system provided an adequate description of both human and machine intelligence. Their representationalist approach to intelligence underpinned the early synergy between classical artificial intelligence and cognitive science (as exemplified in GOFAI 'Good Old Fashioned Artificial Intelligence' and CRUM 'Computational-Representational Understanding of Mind'). Paul Thagard (2005:11) suggests that this model of mind rests upon the following analogy between computer programs and mind:

- **Program** (data structure + algorithms = running programs)
- **Mind** (mental representations + computational procedures = thinking).

Newell was one of the pioneers in the endeavour to construct artificially intelligent machines. Beginning in the fifties, through to the early nineties, research in AI was inspired chiefly by the computational-representational view of mind, but attempts to engineer AI met with limited success. Before turning to more recent developments, it is instructive to consider the strategies employed by AI engineers during these formative decades, and the problems that were encountered. The first programs (in the fifties), exemplified by Newell, Simon and Shaw's "General Problem Solver", attempted to solve problems by using general heuristic principles[19] to search a data base made up of declarative propositions (expressing true facts about the world). The limited success of this model inspired researchers in the early sixties to shift their focus from heuristics to representation. Their goal was to develop representation schemes that not only specified facts about the world, but also organized that knowledge into domain specific categories that implicitly specified how that knowledge could be used. This led AI practitioners away from their initial preoccupation with logics into the field of knowledge representation schemes. A further development occurred in the late seventies, with a shift of emphasis from the form of knowledge representation to the specific information content which such systems were trying to capture. That is, the focus moved out from the internal functioning of the machine to the external world and the human communities of problem solvers that the machine was attempting to emulate. The goal was to formalize the knowledge of human experts into highly specialized "expert systems". For example, in such a system, a medical "agent" would be programmed to diagnose symptoms of medical conditions. Such systems met limited success due to the "brittleness problem".

The *brittleness problem* refers to the inability of expert systems to respond to unexpected situations or communication difficulties with the flexibility and initiative that a human typically demonstrates. Solving the brittleness problem would be tantamount to building a computer that was capable of exercising "common sense". Following the classical, representationalist approach to AI, some computer scientists have attempted to solve the brittleness problem by encoding common sense within an enormous knowledge base. The most ambitious of these projects, "Cyc" (En-Cyc-

[19] 'Heuristic principles', in computer science, refer to a rule of thumb, strategy, trick or any other method that facilitates the search for solutions in a large problem space. Effective heuristic strategies provide "solutions which are good enough most of the time" (Feigenbaum and Feldman, 1963: 6).

lopedia), founded by Doug Lenart in 1984, still hopes to realize its goal by encoding all commonsense in its data base; to that end, users of the World Wide Web are invited to contribute their knowledge in an ambitious attempt to formulate all common sense understandings in declarative semantics.[20] The Cyc approach assumes that commonsense knowledge is made up of propositions and implicit theories that represent our understanding of particular problem domains. These various domains can be formalised in a general ontology made up of appropriately conceptualised sub-ontologies each with its implicit formal theory.[21] By representing all commonsense knowledge in this fashion, Cyc will be able to draw upon the understandings implicit in common sense knowledge concerning the nature of the world and automatically arrive at valid inferences akin to those generated by commonsense. Today, few people outside of the Cycorp community have much confidence that representing all human knowledge within a large formal ontology will produce a machine capable of exercising something approximating common sense. Most researchers have come to the conclusion that human common sense rests upon very different foundations.

1.23 Classical AI and the Symbol-Grounding Problem

The classical 'GOFAI' approach to engineering an agent with a human-like capacity for communicating with people or interacting with a dynamic physical environment was based upon a symbolic model of mind in which the 'intelligence' was entirely a function of internal processes. Newel and Simon's physical symbol system assumed that intelligence was a magic 'black box' that, though linked to the external world via inputs and outputs, functioned as a closed, solipsistic realm of meaning. That is, internal representations and the world that they represent were treated as two distinct entities. Intelligent activity was explained in terms of structures and processes internal to the machine, rather than in terms of the agent's intrinsic connection with structures and processes constitutive of its environment. Hence, the semantic content of a representation was entirely a function of the internal structure of the agent's program. How then should the agent's internal representations be mapped on to the external world that they represent? This 'symbol-grounding problem' is at the heart of difficulties with the classical AI approach to representation.

Classical AI, following a cognitivist strategy akin to the model suggested by Hobbes, sought to solve the symbol-grounding problem by linking the atomic tokens of its internal representation system to simple objects of the external environment. From this standpoint, the external world picked up by an agent's sensors is compounded of simple, invariant features; the solution to the grounding problem is to develop sensors that pick up the relevant simple external features, and map them using the atomic tokens mediating the agent's bottom-level representation of the world. The agent processes this representation and grounds successively more complex layers of

---

[20] See the Cyc homepage at http://www.cyc.com/

[21] While the meaning of ontology within computer science has some affinities with its philosophical usage there are also significant differences. In philosophy ontology is the core subject of metaphysics, and concerns the study of being or existence and the framing of categories or concepts useful for describing the ultimate nature of our reality. Ontology, as used in computer science and information sciences, has a more restricted meaning. It is a set of concepts and relations, specified in a formal theory, which describes how data can be structured and used in a particular domain. It thus forms the basis for classifying and reasoning about objects in its domain. See Tom Gruber's definition (Gruber 2008).

representation in the primary substratum of atomic tokens. In practice, this strategy proved to be difficult to scale up beyond simple 'toy worlds'. Moreover, the model itself is rooted in questionable psychological and philosophical presuppositions concerning the nature of representation and its role in human intelligence. As John Searle (1980) has famously argued, the analogy between human and machine representation are tenuous.[22] One radical proposal for dealing with these difficulties was to rid AI of representation altogether.

1.24 Representation and Alternative AI

The major challenge to the classical GOFAI paradigm has come from Connectionist, Embodied and Situated approaches to AI. I have already referred to the Connectionist paradigm, and its roots in the atomistic view of the brain. This perspective is based on the notion that intelligence is the product of the brain's architecture, and that its massive parallel processes of neural interaction can be modelled mathematically, using artificial neural networks (ANNs). While Classical AI represents through symbols and computes by symbolic manipulation, Connectionist AI represents by associating numerical weights between distributed processes, and computes outcomes by numerically calculating weight matrix products. Such networks can be viewed from within the classical view of mind propagated by Hobbes and viewed as a kind of representation, or adapted to philosophical standpoints critical of the symbolic representationalist view of mind.[23]

Philosophically, the more radical departures from the Classical paradigm are the embodied and situated approaches to AI. These approaches have drawn inspiration from various philosophical and scientific traditions that have rejected the representationalist view of mind. One of the early critics of Classical AI, Hubert Dreyfus (1972, 1992), drew upon the work of phenomenologist, Martin Heidegger, to refute the Classical AI notion that the human capacity to interact intelligently with the world is mediated by representations. Dreyfus argued that the bulk of our everyday activities are performed without representation. Instead of attending to images of the world, we interact directly with the world by drawing upon embodied skills that are 'tuned' to tasks particular to our environmental situation. During the nineties the views of Dreyfus, reinforced by diverse sources of inspiration, such as the psychologist James Gibson (1979), and anthropologist, Lucy Suchman (1987), inspired various members of the AI research community to break with the Classical AI paradigm and develop what is sometimes referred to as 'Alternative AI' (Sengers, 1999).

Using notions such as situatedness, interaction, embeddedness and embodiment, practitioners of Alternative AI highlight the dynamic structural coupling between the

[22] The general import of Searle's mind experiment is illustrated by computer translation systems that use algorithms to spit out "translations" with no cognisance of whether the semantic content of their translation is utter nonsense. Searle's general point is that computers are syntax driven systems that involve no meaning or intentionality. See Searle (1980).

[23] Lambert (1996) discusses a range of approaches to Connectionism, some that adopt the Classical AI approach to representation and some which reject the Classical position (1996:2-4). Amongst these alternate interpretations of the connectionist strategy is the Whitehead inspired work of David Primeaux (2000); Primeaux and Saunders (2006); also Saunders (2008). Thagaard (2005) introduces connectionist theory and in (2008) discusses the epistemological relationship between representation and neurological activity.

agent and the environment. This emphasis is evident in the work of Rodney Brooks, whose extreme anti-representationalist solution to the problem of intelligent action greatly inspired the early alternative AI movement. By developing what he termed ‘subsumption architecture’, Brooks (1990) demonstrated that his biologically inspired robots could be designed to carry out various tasks without utilizing the representationalist approach to planning and acting. He distinguished between ‘symbolically-grounded’ and ‘physically grounded’ agents, arguing that while the former respond to a symbolically coded internal model of their environment, and use that model to plan actions that will achieve their programmed aims, the latter react directly to the environment itself without explicitly representing the form of their external world in their program code.

Philip Agre and David Chapman (1990) drew upon sociological sources of inspiration to give a new twist to Classical techniques, as expressed in their ‘Interactionist’ approach to Situated AI. While recognizing the importance of representation and planning, they differentiate their own Interactionist-Situated agents from the typical Classical agent. The latter internally represents its environment and engages in abstract planning activity before undertaking action, while for Agre and Chapman, plans are a resource that situated agents use flexibly as they engage in moment-by-moment improvised activity – activity that depends upon the unfolding environmental situation. According to Agre and Chapman, plans are not necessary for human’s to engage in sensible action, but when we do draw upon them, they are guides rather than fully determinative instructions: more like cooking recipes than computer programs. Hence, they envisage a form of AI in which plans are “guides to activity, not solutions to problems” (Agre and Chapman 1990:15).

The Alternative AI focus on the structured mode of interaction between agent and environment shifts the focus away from the role of internal structure toward that of external structures. Making artificial intelligence dependent upon the mode of dynamic coupling between an agent and its environment has the effect of situating the resources for intelligent action in the environment rather than inside the agent’s memory. This coupling of intelligent action to the structure of the external environment is suggested by the slogan (courtesy of Brooks) that ‘the world is its own best model’. The world as ‘data base’ is always exactly up to date and complete in every detail. Computer scientists acknowledge that they have neglected the role of environments in multi-agent systems (Weyns 2005a). Agre and Horswill noted this neglect in 1997, and developed a powerful argument for focusing on the role of the environment that foreshadows many of my concerns. The goal of Agre and Horswill’s work was to develop “a theory of intentionality that depends on the agent's embedding in its world, rather than solely upon its internal models of that world” (Agre and Horswill, 1997:138).

Agre and Horswill’s (1997) work reflected a shift away from psychologically and biologically inspired models of intelligent activity toward sociologically informed notions of commonsense. However, it also highlights a key issue for the situated approach; namely, how to incorporate the role of representation and symbolic languages into its model of AI while avoiding the problems that afflicted classical AI? The importance of this issue has been further accentuated by the growing population of robots, web-bots and avatars, functioning in both the real and virtual worlds; how should they communicate with each other, and how should machine agents communicate with human agents? Questions such as these have motivated the development of Socially Situated AI.

1.3 Multi-agent Systems and Social Interaction

Since the mid-nineties, the study of multi-agent systems (MAS) has become recognized as an important sub-field in computer science.[24] Multi-agent systems are, as the name suggests, composed of multiple 'agents'. These agents are autonomous computing elements that vary from simple reactive agents, akin to ants blindly following chemical trails, to sophisticated entities designed to engage in activities analogous to rational negotiation between humans. Modern networked computer systems have created a challenging environment for artificial agents. In order to manage communications with other computers in a networked environment, such as the Internet, computer agents must not simply exchange data with each other; rather, they must negotiate, coordinate, cooperate and perhaps compete with other agents, in order to fulfil their design objectives. As a result multi-agent design is to a large extent preoccupied with issues that are characteristic of human sociality. As Michael Woolridge notes in his popular introduction to the subject, "Multiagent systems seem to be a natural metaphor for understanding and building a wide range of what we might crudely call *artificial social systems*." (Woolridge, 2002: xi)

The notion of a computational agent was largely inspired by analogies with human cognition; however, the need to integrate multiple autonomous agents into a single system raised issues not addressed by theories of cognition – problems analogous to those that arise in social organizations. For example, how should the autonomous agents of a multi-agent system integrate their knowledge so as to achieve a common goal? Communication is obviously a key element, but what type of communication? Informing, questioning, negotiating, bargaining and voting are all possible approaches to facilitating cooperation and the sharing of knowledge, but these strategies raise further questions, such when, how and with whom the agent cooperates? Similarly, how do multiple agents generate collaborative plans? When, for example, might an agent obstruct, ignore, or assist another agent’s plans? These questions led some AI engineers to regard human social interaction as a potential source of solutions that could be mined for strategies for resolving similar problems in MAS-based societies.

The orthodox, classically inspired methodology for creating agents with the capacity to communicate intelligently with other agents, as outlined by Barry Smith (2003a), involves the SMPA view of human action:

S: the agent first *senses* its environment through sensors
M: it then uses this data to build a *model* of the world
P: it then produces a *plan* to achieve goals

[24] Reviewing the twentieth century endeavour to build intelligent machines, Huntbach and Ringwood note a number of factors that led to the emergence of the multi-agent approach to programming. They suggest that the social turn became inevitable as computer science moved away from the notion of a computer as a stand-alone system, "used to model aspects of the world", towards computers as nodes in networked systems, hosting software "agents" that functioned as "active participants in the world"; for "once agents become truly autonomous, the ease with which discussions on the organization of multi-agent systems slip into analogies with the organization of human societies becomes irresistible." (1991: 319, 330)

A: it then *acts* on this plan.[25]
This paradigm was based on the representationalist approach to AI and assumed (in contrast to the embodied-environmentally "situated" approaches to AI) that "intelligence" is a capacity engineered into a program that is prior to the manipulation of inputs and outputs.

Adopting a position of methodological solipsism, representationalists assumed that the best way to study the mind/brain was to abstract it from its real world environment and treat it as an autonomous Cartesian Ego, operating in a sealed black box. From this representationalist standpoint the primary challenge for socially distributed AI is "the communication of one agent's internal state (its beliefs, goals, and willingness to help) to others, and/or the communication of non-shared knowledge about the environment (e.g. the constraints on or goals of the problem as a whole)" (Boden, 2006:1044). In line with the Cartesian assumptions of classical AI, intelligent action is assumed to emanate from the rational processing of information stored in minds. Cooperation, from this standpoint, arises from calculations concerning rational interest, and the chief function of communication is assumed to be the transmission of information from those with information to those that lack information. But does this model capture the essential properties of human social interaction? From the standpoint of a situated approach to social interaction it does not.

The situationalist position assumed that human cognition developed in a real-world environment, and was intimately tied to dynamic interaction with structures in that real world environment. Thinking, from this perspective, is something that takes place as the human organism interacts with its world. That is, thinking is not purely calculations in the head, but involves dynamic interaction with structured environments that steer the agent in a manner analogous to the way that the thermal structure of updrafts steer a soaring bird. In the case of human agents, the physical environment is disclosing socio-cultural structures, manifested in languages, signs, routines and rituals, which provide information and leading clues that steer agent interaction. A great deal of human socialization is about acquiring the set of habitual responses that will enable us to enable us to navigate our way unthinkingly through very complex, but highly structured situations. These socially acquired 'skills' are dependent upon the society's collective maintenance of its carefully structured environments. From this socially situated perspective, our intelligence is a dynamic function of collective, context bound socio-cultural processes rather than calculations performed by autonomous 'minds'.[26] One implication of this non-calculative approach to human interaction is that most classically inspired AI work on communication and negotiation between 'social' agents is not truly social. If AI agents are conceptualized primarily as solipsistic individuals, and their architecture does not take into account the principle that social behaviour "may depend largely on *the dynamical properties of the interaction as such*", then essential characteristics of society will be over looked (Boden, 2006:1045).

---

[25] Smith (2003) draws upon Rodney Brooks' (1991b; 1991c) summaries of the representationalist and situated approaches to agents.
[26] For example, Edwin Hutchins (1995) argues that computational and cognitive properties of human systems transcend the individual, and are better viewed as collectively constituted, context bound, artefact dependent properties of socio-cultural organization. He demonstrates his thesis with an anthropologically inspired study of ship-board navigational procedures.

In contrast to agent centred approaches, socially situated AI focuses on the socio-cultural context of the machine-agents, and emphasises the dynamic properties of social interaction. In fact, the dynamic, 'on-the-fly' creation and maintenance of shared communicative contexts are a key characteristic of human communications (Devlin, 1997:272-3). Hence, socially situated AI goes beyond Brook's concern to explore the potential of situated, reactive agents that lack representation and reason; similarly, it goes beyond Classical AI's focus on the capacity of individual computer agents to process information and generate meaningful communications; it shifts its focus to the dynamic relationship between machine-agent architecture and human modes of interaction. Phoebe Sengers' pictorial representation of the distinction between classical, alternative, and socially situated AI is reproduced in figure 1.

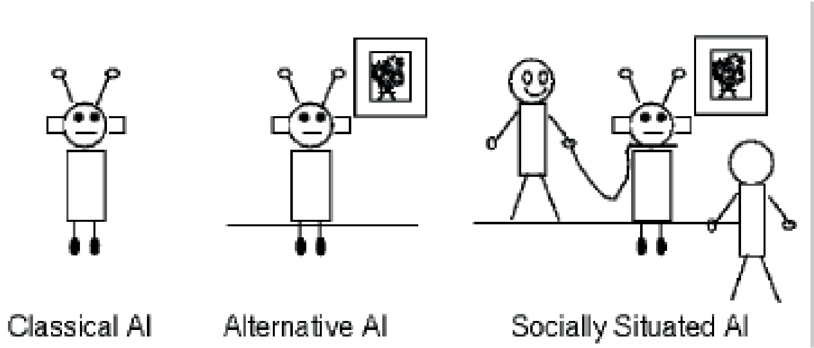


Figure 1. The Environment and AI (© Sengers,2002:427)

1.31 The Socionic quest for truly social form of AI.

One discipline that has struggled against the tendency to reduce the characteristics of a society to the individual qualities of its citizens is Sociology. Neglected by Classical AI, sociology had typically focused on the emergent socio-cultural properties of enduring social collectives. From the perspective of sociology, societies exhibit unique 'structures' that organize and guide the action of the members of a society (just as the grammatical structure of a language shapes the way we speak). As in the case of linguistic structures, the structures that characterize social institutions, such as those defining economic, political or family life, are discerned indirectly through patterns of human activity. Over a century of research has produced a rich vein of empirical data and related theories. Some computer engineers have attempted to mine this body of knowledge. Guided by the metaphor of social action, they have turned to sociology in the hope of finding leading clues as to how to design 'socially' distributed forms of artificial intelligence. Likewise, a number of sociologists embraced the emerging MAS technology as a tool for testing theories of society. One of the first attempts to transform this two-way traffic between computer science and sociology into a unified program of research was the Socionics project developed in the mid-nineteen nineties by Thomas Malsch.[27]

[27] The term Socionic (German: Sozionik) was first used in this context by Thomas Malsch: see (Malsch et al. 1996, Malsch 1997, Malsch 1998a, Müller et al. 1998). As noted in the introduction, the term was coined and used in a more general, but loosely related sense by Nicolas Rashevsky (1966; 1967). The term has also been used to describe an approach to personality analysis, developed by Aušra Augustinavičiūtė in the 1970s; however, the latter work has no relationship to the work of Rashevsky, or connection to Malsch's MAS inspired use of the term.

Responding to his conviction that the direction of MAS based research had created a situation ripe for “interdisciplinary cross-fertilization”, Müller, Malsch and Schulz-Schaeffer came up with a vision of a new discipline based upon a hybrid combination of computer science and sociology. Following Malsch's suggestion, they labelled the new science 'Socionics' - and anticipated that it would emerge out of "a painful, mutual exchange and translation process between the two disciplines [of computer science and sociology]" as professionals from both disciplines engaged in a dialogue oriented toward building "highly complex and highly dynamical multi-agent systems" using "social systems as a paradigm" (Müller et.al.,1998). Thus:

> Socionics aims at working out the advantages of *collectively organized problem solving processes* in order to come up with new forms of generating solutions which find their counterparts in the many facets of our social life and which have not been used significantly in technical contexts up to now. [my italics] (*ibid*)

Malsch's notion of 'collectively organized problem solving processes' is fundamental to the Socionic enterprise. It defines purposeful activity in a manner that allows useful comparisons to be made between MAS based artificial societies and human societies. It also identifies the practical technological basis of the Socionic methodology; it is rooted in the functional capacity of multiagent systems to simulate the collectively organized problem solving processes characteristic of human society. The strength of the Socionic methodology rests upon the degree to which these virtual societies exhibit characteristics that are analogous to human society. The capacity of MAS tools to simulate human society rests upon analogies between the underlying mechanisms generating the patterns of behaviour exhibited by both systems. Thus Socionics joins together theoretical reflection about the nature of society with the building and running of computational models of society; both are essential elements in the socionic vision of a methodology that combines social theory, formal modelling, computational implementation, and simulation-based social experimentation.

1.32 The domain of Socionic research: hybrid societies

In order to illustrate the envisaged domain of Socionic research, Müller and Malsch provided the following example:

> Let's consider an artificial world of Personal Electronic Assistants (PEAs) to exemplify the statement above [on 'collectively organized problem solving']. Suppose every person from child to seniors would be the owner of a team of electronic assistants. Every PEA would be associated with a social context of its owner, e.g. school, work, recreation, family et cetera. Further every PEA would know certain goals of its owner with respect to the specified social context. The aim of each PEA would be to support its owner to reach these specific goals by gathering information, preparing activities, and providing plans. Comparable with its human owner in the human society, the electronic assistant would be part of an electronic society and as such part of the electronic "counter world". (Müller, Malsch, et.al.,1998)

Noting that the process of building these parallel-worlds of electronic societies was well under way, Malsch suggested that a primary task of Socionics would concern the nature of these electronic worlds. Basic questions for Socionics might include:

- What types of societies would emerge? How would the electronic communities be created?
- Do PEAs have an explicit notion of what a society might be? How would the social integration-process work with PEAs?
- What would be the differences between the real society and the electronic society, especially in social contexts? What might be the reasons for the differences? (*ibid*)

Malsch then suggested that Socionics would be interested in going one further step. Given that the real world and the electronic world are connected through the human owners and their electronic assistants, there would also be various kinds of connections and interactions between the PEAs and humans during the problem solving processes. These interactions would be the basis of emergent forms of hybrid society. This, Malsch argued, raises new questions for Socionics. Specifically:

- What forms of hybrid societies would be reasonable?
- What kinds of human-machine interactions would be necessary to join hybrid communities?
- How could real persons and electronic assistants learn about the organizational structures of the corresponding hybrid societies? (*ibid*)

Thus, Socionics, as envisaged by Malsch, combines computer science and sociology in order to develop MAS based 'social' software systems, and utilizes these artificial social systems to develop hybrid societies made up of both human and machine agents. This implies both a narrow technical focus on the construction of autonomous 'virtual societies', and a broader practical analysis of the role that such technology can play in the constitution of hybrid societies. As a 'narrow' science oriented toward engineering advanced MAS based societies, Socionics is an exercise in sociologically inspired software engineering, with the aim of building artificial societies made up of agent-like software entities that communicate, negotiate and act collectively to solve problems. In its broader role, as a science of hybrid societies, Socionics is oriented toward integrating sophisticated MAS based software with human society, using sociological theory and research strategies to analyse and explain the resultant emergent forms of hybrid sociality. In its computational phase, the human agent formalizing social theory and implementing it in MAS software is not explicitly incorporated into the research frame; however, in the applied social phase, the focus shifts to the ways in which humans interact with the computational system. This means that Socionics is concerned with the practical social, political and economic consequences of fusing together human and MAS based societies. Hence, as Malsch states, the practical objective of Socionics is to create ethical societies of agents that contribute not only to our society's economic goals, but also help further the realization of our society's democratic ideals (Malsch, 2001: 23).

As a discipline, Socionics is intended to contribute both to the development of computational theories that shape the architecture of MAS software, and to the sociological theories underpinning our understanding of society. Malsch summed it up this way. "*Socionics*", he said, "*aims to form a new research discipline with the aim of developing intelligent computer technologies by picking up paradigms of our social world. And vice versa, Socionics uses computer technology in order to verify and to develop sociological models of societies and organizations.*" These two arms of Socionic research are intended to feed into and learn from the practical application of these technologies in 'hybrid societies'. Thus, Socionics *"aims at generating answers in the context of (a) the emergence and dynamics of artificial social systems and (b) hybrid man-machine societies."* (Müller, Malsch, et.al.,1998)

A decade has now passed since Malsch began developing his notion of Socionics. The German government has funded a Socionics research program, involving collaborative teams of sociologists and computer scientists working on a number of different projects.[28] This program has generated numerous publications, including several books and a multitude of articles. Articles in Fischer (et.al. 2005) and a 2007 special Socionics feature edition of the principle social simulation journal, *JASSS,* outline the achievements of the Socionics program.[29] During the past decade computational research into multiagent systems has boomed and the sociological application of MAS based social simulation techniques has gained increasing recognition. Yet despite the growing interest in these two forms of research, the Socionics vision of bringing these two research streams together has not excited a great deal of enthusiasm.[30]

Socionics, by engaging with Sociology, is embracing different notions of human agency than those which have dominated the entwined development of AI and Cognitive Science. Sociologists have, for the most part, rejected the Cartesian assumptions of Classical AI and the related computational models of mind that have dominated cognitive science. Practitioners of AI, on their part, have treated sociology with a disdain generally reserved for philosophy. Yet the contingent processes of history have brought these two strange bedfellows together. This brief review of the history of AI has suggested that both opportunities and challenges confront this unlikely union. Essentially, Socionics is building upon and at the same time challenging AI; equally, it is building upon developments in and challenges to traditional paradigms at the heart of the related discipline of Cognitive Science. A core concern, in both cases, is developing a notion of representation and communication that adopts a genuine situated perspective and breaks away from the

[28] Institutional support for the Socionic program was obtained in 1998 in the form of a significant grant from the German Government. (See Heinz-Jürgen Müller, Thomas Malsch and Ingo Schulz-Schaeffer, 1998, "SOCIONICS: Introduction and Potential", *Journal of Artificial Societies and Social Simulation* Vol. 1(3) http://jasss.soc.surrey.ac.uk/1/3/5.html).
29 K. Fischer, M. Florian, and T. Malsch (Eds.) 2005, *Socionics: Its Contributions to the Scalability of Complex Social Systems*. Berlin: Springer. A list of papers published in the course of the six year DFG grant is available at the project web site: http://www.tu-harburg.de/tbg/Deutsch/SPP/SPP_Literatur_II.htm . See also the special Socionics feature in the *Journal of Artificial Societies and Social Simulation* (Jan. 2007) vol. 10, no. 1: http://jasss.soc.surrey.ac.uk/10/1/11.html.
[30] Reflecting on the development of interest in the Socionic approach, Malsch et. al. commented that when "the first projects were launched in October 1999 no one was actually able to predict the boom which Socionics and related work on artificial agent societies would experience in the coming years", and "that the Socionics program was indeed lucky to catch the wave peak at the right moment and to benefit from surfing on an international scientific development" (Malsch & Schulz-Schaeffer, 2007). While the Socionic program has no doubt benefited from these developments, Malsch's vision of a new scientific discipline is some way from being realized. Although the development of MAS technology and its use as a computational platform for social simulation has 'boomed', the disciplines of sociology and computer science have remained largely estranged, and reflecting this estrangement, very few practitioners from both disciplines have engaged in that "painful, mutual exchange and translation process between the two disciplines" that Malsch believed would under-pin the emergence of the new discipline of Socionics.

Cartesian presuppositions of Classical AI.[31] This is the challenge that Dale Lambert addresses, and is why his proposal for an information fusion system represents a fruitful subject for a socionic dialogue.

To digress briefly and foreshadow a later argument, human society is rooted in the continual modification and reproduction of highly structured environments. Model making is part of this process of environmental modification. Viewed from the context of the situated AI paradigm, computational modelling is part of a long human history of environmental modification, of marking trails, making maps, manufacturing tools, rehearsing (often in play) team skills and strategies, and developing organization roles and routines. Viewed from the situated perspective, these modifications to the environment are all part of the social infrastructure that mediates intelligent action. In Andy Clark's (2004) words, we are Cyborgs, organisms that individually and collectively self-regulate via the integration of artificial and natural systems. Or, in more humanistic language, we are embodied, situated subjects, interactively negotiating and collectively constructing the self-understandings, complementary roles and associated technologies through which the world gives meaning to our lives and, vice versa, through which we give meaning to our world. It is this collective interaction with and skillful manipulation of our environment, far broader in its scope than the narrow category of knowledge oriented activities that philosophers have traditionally focused upon, which is the real foundation of human civilization.[32] This discussion however, is premature. Firstly I must describe Lambert's process approach to AI and examine his suggested solutions for overcoming those problems that have dogged Classical AI's approach to representation and the engineering of commonsense.

1.4 Conclusion

My goal in this chapter has been to outline the historical background to computer science and twentieth century efforts to develop artificial intelligence. In the next chapter I will review Lambert's critique of the Classical AI approach to representation, and then, in subsequent chapters, outline his proposal for solving these problems by grounding representation in process metaphysics. This proposal is the foundation of his Post-Classical approach to AI and has since been implemented in Lambert's work on information fusion. Amplifying the process philosophical presuppositions of Lambert's Post-Classical MAS-based techniques, methodology and vision, will help bridge the conceptual gulf between computer science and sociology, and form the philosophical basis of a Socionic theoretical framework. By drawing out the philosophical presuppositions that are common to both Lambert's work in AI and to

---

[31] In this work, I am suggesting that the Multi-Agent System (MAS) technologies advocated by Socionics may fulfil a similar role as Wolfram's cellular automata in unifying and transforming our notions of logic, mathematics and the social sciences. However, how we view MAS depends upon the philosophical perspective that we adopt. MAS can be viewed, at a metaphysical level, as an attempt to represent the basic processes that constitute our universe. It can be used as a technological vehicle for the imagination to discover and explore new horizons, rather than a mechanism of conquest for realizing the positivist vision of mastery and closure.

[32] A strong case for this position is developed by Barry Allen in *Knowledge and Civilization* (Denver, CO: Westview Press. 2004). Arguably our human environments constitute not just technological prostheses, but humanly constituted structures that give form, content and direction to our emotions and cognition. This thesis is essential to an understanding of the role of religious rituals and belief systems.

process philosophy grounded theories in the sociological tradition, it should be possible to develop a shared language and conceptual framework for dialogue and fruitful collaboration between the disciplines, in accordance with the methodology and goals of Socionics.

This narrative of the historical background of the current efforts to engineer information fusion has of course skipped over significant developments and focused on issues that reflect the goals of this current work. To that end the narrative culminates with a discussion of Socionics, a discipline that has yet to establish itself, yet from my standpoint, has positioned itself at a key junction in the future of AI. It is possible to envisage agents from traditional AI standpoints and miss the revolution in thinking implicit in the sociological standpoint embraced by the Socionic program.

According to the Socionic standpoint, intelligent interaction is not simply a function of capacities built into the minds of individual agents; rather, it is the outcome of dynamic social processes and mediating social structures that give rise to "intelligent" decisions. Just as the radical implications of Maxwell's field theory of electro-magnetic forces demanded the casting out of the old atomistic model of matter, so the Socionics vision involves a break from the individualistic, atomistic, substance-based notion of mind, and the acceptance of a social relational, process-based theory of mind. In order to make this transition, it is first necessary to see how the Enlightenment model of mind entered into the development of AI and shaped the metaphors that have dominated modern understandings of mind. This is the philosophical battleground that exponents of Socionics must enter if they are to free their discipline from the legacy that has prevented both AI and Cognitive Science from appreciating the dynamic social foundations of "common sense". Whitehead and Lambert will be my primary guides in this endeavour.

Chapter 2

Lambert's Critique of Classical AI.

Chapter Two – Section Headings

2.0 Introduction

The first chapter outlined the historic conceptual breakthroughs that transformed the ancient art of reasoning into formal symbolic logic and formal ontology. Theorems and propositions whose truth could once only be established by intuitive, discursive or visual examples could instead be demonstrated by mechanical, deductive systems. Turing's universal computer and the related vision of machine intelligence, with a human-like capacity for making rational deductions on the basis of remembered information, was the culmination of this remarkable series of developments.

Much has since been accomplished in transforming the dream of artificial intelligence into reality. The developments outlined in the previous chapter enable multi-agent systems to draw upon domain specific ontologies to respond to inquiries and carry out tasks; systems exhibit elementary capacities to learn and innovate; computer games employ increasingly sophisticated AI to animate agents and create plans of action within virtual worlds. These developments have been shaped by a combination of the formal logical and mathematical procedures outlined in the previous chapter and a mechanistic vision of mind and its rational functions. This combination of method and vision underpinned the 'Classical' vision of artificial intelligence but, as in the case of Newton's view of the Universe, success also exposed limitations. As noted in the previous chapter, Classical AI has tended to produce 'brittle' systems lacking in the intellectual flexibility that we associate with commonsense.

In this chapter I summarize Dale Lambert's analysis of how Classical AI's approach to representation has impeded the AI quest to engineer commonsense. Lambert's critique of Classical AI is developed in his unpublished dissertation, *Engineering Machines with Commonsense* (1996) [hereafter *EMWC*]. Subtitled *Representation Revisited: An Essay on the Foundations of Artificial Intelligence,* Lambert describes his goal in *EMWC* as that of bringing about a "subtle shift" in the "classical Artificial Intelligence paradigm" (*EMWC*:7). In his critique of Classical AI, Lambert establishes the terms of reference for his alternative Post-Classical, process vision of AI. This Post-Classical paradigm in turn frames Lambert's subsequent work in 'information fusion' (he is the chief architect of Australia's DSTO research effort into information fusion) and underpins my vision of a genuine sociological turn in AI. The aim of this chapter is to lay out Lambert's critique, and prepare the ground for an outline of his Post-Classical reformulation of the AI paradigm.

Lambert's campaign to bring about 'a subtle shift' in the Classical AI paradigm is advanced via a dual strategy; firstly, he identifies deep rooted flaws in the foundations of Classical AI, and secondly, by suggesting viable strategies for addressing these flaws, he develops a rival "Post-Classical paradigm". Lambert's not inconsiderable ambitions are underpinned by detailed philosophical analysis, formal proofs and software implementations in support of his rival paradigm. The resultant thesis is challenging reading. However, his work deserves more attention for several reasons.

Although a plethora of critiques have been directed at the classical AI paradigm, Lambert's work is, to my knowledge, the most exemplary case of a process philosophy inspired response to the shortcomings of the Classical paradigm, combined with an implemented architecture illustrating the principles of a process philosophy inspired, alternative AI.[1] That is, Lambert's early vision of how to resolve

[1] The term process is ubiquitous in engineering, and need not be related to process philosophy. The author has identified very little work in the field of AI with strong links to the modern process philosophy tradition, as exemplified in the work of A. N. Whitehead. An exception is the work in region based ontology carried out by the school associated with Anthony Cohn (Cohn & Varzi, 2003). Cohn's work has been an attempt to develop a logical language for computational systems working with spatial information, in which regions are the basic entities. The inspiration for this work goes back to Whitehead's theory of extensive connection (Gotts, Gooday & Cohn 1996:58). The development of a region-based formal ontology, while not a pivotal concern, does play a part in grounding Lambert's project (see discussion of Lambert's process ontology, in chapter 4 of this work).

Other significant work that has come to the attention of the author, include Rogers (1993) unsuccessful attempt to formulate a Whitehead inspired approach to AI. Rogers' work draws attention to the complexity of the task of translating Whitehead's philosophical language into a formal model. Also of significance is the work of Granville Henry (Henry&Geertsen 1986; Henry 1993). Henry has explored the extent to which Whitehead's philosophy can be mechanized, and suggests some analogies between Whitehead's metaphysical system and computer programming structures; however, his chief aim is not to advance AI, but rather, to advance our understanding of Whitehead's philosophical system (Henry, 1993:25). Process philosophy has also been the source of inspiration for David Primeaux's artificial actual entities; however, his work is not an attempt to supplant the Classical AI paradigm. Rather, it is a process inspired adaptation of methods developed for Artificial Neural Nets (ANNs). See Primeaux (2000), Primeaux and Saunders (2006) and Saunders (2008).

Of work in the cognitive sciences that converges with significant aspects of Lambert's critique of the classical approach to AI, and with his vision of a process philosophy inspired alternative understanding of human 'commonsense', the most important is that of the interactivist school, based upon the work of Mark Bickhard. Bickhard, like Lambert, was driven toward a process philosophical position by a detailed consideration of problems associated with the classical symbolic approach to representation the philosophical problems associated with Classical AI serves as the philosophical blueprint for his subsequent practical work in developing a working information fusion system. This work, spanning nearly ten years, and carried forward by a multi-disciplinary team, has been funded by the Australian Government under the auspices of the Defence Science and Technology Organization (DSTO), and represents a significant Government backed research commitment to develop and utilize information fusion technology. Thus, his early work has its own intrinsic value as a proposal for a process philosophy inspired paradigm shift in AI, but this work has an additional extrinsic significance, in that it provides the context for understanding and evaluating Lambert's subsequent work in information fusion.[2]

My appropriation and utilization of Lambert's work as a platform for developing process philosophical and computational foundation for socionics was inspired by synergies between his process philosophy inspired vision of AI and my process philosophy inspired approach to sociological theory. *EMWC* develops a vision of mind, knowledge, language and the world that reframes our notion of human and machine intelligence, and creates a context for speaking about cognitive machines and machine societies. Lambert's process philosophy inspired vision of a paradigm shift in artificial intelligence complements my vision of a process philosophy inspired paradigm shift in social theory. Both theories in turn underpin my vision of a process philosophy grounded synthesis of computer science and sociology, contributing to the clarification and development of socionics.

The vision developed by Lambert in *EMWC* provides the context for my outline of his blueprint for information fusion and of his process vision of a hybrid, human-machine society (or society of societies). These hybrid societies will arise if (or when) Lambert and associates operationalize a high-level information fusion capability. In the spirit

---

(Bickhard, 2009b). However, in contrast to Lambert's engineering background, Bickhard's work developed out of research in psychology, and appears to have made contact indirectly with the modern process philosophical tradition through James Gibson's pragmatist inspired analysis of perception.

A number of researchers working in the area of AI and intelligent robotics have drawn inspiration from Bickard's work (see review article, Sojanov et.al. 2006), but with a minimal indeptedness to the modern process philosophical tradition (as identified by Browning & Myers, 1998; Rescher, 1996 & 2000). In particular, the interactivist approach has not drawn upon the work of A. N. Whitehead; as Bickhard noted recently: "The development of interactivist approaches to a process metaphysics has proceeded largely independently of Whitehead, with at best limited convergences" (Bickhard,2009a:449).

In the author's opinion, Bickhard's work has (contrary to Bickhard's assertion) significant convergences with Whitehead's process outlook. In particular, a Whiteheadian framework such as that developed in this thesis could serve as a basis for comparing Bickhard's and Lambert's work, and their respective approaches to framing a process inspired alternative to the classical approach to representation. Such a comparison would constitute a valuable base for clarifying the presuppositions and implications of Lambert's project for cognitive science.

[2] Thus, for example, Lambert's vision and program for developing high-level information fusion has been criticized on philosophical grounds by Giffin and Reid (see Giffin, 2002; Giffin & Reid, 2003a (Canto 1); Giffin & Reid, 2003b (Canto 2); Reid & Giffin, 2003 (Canto 3)). Lambert's response was to argue that their criticisms were based upon an inadequate grasp of the philosophical presuppositions framing his project, as developed in *EMWC* (Lambert & Scholz 2005:8-11). That is, Lambert explicitly affirms that the philosophical standpoint developed in *EMWC* has framed his subsequent work in the field of information fusion, and continues to defend the philosophy underpinning his *EMWC* position against competing contemporary schools of philosophical thought.

of the socionic project, in subsequent work my aim is to supplement Lambert's work, drawing primarily upon sociological theory in order to explore the implications of Lambert's work for the sociological domain. By focusing on the 'higher-level' aspects of information fusion (those that involve complex social and cultural phenomena and questions) the deeper ramifications of the quest to develop a machine with a capacity for 'commonsense' are disclosed, and Lambert's effort to develop a theoretical framework that encompasses both human and machine participation in a hybrid society becomes the focus of a sociologically informed reformulation. However, this work is only a first step towards that larger goal.

Thus, Lambert's vision of a Post-Classical AI and the blue-print for Information Fusion technology that his early work has inspired, plays a key role in grounding my vision of theoretically and technologically unified foundations for Socionics, and the new forms of social experimentation and innovation that will emerge from this fusion of theory and technology. As Lambert's foundational ideas are developed primarily in *EMWC*, and as these ideas have not received much attention elsewhere,[3] this chapter will outline Lambert's critique of Classical AI, and prepare the ground for a review of his vision for a Post-Classical, process inspired AI.

2.1 Knowledge Representation and Commonsense in Classical AI

In *EMWC* Lambert argues that the effort to engineer a machine capability equivalent to human 'commonsense' is the most difficult challenge facing AI. Classical AI modeled itself upon human cognition, understood as mental processes that intervene between human responses (outputs) to stimuli (inputs). Classical AI equated these mental processes with the running of symbolic programmes; hence (as outlined in ch.1), building 'commonsense' machines was equated with symbolic programming. However, although Classical AI can build machines that perform well in highly specific domains in very restricted ways, these expert systems lack the flexibility and robustness that humans display when they employ their expertise. Lambert argues that the consequent 'brittleness' of Classical AI (its inability to cope with unusual inputs) is the result of inadequacies that have their roots in the philosophical presuppositions that inform Classical AI's approach to representation. Hence, Lambert's critique of Classical AI revolves around questioning the adequacy of its approach to knowledge representation.

Knowledge representation (KR) developed as a branch of AI, but subsequent developments in database design, object-oriented systems, and the design of the semantic web have converged to give the subject an expanded relevance. The demands of information fusion represent a further challenge for KR engineers. KR is a multi-disciplinary subject, cutting across the fields of logic, ontology and computation, and encompasses questions concerning the relationship between knowledge and the real world, the nature of that knowledge, and strategies for mapping that knowledge to a computable form (Sowa, 2000). Knowledge engineers

[3] Although the ideas developed in Lambert's 1996 dissertation are frequently referred to in the many articles published by Lambert and associates since that date (see Bibliography 1998 through to 2009), brief references in short technical papers or seminar presentations to ideas first developed in *EMWC* convey neither the broad context of his critique of Classical AI nor the scope of his vision of an alternative Post-Classical AI paradigm. The only explicit references to *EMWC* in the literature are found in Lambert 2001b, Lambert and Scholz 2005, and in Nowak 2001 & 2003.

are the midwives of Classical AI, in that they attempt to bring out and make formally explicit the knowledge that is often only implicit in human expertise. In going about this task Classical AI engineers drew upon a background set of assumptions concerning human knowledge, the relationship between natural and formal languages, and the relationship between these systems of representation and the world that they represent.

Lambert's critique of Classical AI's representational strategies steps back from the formal details of logic, ontology and computation, and examines the philosophical presuppositions that have formed the largely unquestioned philosophical background of the work of knowledge engineers. This means tracing the notions of language and of rationality that have informed Classical AI back to their origins in philosophical models of the function of reason and the role of language in representing and mediating the mind's understanding of the world. Philosophical considerations of this sort have underpinned the development of Classical AI and its approach to representation. For reasons examined by Agre (1995, 1997b) critiques that address these foundational issues have tended to be treated as marginal to the real work of AI; however, if AI dissidents such as Agre and Lambert are correct, even though the philosophical dimension may appear peripheral to the technical practice of AI, it is in fact foundational and deserves serious attention.[4]

In *EMWC*, Lambert examines the role of representation in Classical AI and suggests that Classical AI's approach to representation can be evaluated in terms of its metaphysical, heuristic and epistemological adequacy. Drawing upon the work of McCarthy and Hayes (1969), Lambert reformulates the role of representation in practical knowledge engineering terms, and sets criteria for evaluating its metaphysical, epistemological and heuristic adequacy.[5] These are specified as follows:

---

[4] Philosophically oriented critiques of AI, such as Dreyfus (1972, 1992), Agre (1997a) and Clancey (1997), remain peripheral to AI for the reasons addressed in Agre (1995, 1997b); namely, AI is primarily a *technical practice.* A range of intuitions and assumptions underpin this practice (implicit in everything from the metaphors to technical methods). Philosophical presuppositions, such as those analysed by Lambert, are rarely made explicit; they instead form the implicit conceptual ground and tacit horizon that mediates and guides our thought. Even if a computer scientist draws upon a different philosophical horizon, and is led by novel intuitions to create the elements of a new technical practice, the artefacts that they produce will tend to be appropriated within the dominant paradigm, and the novel intuition construed as 'nothing new'. In order to grasp the newness of the intuition, and the novelty of the method, it is necessary to laboriously fight your way clear of the existing paradigm and acquire the interlocking structure of methods, intuitions, concepts and theories that instantiate the new paradigm. As Agre (1995, 1997b) notes, this is a tough ask and makes scientific practices very resistant to change.

[5] McCarthy and Hayes (1969) argued that these three kinds of adequacy are necessary for a representation of the world. They are defined as follows:

> A representation is called metaphysically adequate if the world could have that form without contradicting the facts of the aspect of reality that interests us. . . . A representation is called epistemologically adequate for a person or a machine if it can be used practically to express the facts that one actually has about the aspect of the world. . . . A representation is called heuristically adequate if the reasoning processes actually gone through in solving a problem are expressible in the language. (McCarthy & Hayes,1969:469-470)

1. metaphysical adequacy[6] questions the legitimacy of the experts' conceptualization of the world and linguistic primitives chosen to recover it
2. epistemological adequacy[7] question's whether the knowledge engineer has faithfully represented expert opinion
3. heuristic adequacy[8] questions whether the knowledge engineer has faithfully represented the derived theory in a computationally viable form (*EMWC*:43-44)

Lambert argues that the Classical AI paradigm fails on all three counts, due to its Aristotelian and Cartesian philosophical presuppositions, and proposes a Post-Classical, alternative AI.

In framing his vision of an alternative AI, Lambert charts an AI equivalent of a path between the philosophical extremes of cultural-linguistic solipsism on the one hand, and positivist-materialist reductionism on the other. His work thus addresses the challenge enunciated by Agre and Horswill (1997); that of creating a computational model that conceptualizes "agents' adaptations to their environments in ways that neither treat agents as isolated black boxes or dissolve them into one big machine" (Agre&Horswill, 1997:1). In AI terms this means resisting the 'black-box' solipsism of the traditional declarative planning model and avoiding the extreme anti-representational approach of Brooks' subsumption architecture. Lambert attempts to achieve this balancing act by developing a process inspired variant of procedural planning.[9] To implement this dynamic vision of environmentally situated planning and acting, Lambert proposes a multi-agent architecture. Although his model resembles the BDI (Belief, Desire and Intention) architecture associated with Bratman

---

[6] Metaphysics is that branch of philosophy that has attempted to specify the essential and universal form of the entities and processes that constitute reality, as exemplified in the process and substance-oriented characterizations of the external world attributed to Heraclitus and Aristotle. In questioning the metaphysical adequacy of representation in Classical AI, Lambert is questioning Classical AI's presuppositions concerning the relationship between its representations and their external referents.

[7] Epistemology is the branch of philosophy that has attempted to specify reliable principles and procedures for establishing the truth, as exemplified in rationalist and empiricist characterizations of the way in which the mind gains access to the truth outlined by Descartes and Locke. In questioning the epistemological adequacy of representation in Classical AI, Lambert is questioning the adequacy of Classical AI's presuppositions concerning the role of representations in facilitating valid judgments and responses to inputs concerning the world.

[8] Heuristics is concerned with problem solving processes as employed in practice rather than epistemological ideals related to truth seeking. In psychology, for example, a heuristic method denotes a mental short cut that allows a problem to be solved without engaging in an exhaustive rational analysis of the problem domain. People rely on such heuristic methods for decision making and problem solving. A simple heuristic is the trial and error method; this has its computational equivalent in a 'generate and test' search strategy, whereby a program is designed to halt when a 'near enough' solution is attained (as opposed to a proof that exhausts all possible solutions). Philosophically, the question of heuristics focuses attention on our presuppositions concerning the nature of the human person, and their problem solving capabilities.

[9] Machine agents that utilize procedural planning, that is, procedural reasoning systems (PRS) interweave doing and planning by drawing upon a library of pre-compiled plans in the course of action. Strictly speaking, they do no planning as the plans they draw upon have been manually constructed by a programmer. Declarative planning, on the other hand, works from first principles to generate a complete plan prior to action (replanning from first principles if the unexpected occurs). Subsumption architecture is based upon reactive planning; instead of representing the problem domain it employs simple behaviour modules that trigger agent responses. Recent developments in BDI agent-based planning procedures incorporate procedural and declarative planning, as discussed in Meneguzzi, Zorzo, et.al. (2007:2-3) and Meneguzzi and Luck (2007).

and others, Lambert reframes the notions of beliefs, desires and intentions within the terms of his Post-Classical, process inspired AI.[10] In the following sections I will detail the specific shortcomings of the Classical paradigm, as identified by Lambert. This will serve as a prologue to an outline of his Post-Classical proposal for remedying those shortcomings.

2.2 Classical AI's Epistemological Failure

In practical knowledge engineering, the question of epistemological adequacy concerns whether the knowledge engineer has faithfully represented expert opinion. A representation is epistemologically adequate if it provides a practical expression of the facts as articulated by the expert. But how does the expert arrive at their opinion? Not, according to Classical AI, by ‘gut instincts’, experience based ‘hunches’, or some such opaque process; on the contrary, Classical AI engineers regarded human beings as rational problem solvers. Following in the footsteps of Descartes, they assumed that the process of arriving at the truth is a rational process, open to introspection. From this rationalist perspective commonsense was viewed as a process of making judgments (based on intuitions and deductions) that could be simulated by first discursively representing, in natural language, the mind's facts and implicit theories, and then second, by formalizing these beliefs and theories in a logical system. Thus, Classical AI adopted the Cartesian rationalist paradigm, and assumed that commonsense is a rational process that can be introspectively grasped and adequately formalized. Lambert presents two arguments as to why this Cartesian inspired notion of rationality falsely represents the way that we arrive at commonsense judgments or solutions to problems in a particular domain.

Firstly, it assumes that commonsense is based upon the kind of logical completeness that Hilbert associated with adequately axiomatized mathematical systems. Just as Gödel demonstrated that Hilbert's dream of finding a complete and consistent set of axioms for all of mathematics is impossible, so Lambert argues that the dream of formalizing commonsense in a computable system is doomed to fail. As Classical AI moves beyond simple data bases and theories that describe "toy" worlds, and develops large, complex theories of commonsense knowledge, logical systems must contend with both inconsistency and undecidability rooted in the formal and practical incompleteness of human knowledge. Increasing the size and theoretical complexity of the formal representation of commonsense only increases the likelihood of computers succumbing to some form of Turing's halting problem, exposing the brittleness of the system – the very problem the AI simulation of "commonsense" is meant to overcome.

---

[10] BDI architecture contains explicit representation of beliefs, desires and intentions. Beliefs are the information the agent has about its environment (can be false), desires are the agent’s goals, and intentions are the specific things that the agent intends to do in order to achieve its goals. Agent architecture specifies how the beliefs, desires and intentions of the agent are represented, updated and processed in order to determine agent action. Typically, agent action can be viewed as a process, beginning with the agent considering its sensor input and updating its beliefs, then on the basis of these updated beliefs, drawing upon its desires to select a goal, and in turn, drawing upon its plans to select a means to achieve its intended goal, and lastly, carrying out actions specified by the intention. See Bratman, et.al. (1988), Rao and Georgeff (1991), and over-views in Weiss (1999) and Woolridge (2002).

The second erroneous epistemological assumption concerns gaining access to the kind of "commonsense" knowledge that the expert takes for granted. Do the expert's introspective reflections upon their decision making process really disclose how they act? The rational reconstruction of commonsense, in the form of facts and rules, may simply reflect a rationalist myth, rather than the actual causal processes that determine action. The real basis for commonsense, as some Situationalist's argue, might be finely tuned adaptive responses – reflexive 'gut' reactions to richly structured external environments - rather than reflective responses to representations. The external context that influences a judgment may be the product of various factors, including time, physical place, social situation, cultural norms, and contingent negotiations with other actors.[11] From the standpoint of the situated approach to cognition, most problem-solving behaviour is not a response to "laws of thought"; if it were so we would be forever falling victim to infinite or intractable mental computations. Thus, in Lambert's words:

> "Our introspective theories are rationalist reconstruction, not candid insight into our innermost workings. Rationalism provides an inept basis for engineering machines with commonsense". (*EMWC*:62)

2.3 The Heuristic Inadequacy of Classical AI

Lambert deems a computational representation of a problem solving procedure to be heuristically adequate if the representation captures the reasoning processes actually used by a person to solve a problem, and represents that reasoning process in a computationally viable form. [12] Since Lambert has argued that Classical AI fails on epistemological grounds to understand the nature of our human 'commonsense' problem solving capacity, any heuristics based upon Classical AI's overly rationalist reconstructions of human problem solving will fail to capture the real basis of our commonsense capacity. However, these epistemological considerations do not throw much light on the heuristic strategies that underpin commonsense; hence, in order to reconsider the heuristic role of representation and rationality in AI, Lambert contrasts the heuristic strategies of Classical AI with the strategies employed by Situated AI.

2.31 Reason and Reaction

Lambert argues that the Classical AI approach to heuristic adequacy is really about how a machine's internal functions endow it with intelligence. Classical AI maintains that a commonsense machine functions internally by reasoning, and "views reasoning as rational problem solving through the solipsistic manipulation of stored representations" (*EMWC*:63). Lambert notes that this approach to heuristic adequacy is informed by the Classical AI notion that reasoning is a solipsistic, rule-based process of manipulating stored representations, and then contrasts this view of reason

---

[11] Devlin has noted that some AI practioners have responded to the obvious importance of context by arguing for a relativised rationalism, in which universal rules are replaced by domain specific sets of rules. This strategy, though it produces some gains, also succumbs to the brittleness problem. (Devlin, 1997:267-273)

[12]Lambert quotes Feigenbaum & Feldman's definition of heuristic: "A heuristic (heuristic rule, heuristic method) is a rule of thumb, strategy, trick, simplification, or any other kind of device which drastically limits search for solutions in large problem spaces. Heuristics do not guarantee optimal solutions; [. . . but offer] solutions which are good enough most of the time" (Feigenbaum & Feldman,1963:6, quoted in Lambert, 1996:27-28).

with the Situated AI approach to our commonsense ways of assessing and responding to situations.

In contrast to Classical AI's 'black-box' view of intelligence, those advocating a 'situated' approach to AI suggest that externally induced reaction, rather than internally conceived reason, is the "catalyst for intelligence" (*EMWC*:64). Lambert notes that from the Situated AI perspective, it is not reason but reaction to the specific characteristics of our external environment that is the root of our commonsense capacities. Lambert observes that:

> The situated stance on stored representations is possible because the situationist shifts the emphasis out of the head and into the world. The situated world is a world having some agreed ontology and the creature is understood in terms of its interaction with that ontology. Situated commonsense accommodates an unsophisticated creature in a sophisticated world. Classical AI, by contrast, places greater emphasis on what goes on inside the head, and this engenders its solipsistic machines and its interest in sophisticated internal representations. Philosophically this is simply a new battle in an old war.[13] (*EMWC*:66)

Brooks' subsumption architecture exemplifies the reaction-based Situated AI alternative. Situated AI regards the 'world as its own best model', and either devalues or repudiates the notion that stored representations can advance the cause of machine intelligence. Why consult memories when you can directly consult the world? A trial and error based search for the right sized hat, for example, as opposed to a purely rational search based on 3D geometrical modeling of hats and heads, involves interacting with the external environment and deriving the problem solution from that physical interaction: try it on, and if the hat fits wear it.

Lambert notes that hybrid fusions combining elements of both classicist and reactive architectures are possible; and that not all reactive architectures fall within the situationist scheme. In the 1980s, the real world heuristic inadequacy of classical planners (based on the classical commitment to planning via searching and reasoning through the manipulation of stored representations) gave rise to a range of alternative strategies that fell somewhere along a continuum stretching from Brooks' subsumption architecture to the representationalist ideal of Reason. Close to the reactive end of the continuum, for example, universal planners do not devise plans as the need arises, but instead "come with a ready made assortment of stored reactions to possible situations, which are combined as the situation dictates" (*EMWC*:68). Proceeding along the Reaction-Reason continuum brings a greater emphasis on internally manipulated representations (world model) and the sophistication of its representation language. Advocates of Reason emphasize the importance of the human capacity to anticipate events and situations, and form plans of action to achieve their goals, while the reactive advocates argue that representation based reasoning systems are in practice ineffective (*EMWC*:69).

In responding to Brooks' vision of intelligence without either representation or reason, Lambert foreshadowed the current consensus and argued that Brooks had overstated his case. Representation and reason have an obvious role in human intelligence, and

[13] Presumably Lambert is referring to the long running conflict between Idealist and Materialist forerunners of cognitive and behaviorist psychology; a conflict which can be traced back to Kantian notions of mind on the one side and to Locke-Hartley Associationist notions of the brain on the other.

Classical AI's cognitively inspired vision of intelligence has led to the development of forms of representation and symbolic manipulation that are useful. However, the Situationists do have a point in emphasizing the extent to which the intelligent capacity of agents is a function of being embedded in a rich environment. Lambert argues that this characteristic responsiveness cannot be added to Classical planners simply by situating them in the world and equipping them with a reactive capacity via sensors. Scaling-up the reactive capacity of planners in this way assumes that we are rational problem solvers, and that AI can imitate our commonsense capacities by reconstructing the expert theories underpinning our rational problem solving. Lambert argues that most commonsense behavior has non-rational origins in drives and routines. Hence, the main target of his criticism is not representation *per se*, but the rationalist philosophy underpinning Classical AI.[14]

The essential role of the environment in the Situated AI paradigm points to one of the primary defects of the rationalism informing Classical AI; namely, its solipsistic view of intelligence. Classical AI has focused on engineering solipsist machines – minds in a box detached from the dynamically changing world. This detachment from the world is a result of the Classical AI presupposition that representation is able to reliably inform us about the world without directly interacting with the relevant parts of the world. The Classical AI approach to Knowledge Representation is solipsistic in style because it allegedly enables the machine to make claims about the world without interacting with it. Lambert argues that tacking sensors and effectors onto the machine does not overcome this solipsism, as the latter is a property of a style of representation that attempts to deduce the nature of the world without depending on interaction with that world. Implicit in the rationalist notion of representation are notions of the completeness of knowledge and of its capacity to capture the essential characteristics of the world (this will be further discussed in reference to the metaphysical adequacy of Classical AI). Classical AI gives representation a central role in its heuristic strategies; however, its notion of representation is embedded in a solipsist rationalist philosophy. Lambert argues that representation need not entail rationalism, and suggests that Post-Classical AI should reject solipsist notions of representation that "impute a priori claims about the world" (*EMWC*:71). Post-Classical AI should thus take advantage of the symbolic representational techniques associated with Classical AI, but rid itself of the philosophical presuppositions implicit in rationalism.

2.32 Post-Classical Representation

In proposing a more adequate heuristic strategy, Lambert suggests that AI retain the Classical strategy of using stored representations, but in accord with the Situationists that these representative tokens should relate to the world in a way that does not imply solipsism. The contrast between Classical, Situationist and Post-Classical AI is represented in Figure 1. As an example of how representation might apply to the world without entailing solipsism Lambert suggests that it might be possible to

[14] Lambert argues that 'rationalism' is a philosophy that claims to reconstruct the innermost workings of commonsense, and to explain commonsense behaviour primarily in terms of rational problem solving. Lambert contends that this is false; commonsense acquires its rational character through a process of reconstruction through which we seek to predict and explain ourselves. This reconstruction may have some minor effects upon our behaviour, but most commonsense behaviour is a product "of animal and heuristic origins, the details of which we are unable to introspectively explain" (Lambert, 1996:59).

classify the world in the process of interacting with it, using type structures, rather than rationalist systems of knowledge representation "which try to deduce the nature of the world independently of it" (*EMWC*:71).[15] Hence, post-classical AI seeks a computationally viable method of storing and manipulating representations; a method that does enable the machine to deduce further properties of the world independently of the world but avoids the solipsist presupposition that representations adequately capture the essential nature of the world.

	Stored Representation	**Solipsism**
Classical AI	Yes	Yes
Situated AI	No	No
Post-Classical AI	Yes	No

Figure 1: Divergent Paradigms – the heuristic role of representation in AI. (*EMWC*:71)

2.4 Metaphysical Adequacy of Representation in Classical AI

Lambert's philosophical critique of Classical AI calls into question the rationalist presupposition that 'commonsense' is based upon a coherent metaphysical representation of the world. The metaphysical adequacy of representation in Classical AI refers to whether representations recover the facts of interest in the world. Hence, with practical expert systems design, "metaphysical adequacy questions whether a nominated conceptualization of some expert aspect of the world is correct, and whether the representational primitives selected accurately recover that conceptualization" (*EMWC*:72). That is, do the primitive concepts and theory implicit in the expert's knowledge adequately capture that actual aspect of the world it seeks to represent, and is the formal system used to represent this conceptualization of the world able to recover this content.

Lambert notes that the meaning content of a representational system could refer to domains as diverse as space, liquids, and emotions, and notes that various specialist axiomatic ontologies have attempted to represent the expert's understanding of these domains. For example, knowledge engineers have attempted to capture the way in which people represent the physical world to themselves when they make commonsense judgments about everyday mechanical tasks, such as filling a glass with water, and make explicit the 'naïve physics' of this commonsense knowledge of the physical world within axiomatic ontologies. However, Lambert's focus is not on the metaphysical adequacy of these specific ontologies; instead, he looks at the shared assumptions underpinning the effort to specify the meaning content of commonsense. His conclusion, based on two grounds, is that the idea of axiomatizing commonsense knowledge is itself mistaken.

Firstly, the notion that a formalized 'commonsense' ontology (that makes explicit the nature of the world, as grasped by our commonsense), can facilitate machine intelligence, assumes compatibility between the human and machine modes of interfacing with the world. That is, it assumes that the cognitive architecture of human and machine are sufficiently alike that the human mode of interacting with the world

[15] Lambert is referring to the notion of types employed in model theory. This will be discussed further in later chapters, but loosely speaking types describe the possible elements of a mathematical structure.

can be used as a guide to engineering the machine mode of interaction with the world. However, given the architectural and sensory differences between human and machine 'cognition', how a machine appraises and interacts with the world are bound to differ from a human's 'commonsense' appraisals and interactions. "The engineered theory of commonsense is therefore likely to be metaphysically inadequate for the robot, since the commonsense tokening of [the ten year old's commonsense] is likely to irreconcilable with the tokens of the robot's [commonsense] view." (*EMWC*:73) The architectural differences between the sensors and processes that mediate the embedding of a machine and a human in the world are one reason why the ontology implicit in human commonsense is unlikely to be of use to a machine.

The second limitation relates to the linguistic realism that has underpinned Classical AI's notion of the role of representation in implementing commonsense, and concerns the way in which common language is presumed to intentionally relate to the things it represents. Most advocates of Classical AI have assumed that commonsense involves grasping, via common language, a realistic image of the world and have envisaged translating this realistic image of the world into a formal symbolic language. In this context 'common language' refers to words that serve as tokens for allegedly mundane and obvious knowledge concerning things and activities. Classical AI assumes that common sense can be expressed in common language that reliably reports the nature of reality (linguistic realism), and that this description can serve as the basis of a computational implementation.

Lambert argues that representation in Classical AI embraces a commitment toward the manner in which knowledge should be represented *internally within a machine,* namely that the "*symbolic workings of the machine should accord with our common language conceptualizations" (EMWC:73-4).* Under this 'Internal Language Hypothesis', machines internalize tokens that represent the semantic relationship between common language and the nature of reality. It is assumed that the image of the world represented in common language corresponds, more or less, with reality ('linguistic realism'), and that this image is an adequate basis for guiding the machine's interaction with the world. Hence, Classical AI is based on the notion that human commonsense notions of the world (mediating our everyday 'manifest image' of the world), rather than scientific models of the world, constitute the most reliable guide for building an ontology of the world.[16]

Lambert asserts that common language 'linguistic realism' and the related 'manifest image' realism both assume the metaphysical adequacy of commonsense and its implicit theories of the world. Denial of either linguistic realism or manifest image realism implies a denial of the metaphysical adequacy of commonsense theory. Lambert argues that there are several grounds for questioning the metaphysical adequacy of commonsense implied by realist notions of common language, some of which are implicit in the Situationist and Connectionist approaches to AI. However, the major shortcoming of linguistic realism is its notion of intentionality.

[16] Wilfrid Sellars distinguished between the 'manifest image of man-in-the-world' and the 'scientific image of man-in-the-world'. Put loosely this refers to the distinction between the commonsense view of the world composed of every day objects and the scientist's theoretical view of the world, made up of abstract theoretical entities (*EMWC*: 74). Classical AI has resolved the tension between these two images by not concerning itself with the scientific images of the world and embracing the manifest image view; hence, "the manifest image is the metaphysics of Classical AI" (*EMWC*:75).

In the twentieth century, both Continental and Analytic philosophers have focused attention on the intentional relationship between language and the world. Lambert's focus is on linguistic analysis, and the two main approaches to intentionality developed by philosophers working within the Anglo-American Analytic tradition. He begins by asserting that the crucial feature of linguistic realism is the *intentionality* that connects semantic content via tokens to things in the world; that is, his focus is on the referential property of intentionality that enable linguistic tokens to mean things *about* things in the world. Propositions are assumed to be true because their content is validated by the connection between symbol and its referent in the world. For example, the symbols used to express the proposition that 'it is snowing on Mont Blanc', are intentionally related to the real Mont Blanc, and express a true proposition if it really is snowing on that mountain. The manifest image metaphysics of Classical AI rests upon the assertion that meanings are grounded in the structures of reality, and hence that language is in some sense a mirror of reality. Lambert's analysis demonstrates the failure of attempts to justify some version of this 'linguistic realism'.

2.41 Referential Linguistic Intentionality

This is the most difficult section of Lambert's critique of Classical AI. Through some fifty densely argued pages, Lambert charts a path through a philosophical labyrinth of complex debates and provides a justification for calling into question the metaphysical adequacy of the main approaches to representation adopted by Classical AI. However, before following Lambert into this labyrinth it may be helpful to step back and provide a simple overview of what lies before us – a sandbox representation of where we are heading.

As noted previously, intentionality is at the heart of unresolved debates in the Continental and Anglo-American versions of the philosophical 'linguistic turn'. Since both traditions generated a host of internal contradictions and philosophical aporias and disintegrated into a veritable philosophical Babel of conflicting voices, any overview will be to some extent an artifact of the author's intentions. In Lambert's case, his intention is to expose a particular weakness in the Classical AI tradition, and he does this by focusing on two strains of thought within the tradition of linguistic analysis that have influenced the development of Classical AI.

Hence, in creating this sandbox landscape, the first step involves making a sharp distinction between analytic and phenomenological approaches to intentionality. Though both analytic and phenomenological traditions had similar philosophical origins in investigations (initiated in the late nineteenth century by thinkers such as Franz Brentano, William James, Ernst Mach, Gottlob Frege, Edmund Husserl and Bertrand Russell) into the relationship between consciousness, language and logic, in the early decades of the twentieth century the cross-fertilization of ideas between pioneers came to an end, and two separate traditions began to emerge, each advocating very different approaches to the analysis of intentionality. The goal of the phenomenological tradition, as exemplified in the works of Husserl, was to drill down to the phenomenological foundations of logic and science; his method was a radical empiricism (in James' sense) that was prepared to take nothing for granted. He initiated critical inquiries into the nature of logic and of its relationship to the sciences that lead toward a fundamental questioning of the accepted psychological notions of intentionality and prepared the ground for reframing our notions of intentionality in terms of a new, phenomenologically inspired metaphysical system.[17] The analytic tradition, on the other hand, was hostile to this notion of philosophy.

Analytic philosophy has its roots in an approach to philosophy defined by Russell and Wittgenstein. Instead of questioning the grounds of symbolic logic and behaviorist psychology, they sought to reconstruct our understanding of language and truth on the basis of these sciences. Faith in science, rather than a phenomenolgical enquiry into the metaphysical universals of human experience, was the key to disclosing reality. In the *Tractatus,* Wittgenstein's early statement of faith in the capacity of logical, empirically based propositions to exactly map the properties of the real world, the vision that inspired many later analytic philosophers was stated with great precision:

> "The totality of true thoughts is a picture of the world" (3.01). "The picture is a model of reality" (2.12). "The proposition is a picture of reality, for I know the state of affairs presented by it, if I understand the proposition" (4.021). "Reality is determined by the truth or falsity of the proposition; it must therefore be completely described by the proposition" (4.023).

Those sympathetic to the early Wittgenstein's approach, such as Carnap and other 'logical positivists', drew upon a combination of behaviorist empiricism and apriori rationalism to analyse language and language use. Their approach to language assumed that every valid proposition is based upon either a logical tautology or an empirically verifiable fact. Metaphysics was derisively grouped with theology and regarded as the mother of useless concepts; from this standpoint, scientific concepts were a product of the sound use of logic and empirical observation while metaphysical postulates were the product of speculative fancy. Hence metaphysical concepts had no valid referential sense – they were simply nonsense – in contrast to concepts that satisfied positivist criteria of meaningfulness. The valid referential sense of language could be explained in two complementary ways, an 'externalist' non-psychological and 'internalist' psychological approach, reflecting the logicist and behaviorist sides of the positivist tradition.

As an example of the externalist approach to explaining the referential form of linguistic intentionality, imagine a person who points at an animal and says either "Norty is a cat", or "Norti es un gato". The externalist explanation of this act focuses on the meaning content of the proposition, or eidetic essence indicated by both phrases, and the role of this intended meaning in mediating between the phrase and the external state of affairs corresponding to the proposition. According to this externalist theory, a particular mental act is meaningful if it refers to a propositional essence of this sought, and the proposition is true if there is a state or entity in the material world that corresponds to the idea signified by the proposition. The meaning of the proposition lies in the linkage between the words and the existence of the cat called Norty. This correlation is assumed to be between a substance and its properties, and the content captured in the corresponding subject-predicate structure of the words, such that *Norty is a cat* has the same meaning as cat(Norty) because the logical

[17] This quest for a more adequate metaphysical ground is reflected in the work of Husserl (1970c) and (though his work developed independently of Husserl) in the work of Whitehead (1929). Lambert's Post-Classical AI proposal can be viewed as a speculative endeavor in the same radical spirit; one that builds upon a process metaphysical approach to representation in order to develop an alternative approach to intentionality.

components of each term have the same external referents. This external approach to grounding meaning contrasts with the internal approach.

An internalist theory might suggest that the referential form of linguistic intentionality in the previous example is simply the set of mental sensations associated with thinking and perceiving "Norty is a cat". Truth is defined as the psychological correlation between the set of mental sensations associated with the intentional thought/utterance, and a corresponding set of sensations and perceptions associated with the event of seeing what the subject experiences as 'the cat named Norty'. Meaning in this case is located internally, in the subject's mental experience. These two realist approaches to meaning were dominant in the analytic tradition during its formative decades, however during the thirties Wittgenstein began to question his previous work and this resulted in a very different approach to language.

The later Wittgenstein's work called into question his earlier 'picturing' model of the relationship between language and the world. Apart from describing an object, we can use language to tell stories, make up riddles, curse, greet and pray. As he stated in *Philosophical Investigations*:

> There are *countless* kinds: countless different kinds of use of what we call "symbols", "words", "sentences". And this multiplicity is not something fixed, given once and for all; but new types of language, new language-games, as we may say, come into existence, and others become obsolete and get forgotten. . . Here the term "language-*game*" is meant to bring into prominence the fact that the *speaking* of language is part of an activity, or of a form of life. (Wittgenstein, 1953:1.23).

The 'later' Wittgenstein's influence grew in the post-WWII decades and the positivist ideal of a scientifically purified language was supplanted by a new version of analytic philosophy that involved fine grained logical analyses of all kinds of scientifically unsound, but conventionally acceptable forms of language use. Names such as Austin, Ryle, Quine and Sellars came to the fore, and analytic philosophers focused primarily on the communicative function of forms of language used in everyday discourse. Typical of this new focus was Wittgenstein's suggestion that everyday propositions, as employed in normal conversation, derived their meaning from their conformity to the rules of various culturally defined 'language games'. His analysis of such 'games' motivated new approaches to formalizing language based upon the need to recognize the context or situation of language usage. This approach to meaning suggested that the meaning of words - semantics - was derived, not from logical tautology or empirical fact, but from the context of the usage of the word.

In order to give rigor to this insight, some analytic philosophers attempt to adapt Tarski's (1933) semantic approach to truth to the diverse and dynamic contexts that define the meaning of words. Following Tarski, earlier analytic philosophers had attempted to avoid the complexities of relating language to the real world by restricting the referential claims of language to formal languages and models. Those with positivist inclinations had dreamed of one universal model of reality, a logical calculus from which all possible true propositions about the world could be derived, all of which would map precisely onto the real world. Later philosophers, such as Kripke, rejected the dream of a single all-encompassing mathematical model, assumed a multiplicity of models, representing the multiplicity of possible worlds that language can constitute. Barwise and Perry (1983) drew upon this model theoretic approach to develop a system which they hoped would be able to deal with diverse real-life language-use contexts, as described by the later Wittgenstein, and adopted model theory to partial worlds, called *situations*. Barwise and Perry's application of model theoretic mathematical techniques in situation theory reflected an effort to take the mathematical form of symbolic logic from the artificial and narrow mathematical confines of its early development, and apply it within the complex domains of real-life; this in turn reflected the movement in analytic philosophy away from the construction of an ideal logical language, toward a preoccupation with how in fact natural languages function.

Thus, in the post-WWII years, analytic philosophers became more interested in natural language usage, and shifted their focus to the practical use of language. Instead of seeking essentialist meanings for notions such as 'reality', 'truth', or the 'good', natural language philosophers looked at the context of word usage to arrive at the meaning of the word. For example, if a person said "this is not *real* coffee!" they are probably not making a metaphysical statement about reality; rather, they are simply expressing a preference for expresso over instant coffee. Likewise, when the same person states "now that is a *good* coffee!" they are probably not expressing an ethical proposition; rather, the 'good' in this context is an expression of satisfaction over having been given the coffee that people sharing their cultural aesthetic define as desirable. This would probably be true if the conversation was taking place in a café between friends; on the other hand, if the conversation was taking place in a philosophy tutorial the speaker may well have been using the sentences to illustrate a philosophical point about metaphysics or ethics. In either case, meanings can be understood by grasping the appropriate linguistically mediated, socio-cultural context or horizon of meaning. In a typical interaction, this semantic context is indicated by various factors, including voice and bodily expressions that express serious or playful intent, and normative indicators, such as when, where and to whom we are speaking.

From this perspective, language is not about representing an 'objective' reality coded into the essence of words; rather, language use is a form of activity that has multiple functions. This shift in focus away from the positivist foundationalism of the pre-war years toward natural language analysis has marked a major transformation in the analytic tradition, and signals a series of developments that take us to the limits of the analytic / Classical AI period covered by Lambert.[18]

---

[18] The boundary between the analytic and post-analytic approach to language is crossed when, in the words of John Sowa, we recognize that to "deal with meaning, semiotics must go beyond relationships between signs to the relationships of signs, the world, and the agents who observe the world, interpret it in signs, and use the signs to plan further actions upon the world." (Sowa 1999). A concern with natural language usage thus has led some analytic philosophers toward a socio-culturally inspired convergence with phenomenologically oriented philosophers. Analytic investigation into language had ultimately led to recognition of the importance of socio-cultural context in the framing of meaning, and this led into anthropological and sociological investigations into how people create and maintain these shared linguistic horizons. Inquiries of this type converged with similar sociological and linguistic investigations that had been inspired by the phenomenological tradition.
This convergence was significant in terms of Lambert's fundamental questions, as the inquiries into the nature of intentionality that underpin the work of Husserl and Whitehead were based upon phenomenological examinations of the foundations of the logic, mathematics, physics, psychology and the social sciences; the result was expressed in phenomenologically more adequate metaphysical understanding of reality, and significant new approaches to formulating the conceptual basis of

Several observations emerge from this depiction of analytic philosophy's approach to 'referential linguistic intentionality'.

- Firstly, Lambert's argument is not with intentionality or realism per se, but with the particular philosophical forms of linguistic realism promoted by analytic philosophers.[19]
- Secondly, Lambert targets analytic concepts of intentionality because they falsely assume that the world represented in every day language (potentially) provides a metaphysically adequate description of reality (which when formalized will provide machines with a grasp of reality).
- Thirdly, in suggesting that there are two main ways in which analytic philosophers have attempted to explain the intentional connection between the semantic content of a token and its referent (the external logicist, and internal psychological explanations), Lambert's focus is on difficulties that undermined each position, and solutions that were proposed. His examinations of solutions carry him from positivist to post-positivist versions of analytic philosophy.
- Lastly, Lambert's distinction between the external non-psychological and internal psychological explanations provides his link to classical AI. The former explanatory strategy corresponds with the Knowledge Phase in the AI tradition (key issue was to select *what* the machine needed to know), and the latter strategy with the Representation phase (key issue was *how* to represent what the machine knows).[20]

---

scientific theories in various disciplines, including sociology. These traditions provided a new source of ideas for AI engineers.
Thus, in the early nineties, when researchers were investigating distributed artificial intelligence (DAI) and had begun to consider how to combine autonomous agents into self-organizing, problem solving, multi-agent systems, some AI engineers began to look toward sociological work that had been philosophically oriented by the work of Husserl, Whitehead, Dewey, Mead and others (giving rise to Socionics, as outlined in the previous chapter).
Lambert's questioning of the analytic presuppositions of Classical AI and his appeal for a process philosophy based reformulation of AI should be considered in the context of this mood of growing scepticism about the philosophical presuppositions of analytically inspired approaches to cognitive science and AI, as reflected in works such as Hutchins (1995), Chalmers (1996), Clark (1997), Agre (1997), Clancey (1997), Sowa (1999) and Dourish (2001).

[19] His critique of the metaphysical adequacy of representation in Classical AI is an attempt to show that analytic notions of 'referential linguistic intentionality' are problematic and that associated notions of linguistic realism are false. Thus, for example, he argues that analytic notions of intentionality are an inadequate basis for explaining the connection between "the token 'Buckingham Palace' on this page with the principal royal residence in London and the token 'Norty is a cat' on this page with my cat" (*EMWC*:80).

[20] Lambert (1996) speaks of three phases in the development of the AI tradition, Search [phase 1], Representation [phase 2] and Knowledge [phase 3]. "The Search Phase promoted a systematic study of general purpose heuristic searches principles as sufficient to account for intelligence." (*EMWC*:30) The Representation Phases "promoted a systematic study of general purpose representation principles for encoding knowledge as sufficient to account for intelligence." (*EMWC*:30) The goal was to represent knowledge systematically, either using the declarativist's data form of 'knowing that' or the proceduralist's data form of 'knowing how'. In the Knowledge Phase the concern was with the content of knowledge: "The central problem was that of deciding what the machine needed to know, and how that knowledge was to be acquired" (*EMWC*:34). This spawned two avenues of inquiry; one concerned with domain specific knowledge (formalizing non-expert and expert domain specific knowledge); the other concerned general knowledge practices, such as how we reason under conditions of uncertainty. (*EMWC*:35-6).

Thus, Lambert's focus is on an implicit metaphysical presupposition that he attributes to analytic philosophy, as exemplified in different ways of explaining the intentional connection between word and reality. Lambert contends that analyzing the flawed analytic notions of intentionality that underpinned the development of classical AI helps to explain why it failed to deliver on its promises. He divides his analysis into consideration of non-psychological explanations of linguistic intentionality, and psychological explanations of linguistic intentionality. His goal is to show that two fundamental intuitions have informed these notions of linguistic intentionality (either that tokens are about things in the world, and hence content is reference into the world, or that tokens are about interpretations in the head, and hence content is psychology) and that when the logical implications of each intuition are in turn developed, they disclose weaknesses that seem to point to the validity of the opposite position. This polar tension underpins Lambert's rejection of the classical stance, and his endeavor to resolve this tension leads him to his proposal for an alternative, post-classical approach to representation. The arguments that lead Lambert to his conclusions will be considered in the next two sections.

2.42 Non-psychological explanations of Linguistic Intentionality

Lambert argues that most advocates of Classical AI assume – as per Naïve (Linguistic) Realism – that the content of tokens is derived from what they refer to. Thus words and sentences derive their meaning from their reference to salient features in the external world. One of the problems with this position, discovered by Frege, is that two tokens can have a different content (sense) but refer to the same thing, as in the case where the morning star and the evening star both refer to the planet Venus. Thus tokens can have a differing content but the same referent; hence, contrary to the naïve linguistic realism hypothesis: "Content and reference need not coincide" (*EMWC*:81). A further problem is that of content with a non-existent referent, such as a unicorn; thus: "Content can survive without reference." Third, even when the referents of both subject and predicate do exist, they may not coexist in the real world, resulting in false but meaningful statements, such as "Socrates is my Mother". Hence: "Content requires truth". These problems illustrate that "content, while perhaps being mindful of reference, is not encapsulated by it, and as a consequence, naïve reference fails to explain the miracle of linguistic intentionality" (*EMWC*:81-82).

Thus in order to preserve the basic intuition of naïve realism, concerning the link between language and the external world, it was necessary to develop a more sophisticated explanation of the relationship between content and reference. One way to deal with the previous difficulties is to break the sentence up into parts that do not presuppose the existence of referents. For example, instead of saying 'Socrates is my Mother' [mother(Socrates)], I say there exists some x, where x is called Socrates, and x is my mother [$\exists\, x(\text{Socrates}(x) \wedge \text{mother}(x))$]. Thus the content of the sentence 'Socrates is my Mother' is no longer a reference to the world, but a reference to a logical relationship between two propositions. This creates truth conditions that may or may not be satisfied in the real world, enabling a statement expressed in this way to be both meaningful and false (*EMWC*:83-4). The problem with this approach is that the contents slip from our grasp and there is the danger of infinite regress, such that formal expressions refer to further formal expressions ad infinitum. Lambert analyses two non-psychological solutions to this problem; the first "appeals to intensional entities as the content of sentences", while the second is "an extensional solution which leaves the content of a sentence as a sentence of a known language" (*EMWC*:84).[21]

The latter, extensional solution is developed in Tarskian semantics whereby the conditions that make the content of a sentence true are specified, not by affirming a connection to the external world (as in naïve linguistic realism), but by recursively specifying the truth conditions in terms of a meta-language (eg. 'Snow is white' is a true sentence if and only if 'snow is white' is a true statement within the meta-language to which the sentence refers), and by giving the necessary and sufficient conditions for the truth of the sentences that make up the meta language. Using language to define what language is about sidesteps foundational epistemological and metaphysical questions concerning the truth-claims a sentence makes about reality.

While Tarskian semantics provides a way of sidestepping many thorny issues, as Lambert shows, analytic philosophers have identified various instances that demonstrate that the meaning content of an utterance cannot be resolved simply because the linguist has found a way of specifying the necessary and sufficient truth conditions of the meta-language. Knowing the truth conditions of a language is not the same as knowing what an utterance means.[22] Hence, this approach to

[21] Intension refers to what a term *means*, while extension refers to what it *denotes*. The distinction between intensional and extensional definitions of a term can be illustrated by two opposing ways of defining a sky-scraper; the intensional definition specifies the qualities or properties associated with sky-scrapers, such as their height, their function, the material from which they are constructed, method of construction, and so on, and defines sky-scrapers as the class of all buildings that possess the specified attributes. The extensional definition would list concrete examples of sky-scrapers, such as the Empire State Building, the Rialto Building, etc. and state that sky-scrapers are the class of all such buildings. Intension and extension signal similar distinctions as the paired terms 'connotation and denotation', 'sense and reference', etc. Clearly, meaning determines designation, but is not synonymous with it, as in the famous example, of "the morning star" and "the evening star": both designate the planet Venus, but don't have the same meaning. In classical first-order logic intension plays no role. However, while extensional definitions were adequate for modeling the reasoning used in mathematics they proved inadequate for formalizing the language used in everyday reasoning; hence the development of intensional logics that utilize formal expressions of both designation and meaning and investigates the relationships between them.

Note that 'intension' (with an *s*) and 'intention' (with a *t*) are different concepts; intension and extension are alternate ways of defining words (by intension or extension), while intention is a broader concept that refers to the vector like way in which consciousness of a thought or word is connected to the thing or action in the external world to which it refers. Notions of intension and extension have a specific reference to the semantic content of concepts, while intentionality is a property of consciousness that can refer to thoughts, feelings, perceptions or even emotions, provided they direct the attention of the cognizing subject toward the object eliciting the thoughts, feelings, perceptions or emotions. In the Continental tradition, Phenomenologists, such as Edmund Husserl, have argued that intentionality is the defining feature of all forms of human consciousness.

[22] For example, if an anthropologist learns that a valid word in the language of a people is "Blegallich", and suspects from the example of his informants that this is a true thing to say if it is uttered in the context of the dawning of the Sun (the anthropologist tests this hypothesis by imitating his informant at the next sunrise, and stating "Blegallich"; his informant gestures an affirmation, indicating that it is a true statement in that context), the anthropologist still does not know if "Blegallich" is simply a word for sun rise, or whether it is a ritual statement of reverence meaning "the Sun-god has risen", or any number of other possible meanings. The problem is that assignment of values to a sentence does not deal with contextual and indexical factors (that relate to a specific act of communication) that determine the meaning of a particular utterance. As David Banach has argued, even where we establish relations between sign and object, given that formal systems are "sets of meaningless symbols and can be interpreted as applying to any domain onto which they can be isomorphically mapped", there will be

extensionally specifying the meaning of something has failed to provide grounds for validating Classical AI's metaphysical assumption of linguistic realism.

The alternative to the extensional option suggests that intensional entities define the content of sentences. This approach shifts the focus from the actualities denoted by the sentence to the ideational connotations implicit in the sentence. Three forms of this approach have gained currency in Classical AI: proposition, possible worlds, and abstract situations (*EMWC*:89). Taking each in turn:

- propositions (or 'thoughts' in Frege's language) mediating between tokens and their referent are cited as the meaning of formal expressions[23] and hence "the content of declarative tokens is reduced to a theory of propositions" that is formulated in terms of a 'third realm' that is neither of the mind nor of the world (*EMWC*:90). Lambert describes the 'propositions' of this third realm, mediating between the sense of tokens and their referent in the world, as "little more than ontological flights of fantasy designed specifically to plug problems with linguistic content". (*EMWC*:92)
- possible worlds (Kripke) refer to the distinction between possible outcomes and actual outcomes; the way things might have been as opposed to how it is.[24] "The possible world theorist reduces meaning to propositions understood as function[s] mapping possible worlds to truth values" (*EMWC*:92). The possible world solution shares in the problems that attend the situation solution to explaining the content of sentences.
- situations (Barwise and Perry) resemble possible world solutions, but differ in their more fine grained distinctions between the actual, factual and non-factual; they may also present information that constrains the meaning we derive from other situations. Lambert argues that both the possible worlds and situation approach suffer from a metaphysical irrelevance problem, in that both multiply non-actual worlds or non-factual situations to explain contents that are meaningful yet have no actual referent (*EMWC*:94).

These and other problems to do with identity and relations suggest that intensionalist logical strategies fail to solve the linguistic intentionality problem of explaining how the semantic content of a token is intentionally connected to the actuality to which it refers. The alternative to an external strategy for explaining linguistic intentionality

too many alternative relationships, and thus although "we can determinately interpret our symbols, there will be too many different alternative interpretations". (Banach, 1987: chapts. 7&8).

[23] For example, Frege concluded that propositions, or 'thoughts', define the sense of declarative sentences. A sentence is a series of sounds or marks that express a sense; the truth of a sentence refers to the truth of this sense. According to Frege, 'thoughts' (propositions) are language independent ('Norty es un gato' and 'Norty is a cat', refer to the same proposition), and are neither the mind's ideas nor a thing existing in the outer world; instead they refer to a third realm of propositions that correspond to ideas but are true independently of whether anyone has thought of them, or whether their object exists in our world. The thought, for example, expressed in Pythagoras's theorem, is timelessly true, independent of when it was discovered; it exists, like a planet, in its own realm. In thinking we do not produce thoughts; rather we apprehend them, and their truth or falsity exists independent of my recognition. This third realm of thought objects needs a thinker in order to become effective, yet the thinker does not create them; rather, they are neither wholly real nor wholly unreal, but belong to a special domain of reality that is brought into operation by the thinker (akin to Whitehead's abstract domain of eternal objects). See Frege ([1918] 1956:292-311) and Lambert (*EMWC*:90-91).

[24] When a die is thrown, for example, there is only one actual outcome, but five other possible outcomes. Possible worlds include all the ways the world might have been, including the way it actually is.

(based on extensional or intensional notions of sense) is an internal, psychological strategy.

2.43 Psychological explanations of Linguistic Intentionality

Externalist attempts to explain the intentional connection between token and its content focus on the external logical connection between the token's content and its referent (as in naïve linguistic realism). The alternative form of explanation that has been utilized by Classical AI draws upon the internalist, psychological thesis that the capacity of language to refer to the world is "because minds make the intentional connection on behalf of language" (*EMWC*:97). Internalism grows out of the Cartesian intuition that the phenomenological actuality perceived by a subject is a real state in its own right, and that it makes no difference whether it is a veridical (i.e., something corresponds to it in the outside reality) or a non-veridical state. On this account, the content of the proposition "the arthritis in my leg" is meaningful to the subject, even if a doctor finds it hard to believe.[25] What is important, from the internalist standpoint, is not some external referent, but the role that the concept plays in the mind of the believer; thus, to use the arthritis example, the relevant question is not what doctors say arthritis is, but how I *intend* arthritis when I believe that I have it. Intentional content, from this standpoint, are what my subjective beliefs, desires, fears, hopes and other subjective attitudes, are about.

Lambert's discussion focuses on a range of internalist positions and analyses their shortcomings. A key problem with one form of the internalist position was identified by Hilary Putnam. Putnam proposed a thought experiment based on the notion of a Twin Earth, identical in every way to our own except for one anomaly; although water on both planets appears to have the same properties, the water on Twin Earth has a different atomic structure. Thus, for people on both Earth and Twin Earth, 'water' is experientially identical and from a psychological standpoint, people intentionally referring to water have the same psychological experience. Yet the word water has a different referential meaning content; it refers to the different fluids peculiar to each planet. Hence, though people on both planets experience the same psychological states when they say "it is raining", they do not mean the same thing since the moisture that forms rain on both planets is in fact different. That they had always intuitively appreciated the significance of the external referent would be evident if scientists from one planet visited the other and discovered the unique composition of each planet's 'water'; there would be no doubt in the minds of the scientists concerned that what they referred to as 'water' was now and always had been the water from their own planet. This thought experiment demonstrates, in Lambert's words, that it is possible "for two minds to appropriate identical psychological conceptions toward the same token which nevertheless intuitively possesses different contents by virtue of different environments" (*EMWC*:100). Hence, the internal psychological state of the speaker alone is not enough to explain what a word is about.

---

[25] As a result, psychological explanations are often framed in terms of propositional attitudes, where there is a subject (I), an attitude (hope/believe/fear/) and the intended object (*that* <it is raining/snowing on Mont Blanc/there is arthritis in my leg>) and where the intentional content refers to the object of the *that-clause*.

While the external mode of explanation corresponds with the Knowledge Phase in the AI tradition and was concerned with the content of knowledge, the internal, psychological mode of explanation corresponds with the Representation phase and was concerned with how to represent the content in systematic manner. In the Representational phase of Classical AI the focus of research shifted to the study of representation principles; the goal was to represent knowledge systematically, either using the declarativist's data form of 'knowing that' or the proceduralist's data form of 'knowing how'.[26] The declarativist focused on representing knowledge as data structures and minimized the sophistication of procedural operators, while the proceduralist reversed this focus. Although both strategies attempted to explain linguistic intentionality psychologically, each was based on a different semantic outlook. In Lambert's words:

> For the proceduralist, the content of a mental structure usually comes by its use or potential use. For a declarativist, the content of a mental structure usually derives compositionally from its referents. This referential-use debate is of course a favourite among philosophers, and Classical AI's attempts to psychologically resolve [semantic] content reignites those debates. (*EMWC*:98)

The proceduralist-declarativist controversy can be characterized by the relative narrowness or broadness of the relationship between the internal operations of the mind/computer and the characteristics of the world to which those operations refer.[27] For an extreme psychological reductionist, operations internal to the mind produce semantic content procedurally without recourse to referential properties. Lambert notes that in the Classical AI literature Sloman advances a position akin to this whereby the *narrow content* defined by the 'how' of the computational procedure that solipsistically defines what is referred to, in contrast to the other extreme which involves a *broad content* arising from specifying and relating an internal procedure to facts that define the external context. This allows "reference to dominate content in a psychological setting. This was the position adopted by the declarativists during the Representation Phase." (*EMWC*:106) Pat Hayes and Fodor took this position. By returning to a referential theory of content, though one based on a psychological platform, declarativists such as Fodor explain content in terms of "a reference relation giving rise to truth conditions. Content is broad content." (*EMWC*:107)

---

[26] An intuitive grasp of the distinction between these two approaches is provided by two contrasting ways of giving directions from locations A to B. The proceduralist states, "take the road north from A for 5 kms, turn left," and so on, with precise instructions all the way to B. The declarativist hands over a map showing the roads linking A to B and lets the person work out the best route.

[27] Terry Winograd describes the procedural/declarative controversy as follows:

> It is an artificial-intelligence incarnation of the old philosophical distinction between 'knowing that' and 'knowing how'. The proceduralists assert that our knowledge is primarily a 'knowing how'. The human information processor is a stored-program device, with knowledge of the world embedded in the programs. What a person (or robot) knows about the English language, the game of chess, or the physical properties of his world is coextensive with his set of programs for operating with it . . . The declarativists, on the other hand, do not believe that knowledge of a subject is intimately bound with the procedures for its use. They see intelligence as resting on two bases: a quite general set of procedures for manipulating facts of all sorts, and a set of specific facts describing particular knowledge domains. In thinking, the general procedures are applied to the domain-specific data to make deductions. Often this process has been based on the model of axiomatic mathematics. The fact are axioms and the thought process involves proof procedures for drawing conclusions from them (Winograd,1975:186)

Lambert examines a number of positions midway between the narrow and broad psychological explanations of intentionality and concludes that none of the explanations of the miraculous connection of linguistic intentionality, between token and the world, is without problems. Hence, he concludes that "with the semantic motivation for the metaphysics of the manifest image in jeopardy, the metaphysical adequacy of Classical AI's linguistic representation must be equally problematic" (*EMWC*:112).

2.44 The failure of Linguistic Realism: why Lambert rejects the Linguistic Turn.

Since Classical AI inherited the analytic outlook it also inherited the analytic movement's inadequate notions of linguistic intentionality. As described above, both the psychological and non-psychological explanations share certain attributes. The two basic intuitions that have informed linguistic intentionality assert that either:

- tokens are about things in the world, and hence content is reference into the world, or
- tokens are about interpretations in the head, and hence content is psychology.

The apparent polar relationship between these positions was undermined by weaknesses that led each back toward its opposite. Thus, the former non-psychological position encounters Frege's demonstration that content need not always coincide with reference ("morning star"/ "evening star" = venus). This leads to intensional entities that are used to qualify extensionally based meanings, which has the effect of dislodging content from the world. The latter psychological position encounters Putnam's (1973) classic 'twin-earth' demonstration that content needn't always coincide with psychological states (Earth water & Twin Earthwater $\neq$ "water").[28] Lambert concludes that just as Frege can be used to dislodge content from the world, Putnam can be used to dislodge content from the head (*EMWC*:113).

If the failings described in the previous two sections are genuine (externalist and internalist arguments for linguistic realism), then linguistic realism is false, and Classical AI's theories of commonsense are metaphysically inadequate. This is tantamount to a rejection of the 'linguistic turn' and "the analytic vision that philosophy can be understood through language" (*EMWC*:112). However, the failings of these analytic versions of linguistic realism provide a valuable clue as to how to construct a notion of representation that meets the criteria of metaphysical adequacy.

Lambert explains the metaphysical tensions mirrored in psychological and non-psychological explanations of content by recognizing that intentionality is anchored in attributes of both mind and the world. He argues that fluctuations between broad and narrow content are possible only "if both world and psychology are able to independently influence content, with the world contributing what we view, and the psychology contributing how we view" (*EMWC*:115). Either can overshadow the other, depending on the situations and the succession of communications between

[28] Putnam's experiment is like a mirror of the planet Venus 'Morning Star' / 'Evening Star' experiment. The latter demonstrates that external reality alone cannot account for meaning, while Putman's thought experiment demonstrates that internal reality 'Water' fails to account for the different intentional contents in Earth / Twin Earth. Hence, Putnam's experiment suggests that external reality is intrinsic to the meaning content of intentional acts and highlights the failure of the declarativist's 'knowing that' or the proceduralist's 'knowing how' to explain the meaning content of words.

people. For example, if we define a token 'skyscraper' intensionally, using terms such as 'tall building', the stated properties might not pick out all and only skyscrapers. While if we appeal to an extensional definition, all buildings that belong to the Natural Kind exemplified in the class of skyscrapers might not conform to what the public in a particular city recognize as a skyscraper. The first example rejects narrow explanations of content by demonstrating that such explanations can result in a broader reference than was intended. The second rejects broad reference as an explanation of semantic content on the grounds that it fails to explain the narrower range of references utilized in practice. This is possible, Lambert argues, because in reality, language content is not fixed. Rather, as the examples cited illustrate, the referential 'what' of the world and the psychological 'how' of the mind are interdependent factors in determining valid linguistic content (or sense). Broad referential content constrains narrow psychological precepts; the latter must match the world, but only to a limited extent, as the broad content is only ever presented to us in primitive information that correlates in an imperfect fashion with the narrow content of the psychological precept.

2.45 Why language functions adequately despite the metaphysical inadequacy of linguistic realism.

Lambert argues that language works because of a 'charitable contract' established between speaker and audience in the course of the communication, to bridge differences and inconsistencies.[29] For this to work there must be a "sufficient overlap between what is viewed and how it is viewed." (*EMWC*:116) Manifest image realism would attempt to anchor the meaning of the token 'cat' in a causal structure of perturbations corresponding to an experience of actual cats.[30] Word and object are tied together intentionally, such that the world disclosed in language is in some way grounded in, and disclosing the metaphysical actualities of the world. The analytic tradition hoped to make the form of these intentional connections explicit and developed theories positing external or internal linkages that provided logical or psychological explanations as to how valid assertions were grounded in reality. However, language appears to be far more flexible than this, with words capable of extension in infinitely many ways. Hence Lambert concludes that "Common language terms do not directly anchor into the representative causal structures responsible for intentionality, and in that sense, manifest image realism is a false doctrine which

[29] Lambert's notion of a charitable contract is indebted to Willard Quine and his reflections on the empathic stance that must be taken while translating. From this standpoint, when translating another person's utterances into our own language, we should avoid imputing to them apparent absurdities or irrationalities implied by our literal translations, and instead seek to draw from their words a meaning that we can maximally agree with. We do this by making the charitable assumption that our difficulties arise from our ignorance of the other's language, and compensate by projecting our own rational sensibilities onto the other (Quine 1960:59).

[30] Lambert argues that the world presents itself as perturbations, in the form of causal influences that manifest structural patterns. Broad content only consists of these perturbations, and delivers information directly in the form of these causal structures, not the objects posited by the manifest image model of the world. If a similar approach is adopted toward the narrow content of the mind, then mental precepts are also reports of causal structures. "We can have causal structures in the world, similar representative causal structures in the mind, and homomorphisms of varying exactitude identifying intentionality. The proposed basis for intentionality is then not greatly different from the familiar forms of maps, portraits, and sculptures, since it too is then reliant on structural homomorphism." (*EMWC*:116)

obscures the basis upon which intentionality involving language works." (*EMWC*:117) Hence, the manifest image cannot be used as a metaphysical basis for engineering commonsense. The brittleness problem is symptomatic of this.

This does not mean that common language should be ignored. The proper function of common language, Lambert argues, is not that of opening a metaphysical window on to the world. Rather, its merit is "that it promotes efficiency and flexibility at the expense of detail and accuracy. . . Language, as an assembly of sounds and shapes embedded within causal structures to which we are attuned, serves as an efficient indexing mechanism for rapidly selecting sufficiently matching mental structures which can be adapted to the circumstances confronted. . . . Common language has flourished, despite its inaccuracies, because of this efficient indexing capability" (*EMWC*:117). Lambert's goal, in his Post Classical AI, is to weave together this functional view of language with a process view of mind and world.

2.5 The flawed philosophical presuppositions of the Classical approach to Engineering Commonsense

Lambert concludes that there are a number of shortcomings associated with the Classical approach to engineering machines with commonsense. When assessed against the McCarthy and Hayes criteria of representational adequacy, Classical AI is:

1. metaphysically problematic, in that it adopts the position of linguistic realism, and assumes (incorrectly) that common language provides a metaphysical window to the world, and hence that common language provides the foundations for a metaphysically adequate view of reality (the linguistic realism fallacy).
2. epistemologically inadequate, due to the presupposition that the problem solving procedures of commonsense are rationally transparent and expressible as rational calculations, and hence that human problem solving capabilities can be implemented in a machine using the same rational procedures (the rationalist fallacy).
3. heuristically inadequate, due in large part to the solipsist presupposition that the heart of a commonsense machine is a rationalist system of knowledge representation that enables an agent to deduce the nature of the world independently of the world (the solipsism fallacy) (*EMWC*:118).

Lambert's list of the inadequacies of the classical approach to representation identifies problems that are not "repairable superficial anomalies"; rather, the fallacies identified are symptoms of deep-rooted problems inherent in the philosophical presuppositions informing the classical understanding of human intelligence and of potential strategies for its machine implementation. That is, the inadequacies in Classical AI's approach to symbolic representation are a function of the classical mode of understanding, "since it is the classical understanding that promotes classical representation as a means of both understanding and engineering commonsense machines, and it is the classical understanding that promotes the rationalism, solipsism, and linguistic realism in classical representation" (*EMWC*:119).

Lambert, following Philip Agre (1993) suggests that research communities arise as a consequence of largely tacit shared understandings, embodied in the vocabulary, methods and values that together constitute the community's 'worldview'. Scientific worldviews are rooted in philosophical presuppositions concerning the nature of existence, mind, meaning, method, knowledge and truth. These presuppositions and their intertwined dependencies are communicated to the members of a research community through a dialectical interplay of the questions asked and answers proposed within the public discourse of the community "with each question and answer potentially inciting a further instance of the other" (*EMWC*:121). The ways in which these presuppositions are woven into questions and answers reflect the dogmatic biases of the particular worldview. Without such dogmatic foundations any attempt to communicate would be subject to a skeptical regression, resulting in communication failure. Rationalism, solipsism and linguistic realism are expressions of a dogmatic orientation that for the most part has been accepted by the Classical AI research community. This orientation reflects the principles of understanding that have defined the shared horizon of Classical AI.[31]

Lambert follows Michael Dummett (1994) in suggesting that changes in the history of Western philosophy were defined by foundational biases manifested in the questions asked and answers provided. Over time, skepticism undermines the dialectical rhythm of an era, leading to new dominant questions and chains of dependency in the principles of understanding underpinning a worldview. Hence, the Ontological Age, exemplified by Aristotle, secured its dogmatic basis by querying *what there is*; then the Epistemic age, exemplified by Descartes, reached its dogmatic conclusions by questioning *what is known*; and lastly, the Age of Language, exemplified by Wittgenstein, reached its dogmatic conclusions by pondering *what is represented* (*EMWC*:122).

Lambert argues that Dummett's analysis can be used to differentiate the worldviews that can inform AI. On this basis it is possible to define Ontological AI, Epistemic AI and Semantic AI. Each has its dogmatic foundations and chains together a dependency structure that link *what there is*, with *what is known* and *what is represented*. On this basis Lambert is able to relate phases in the development of Classical AI to a shift in orientation from Epistemic to Semantic AI. Hence, from this perspective, possible orientations to AI include:

- *Ontological AI (object oriented view – more important in mainstream computer science than in Classical AI): what there is* → (grounds) *what is represented* → (grounds) *what is known [Aristotelian AI]*
- *Semantic AI (representation phase in Classical AI): what is represented* →(grounds) *what there is* → (grounds) *what is known [Wittgensteinian AI]*
- *Epistemic AI (knowledge phase in Classical AI): what is known* → (grounds) *what is represented* → (grounds) *what there is [Cartesian AI]* (*EMWC*:123).

Lambert then puts these classifications to work in both explaining the failure of Classical AI and in justifying his alternative vision of a Post-Classical AI. In summary form, his investigation discloses two competing standpoints, "a less celebrated object

[31] According to Lambert, a community's principles of public understanding rest upon dogmatic foundations, however skepticism remains the permanent partner of dogmatism, such that the foundations of a worldview may prove transitory, with key features of a shared horizon in time perhaps receding to be surpassed by another (*EMWC*:121).

oriented proposal, and a commanding epistemic Cartesian AI harboring Wittgensteinian-like sentiments" (*EMWC*:129).

The object oriented proposal, Lambert argues, has Aristotelian foundations, and possesses several drawbacks. First, its ontology is based upon the notion that the objective nature of things can be recovered from the 'manifest image' of things implicit in common language. Second, it reduces relations to the properties possessed by specific things and is unable to express relations that exist but are not possessed by any particular thing. Third, Aristotelian foundations provide a poor basis for dealing with change (*EMWC*:129-130).

Cartesian AI is also beset by problems, as manifested in Lambert's analysis of the problems arising from rationalism, solipsism and linguistic realism. As previously argued, rationalism is a flawed reconstruction of human problem solving behavior. Likewise, solipsism mirrors the Cartesian mind-body divide, and fails to capture the vital role of our environment as the context mediating our embedded commonsense capacities for judgment and action. Lastly, the analytic influence in Cartesian AI is equally unsatisfactory. Reflecting the bias of Wittgensteinian AI, it promoted linguistic realism and endorsed the notion that the manifest image of the world implicit in common language provided an adequate metaphysical foundation for commonsense machines.

The source of difficulty with Cartesian AI, Lambert argues, extends beyond the philosophy of meaning implicit in its notion of representation to the Cartesian philosophy as a whole. Its epistemological foundation in a particular approach to subjective knowledge, the role of rationalism in deducing knowledge from these subjective foundations, the mind-body dualism implicit in this epistemology along with its attendant solipsism, all find forms of expression in Classical AI. Hence, Lambert's conclusion that it is "the Cartesian understanding as a whole that denies Classical AI its success" (*EMWC*:124-131).

2.6 Toward a Post-Classical AI

Lambert's critique of Classical AI was driven by the practical concerns of a computer engineer. This is an important factor to keep in mind when comparing his critique of representation in Classical AI with critiques motivated by philosophical concerns (such as Dreyfus), or the psychological concerns of cognitive science (such as Bickhard). Lambert's aim was not to advance an alternative philosophy of AI, or to promote a new approach to cognition; rather, his motivation was to find strategies for overcoming engineering problems associated with the Classical paradigm.

Lambert's aim was to uncover why the Classical paradigm's model of human 'commonsense', and its related strategies for formally modeling how people make sense of, and solve everyday problems had ended in failure. To that end he identified flawed philosophical presuppositions that have guided the classical endeavor to capture the formal essence of commonsense. For Lambert, the big philosophical questions (what is the nature of reality; what are beliefs; how is a true belief related to the nature of reality; what role do true beliefs play in the human capacity to propose commonsense solutions to real problems?) are a means to an end: that of rethinking the nature and role of representation and rationality so as to promote a different model of artificial intelligence.

In proposing his Post-Classical AI paradigm, Lambert attempts to reconfigure the classical presuppositions concerning the interaction between mind, method, representation, knowledge and existence, while retaining the computational technique developed by the Classical AI research community. In seeking a new, process oriented worldview, he rejects the historical alternatives implicit in Aristotelian, Wittgensteinian and Cartesian versions of AI. As will become evident in later chapters, Lambert

> . . . subscribes to neither the age of epistemology account of understanding, nor its established age of ontology or age of language alternatives, preferring instead a non-well founded doxastic and methodological interdependency. Understanding is not the product of a hierarchy rooted in knowledge, nor existence or meaning. It is about interplay between dogmatism and scepticism. It is about balancing a network of tensions over time, between an evolving world of processes (the methodological) and embedded individuals who form beliefs about that world (the doxastic). (Lambert & Scholz, 2005:10)

In *EMWC* Lambert engages in a sustained attempt to show, in practical detail, how the difficulties associated with the classical notion of representation, as described in this chapter, can be addressed by adopting a perspective derived from process philosophy. Process philosophy provides Lambert with the conceptual resources for re-negotiating the extremes of the classical and situated paradigms, and of arriving at an understanding of representation congenial to a socially situated view of intelligence that is consistent with the Socionic motivations of my project. While this sociological reading of AI is not made explicit in *EMWC*, there is a strong implicit recasting of AI in a sociological vein. Hence, *EMWC* is a pivotal work in establishing process-philosophical foundations for a post-classical AI congenial to the development of the Socionic vision. The next chapter will introduce the process philosophical principles and presuppositions that frame my reading of Lambert's Post-Classical paradigm.

Chapter 3

A Whiteheadian Process alternative to the Cartesian world-view of Classical AI

Chapter Three – Section Headings

3.0 Introduction: Lambert's process alternative to Classical AI

Lambert's proposal for *Engineering Machines with Commonsense* involves two parts. As outlined in the previous chapter, the first part of *EMWC* involved a wide ranging consideration of the adequacy of the epistemological, heuristic and metaphysical presuppositions of Classical AI. The focus of Lambert's analysis was on Classical attempts to model the presumed role of rationality and language in intelligent behaviour. His conclusion was that the Classical approach to AI was based upon a false understanding of human commonsense capabilities; it had followed a dead-end path inspired by a false, Cartesian notion of human rationality and a mistaken hope that natural language, when broken down into Wittgenstein's logical atoms, could capture the reality of the world.

The second part of *EMWC* has a different flavour. It is not a wide-ranging consideration of the philosophical alternatives advanced by critics of Classical AI. There is no consideration as to why, for example, process philosophy provides a better prospect for a Post-Classical AI than Heideggarian AI.[1] Instead, with little attempt to justify his choice Lambert adopts process philosophy as the philosophical foundation of his Post-Classical paradigm, and focuses on clarifying the engineering implications of his choice. To that end Lambert shows how, beginning with a formal ontology inspired by process metaphysics, it is possible to formalise Post-Classical notions of representation and belief and then use this logical tool box to fashion the architecture of an artificial mind.

Lambert's exposition can be viewed as a practical demonstration of the process philosophy alternative to the Classical understandings of representation and rationality. Rather than engage in a general discussion of the epistemological, metaphysical and heuristic implications of process philosophy for our understanding of cognition and commonsense, Lambert attempts to 'cash-in' on the process perspective by providing a formal blue-print for building a Post-Classical multi-agent mind. Lambert does not develop a software implementation of that design; rather, he provides a step by step formal outline and 'proof of concept' of his Post-Classical strategy. This strategic blue-print takes the form of a series of chapters that are akin to a set of instructions for an exceedingly complex lego® block model, showing which blocks you begin with, how you clip them together, and so on, with a focus on the minutiae of the how and why. The style is akin to a mathematical treatise; it provides the technical details of an alternative strategy, but lacks a discursive overview of process philosophy and its implications for the attempt to engineer commonsense.

If Lambert was treading a path within a well known paradigm, the reader of *EMWC* would understand the vision and guiding presuppositions motivating Lambert's proof by demonstration strategy. However, that is not the situation. The AI community has virtually ignored the modern process philosophy tradition. Hence, Lambert is promoting a radically different AI alternative, rooted in a fundamentally different philosophical vision than those that he discussed in his critique of Classical AI. What the reader of *EMWC* searches in vain to discover is a concise outline of the alternative philosophical vision that frames and informs Lambert's Post-Classical paradigm. Instead, they must tread an arduous path, first researching the nature of process philosophy, and then carefully unpacking Lambert's 'proof by demonstration' to find

[1] Although Lambert refers to Dreyfus's work in his critique of Classical AI, and to the work of AI researchers inspired by Dreyfus's Heideggarian approaches to AI, such as Winograd, Brooks and Agre, his advocacy of a Post-Classical AI does not involve reviewing the philosophical presuppositions that have framed these Post-Classical alternatives. Dreyfus (2007) reviews past developments and current prospects of Heideggarian AI.

the implicit philosophical presuppositions that Lambert has drawn upon to guide his alternative Post-Classical engineering strategy.

Grasping a new paradigm is always a difficult exercise, as it means breaking free of entrenched conceptual habits, and finding a new standpoint from which to evaluate a situation. That is why innovative scientists seek easily grasped metaphors that communicate clues and awaken curiosity about the nature of the intellectual world that their new paradigm seeks to disclose. Why embark on a demanding voyage without a prior sense of where you are going and how you are going to get there? Columbus's practical endeavour to find a new route to the Indies would not have made sense if he had still believed in a flat earth. He needed the vision of a spherical Earth. Lambert had an alternative vision of how to engineer commonsense; a vision that guided and inspired his set theoretical formulation of a Post-Classical artificial mind. However, it is difficult to find. His 'proof by demonstration' is akin to having the plans and pieces of a lego® kit without the inspirational picture on the front of the box. In his blue-print for a multi-agent system (MAS) we see the fruit of that vision, but the vision itself is largely implicit in the detail of the design; there is no summary statement of the presuppositions and guiding ideas of the Post-Classical philosophical perspective.

What then is Lambert's alternative philosophical perspective? Although his work contains very few specific references to the work of process philosophers, he is quite explicit in his debt to the process philosophy tradition initiated by Heraclitus.[2] As with Heraclitus, the basal events of Lambert's universe are processes. Lambert uses set theory to develop a formal language for defining processes, and for specifying the valid sentences that can be used to describe his formal process universe. Process is then used to explain representation; representation to explain machine beliefs; and beliefs to explain a machine version of mind. Lambert's basic strategy is a process inspired version of *Ontological AI,* as outlined at the conclusion of the previous chapter.

Process thus provides the ontological foundation for Post-Classical AI, however consideration of the philosophical implications of this move remain secondary to the engineering objectives of the thesis; namely, developing a mathematically valid model of mind that can be implemented as a computer program. This perhaps explains why Lambert makes few references to process philosophers or their works.[3] Yet

---

[2] Apart from a quotation from Hartshorne and an aside on Heraclitus (see next chapter), the sole explicit reference in *EMWC* to process philosophers occurs in a footnote where Lambert writes:

> Exploiting the concept of process is no novel idea. Several important philosophers have founded outlooks based upon the notion of process, among them Bergson, Peirce, Dewey and Whitehead (*EMWC*:136).

In subsequent publications he also makes references to Whitehead and process philosophy, but these references are simply to acknowledge the orientation informing his work.

[3] Lambert has explained the lack of explicit reference to philosophical sources of inspiration, as follows:

> [In *EMWC*] I have few references to processes philosophers. This is true. The reason, in case you are interested, is that for most things I do not read the literature and then formulate my views based on a fusion of those ideas. I instead tend to think about things and then formulate my views. Then I go to the literature and see what I can find that relates to my ideas and might refine those ideas. This was the case with the process philosophy, which I conjured up in its first order logic form in a garage while living at Port Adelaide. I then went to the literature and discovered Heraclitus, etc. – hence the lack of detailed references. (Lambert, 2009b: 1)

Lambert has since affirmed that he was then and remains “strongly committed” to process philosophy, and that it has been “extremely valuable at a practical level” in the development of the Post-Classical vision first expressed in *EMWC* (Lambert 2009b: 1).

Thus, Lambert’s primary concern in *EMWC* is not to promote a philosophical vision. The reader must go elsewhere to learn about process philosophy. His work does, however, demonstrate how process philosophy can be deployed to underwrite an alternate approach to AI. This concrete demonstration, rather than the implicit motivating vision is his primary contribution. Hence my aim is not to engage in conjecture as to Lambert’s understanding of process philosophy, but rather to demonstrate how process philosophy can be used to provide an understanding of what Lambert has achieved. I shall therefore interpret Lambert’s project through my own vision of process philosophy and attempt to provide the missing discursive outline of Post-Classical AI’s solution to engineering commonsense. This process philosophical horizon not only provides a context for making sense of Lambert’s work; it also provides a shared philosophical context for linking Lambert’s project with my Socionic concerns.

To that end I will outline a Whiteheadian ‘process alternative’ to the Cartesian presuppositions of Classical AI. I will begin with a brief sketch of the contrasting Cartesian and Whiteheadian philosophical visions, and then provide a more detailed analysis of Whitehead’s metaphysical system. This analysis will lay out the set of philosophical presuppositions that underpin a process alternative to the Cartesian explanation of our ‘commonsense’ problem solving capabilities. I will then provide a brief over-view of how these philosophical presuppositions are put to work in Lambert’s Post-Classical proposal for engineering commonsense. Subsequent chapters will examine in greater detail how Lambert goes about realizing the promise of this process approach.

3.1 Process Ontological foundations for AI

As outlined in the previous chapter, Lambert appropriated Dummet’s (1994) analysis of the foundational biases that have informed the development of Western philosophy and suggested that they find their echo in the development of Classical AI. He noted that the Ontological Age associated with Aristotle expressed its philosophical bias by adopting as its dogmatic foundation a substance-based ontology and argued that this bias informed a variant of Classical AI that he refers to as *Aristotelian AI*. Lambert’s Post-Classical strategy for engineering commonsense is a process inspired version of ‘Ontological AI’, whereby the ultimate building blocks of reality are not substances but processes, and this different foundational presupposition concerning the nature of existence supports different notions of representation, belief and mind. The result is a process version of the chain of conceptual dependency, where an ontology of *what there is* → (grounds) *what is represented* → (grounds) *what is known.*

In order to highlight the distinctive consequences of adopting process as the ontological foundation, I will outline linkages between the Aristotelian and Cartesian versions of the chain of conceptual dependency, and contrast these with a process philosophy based sketch of the same chain. For the sake of providing an identifying tag, I will label the presuppositions underpinning the Classical AI as the Aristotelian-

Cartesian paradigm. This reflects Lambert's contention that Classical AI presupposes an Aristotelian metaphysical outlook allied to Cartesian notions of rationality. Likewise, I will label the philosophical presuppositions underpinning my reconstruction of the Post-Classical form of Ontological AI as the Whitehead paradigm, as my framework for interpreting Post-Classical AI is grounded primarily in Whitehead's process philosophy.

This sketch will begin by focusing on the guiding metaphors of Aristotelian and Cartesian philosophy, and then contrast these with the guiding metaphors employed by Whitehead in developing his process philosophy. The core concern will be to illuminate distinctions concerning their respective approaches to representation and commonsense.

3.11 Substance-based presuppositions of Classical Ontological AI and the bi-furcation of reality

The Aristotelian version of the chain of conceptual dependency begins by explaining the ultimate nature of existence in terms of Aristotelian substances. The world is viewed as an objective, substance based domain that is independent of the knowing subject. Autonomous units of substance form the ontological foundation of this system. Every substance has its own unique primary attributes, and it retains the identity of its primary attributes even while its secondary attributes (such as accidental qualities and relations) vary. Thus, Aristotelian substances were imagined to be permanent, unchanging grounds of the changing stream of events. This metaphysical foundation remained influential in Cartesian philosophy. In the Cartesian-Newtonian vision of metaphysical bed-rock, the unchanging primary form of substance was reduced to spatial and temporal locations of bare mass. Hence, the Newtonian Universe could be adequately represented as a sequence of states; each state being a snap shot of spatially distributed particles of matter, with successive states reflecting changes in the spatial configuration of matter through the passage of time.

In this Aristotelian-Cartesian-Newtonian paradigm the prototypical example of representation is vision: through vision the activities that take place in the world are mirrored like pictures in the mind's eye. This mirroring relation is possible because the material world and the mind have a similar substance-attribute structure. Language can also represent the world, as it too reproduces the substance-attribute view of the world in the subject-predicate structure of its representations.[4] In the previous chapter I dealt with some of the difficulties that modern logicians have encountered in bridging the metaphysical gap between the intentional act, and the intended object of that act. Whitehead had focused on the same problem.

Whitehead argued that there is a 'fatal' gap in the Cartesian 'theory of representative perception' between "mental symbol and actuality symbolized" (*PR*:76). The theory of substances asserts that minds and natural entities are different kinds of particulars, and suggests that the perception of a natural entity, such as the sun, is akin to a picture in the mind. However, Descartes could find no satisfactory way to explain how the

[4] Aristotelian philosophy promoted the subject-predicate dogma, whereby a human mind (substance of the mental type) encounters another substance of the dog type (dog) and then expresses this idea via a universal that predicates the subject (actual dog) with its idea of a dog (universal).

object of a perception is present before our mind's eye. At times he implied that the sun itself exists in the mind, but this led to insoluble epistemological difficulties, so he fell back to the bi-furcated position of "the mind with its private ideas which are in fact qualities without intelligible connection to the entities represented" (*PR*:76). Whitehead concluded that Cartesian thought fails to explain how the ideas that arise in mind are connected to the object that they purport to represent; hence, the theory of "'representative perception' can never, within its own metaphysical doctrines, produce the title deeds to guarantee the validity of the representation of fact by idea" (*PR*:54).

The consequences of this philosophical outlook was, in the words of Whitehead's oft-quoted passage, "the bifurcation of nature into two systems of reality"; on the one hand, nature as the objective realm, made up of particles of matter, and on the other hand nature as we perceive it, as a realm of subjective feelings and appearances (Whitehead, 1920:31). The logical consequence of this vision was that qualitative values are not real properties of Nature; rather, they are attributes of the human mind. From the standpoint of the prevailing substance based metaphysic, we should not give credit:

> [to] the rose for its scent; the nightingale for his song; and the sun for its radiance. The poets are entirely mistaken. They should address their lyrics to themselves, and turn them into odes of self-congratulation on the excellency of human mind. Nature is a dull affair, soundless, scentless, colourless; merely the hurrying of material, endlessly, meaninglessly." (Whitehead, 1926: 54).

Thus, the Cartesian-Newtonian vision built upon and simplified Aristotle's ontology, but created a chasm between subjective awareness and the objective world. Aristotle's pantheon of physical substances, each qualified by a unique formal essence, was reduced to configurations of matter, defined in terms of spatial extension, time and mass. All else was a function of the human mind. Nature and Spirit were divorced, but as in the Classical Greek vision they remain aligned within a divine order. Manifesting this order, Reason or Rationality was both the law-giver to Nature, and the lynch-pin of human freedom. Although we perceive Nature through the subjective and at times distorted filter provided by our senses, our Reason draws upon a realm of transcendent logical principles and ideas to penetrate this fog of appearances and disclose the true form of the objective realm.

Thus, in the Cartesian paradigm our understanding of the world is reliant upon the human mind and its rational capabilities. The Cartesian mind is like a cabinet, full of symbolically coded information; through manipulating these symbols we represent the state of our world and anticipate its transitions. Through an inexplicable process of co-ordination, mental operations of the mind and physical operations of the brain are correlated, and translated into bodily action. Human agency and freedom, from this perspective, is a form of mental causal determination that our reason exercises through its transcendence of the physical world, and its conceptual grasp of that world. The mathematically formulated empirical sciences have strengthened our conceptual grasp of the world. The scientific process of discovery is the basis of the ongoing 'purification' of our linguistically coded store of 'commonsense' knowledge, and is fundamental to our capacity to make valid judgements about the world.

In sum, the chain of conceptual dependency in Aristotelian and Cartesian-Newtonian philosophical paradigms have their starting point on different sides of the matter-mind equation which begins with self-sufficient substances and the secondary properties and accidental attributes that these substance convey to the mind and ends with mental representations that express a rationally purified linguistically articulated model of the world. Though Aristotelian philosophy reflects an ontological bias, and Cartesian philosophy an epistemological bias, both philosophies share a common faith in the substantial foundations of our world, and the capacity of language to mirror that reality. Hence knowledge, in the form of valid propositions can capture the essence of reality; and these beliefs form the store of knowledge which the mind manipulates according to rational rules of inference in order to make 'commonsense' decisions. This, in simple terms, is the Aristotelian-Cartesian philosophical background that underpinned the Classical AI attempt to computationally engineer commonsense.

3.12 Process-based presuppositions of Post-Classical Ontological AI – an initial outline of Whitehead's resolution of the bifurcation problem.

The Aristotelian-Cartesian paradigm that underwrites the Classical model of commonsense is fundamentally flawed. At the heart of the philosophical inadequacies detailed in the last chapter are questions concerning the nature of representation, and the role of language in mediating human commonsense capabilities. Whitehead's 'bifurcation' critique addresses similar questions concerning the relationship between representation and the thing represented. How am I as the embodied subject in this physical room, related to the mental room in my mind's eye? How do symbols assist in the process of connecting a mind to the thing intended? What is the basis of our commonsense capacity to deal with the diverse situations that occur in real life? Whitehead's process philosophy suggests an answer.

Whitehead pioneered a new approach to metaphysics. John Dewey, a contemporary of Whitehead's and significant process philosopher in his own right, argued that Whitehead had "opened an immensely fruitful new path for subsequent philosophy to follow" (Dewey, 1951:659). Dewey notes that traditionally, philosophers based their doctrines "upon what each one happened to regard as 'first principles' in their capacity of coming logically first"; these principles were viewed as the immaculate gift of God, or Reason, and there was no acknowledgement of the philosophical background of presuppositions "which determines the function which they [the postulated first principles] perform" (ibid:643). Dewey argues that what sets Whitehead's metaphysical system apart from this tradition was his consciousness of the need to be explicit in stating "the nature of the region from which he sets out" (ibid:644). Whitehead does this, Dewey argues, through his critical appeal to the evidence of human experience; experience that is "rich beyond the possibility of exhaustion and subtle beyond the reach of human wit" (ibid:646).

A principle concern driving Whitehead's examination of human experience was the bifurcation problem.[5] He shared the disquiet felt by poets such as Wordsworth and

[5] From Whitehead's perspective, the bifurcation problem was more than an irritating epistemological flaw. He believed that the Cartesian reformulation of Aristotelian metaphysics had created a Natural Cosmology that stripped value from Nature, and a Humanism that deprived human values of any objective ground in the real world. This offended Whitehead's basic intuitions concerning the nature of reality. He was a child of the Romantic era and believed that Mind and Nature, subjective value and

Shelley towards a mechanistic Cosmos in which the stars "blindly run" and believed that the Romantic vision of an internally connected universe provided a more authentic witness to an organic unity encompassing Mind and Nature. Citing the lines from Shelley's poem, *Mont Blanc:*

The everlasting universe of Things
Flows through the Mind . . .

Whitehead argued that Shelley is an "emphatic witness to a prehensive unification as constituting the very being of nature" (Whitehead 1926:106). Whitehead sought to transform this vision of the 'prehensive unity' of nature into a metaphysical system.

As Dewey notes, rather than privileging either Mind or Nature, or accepting some form of substance-based dualism, Whitehead sought to combine both sets of evidence, seeking to interpret human experience as a specialization of traits of Nature and, turning this around, to interpret the events of nature in terms derived from human experience. Whitehead's core premise, Dewey argues, is that "the direct evidence as to the connectedness of one's immediate present occasion of the experience with one's immediately past occasions can be validly used to suggest categories applying to the connectedness of all occasions in nature" (Whitehead [*AI*:284], quoted by Dewey, 1951:647). Dewey contends that by building upon this premise, Whitehead was able to develop a system of thought that resolves "the problem of the subject-object relation, and the problem of the discrete-continuous, or individuality and relativity" (ibid:650).

Thus Whitehead developed a solution to the bifurcation problem; but his solution was a hard pill for most philosophers to swallow. Today, outside of a relatively limited circle of admirers, Whitehead's system of thought is a closed book. However, as the Belgian philosopher of science, Isabelle Stengers recently stated, Whitehead "is still a philosopher for our time" because he squarely faced the issue most philosophers evade: that is "the bifurcation of nature into a causal, or objective nature, as it is explained in scientific terms, and an apparent nature, the nature as we perceive and feel it" (Stengers 2005:3). In the light of Lambert's project, Stenger's remarks are confirmed. Whitehead's proposal for solving the bifurcation problem is still pertinent to outstanding problems such as those addressed by Lambert. However, Whitehead's philosophy is difficult to grasp.

The set of concepts that Whitehead used to explicate the nature of process is of bewildering complexity.[6] The challenge in dealing with Whitehead's thought is to push on through the forest of conceptual tree-trunks until a vantage point is attained from which it is possible to survey the forest. In negotiating this forest I will tread a minimalist path, introducing those elements that are central to Lambert's project; for

---

objective fact were part of the same unified reality. From his perspective, the bifurcated view of reality does not reflect the reality of human experience; rather, it is an artefact of the Aristotelian-Cartesian paradigm. By changing the metaphysical ground-rules of philosophy and developing new concepts capable of more accurately representing the full spectrum of human experience Whitehead believed that the bi-furcation problem could be dissolved.

[6] Frederick Copleston, whose nine volume *History of Philosophy* is one of the great intellectual achievements of the last century, was a consummate master of philosophical interpretation and compression, however after praising the importance of Whitehead's work he apologetically noted that it was "far too complicated to summarize" and that after some consideration he had "decided not to make the attempt" (Copleston, 1985 [vol.8]:400).

those seeking further detail, I have provided extensive footnotes.[7] My selective presentation of Whitehead's thought will be shaped by a 'mountain top' goal, and a guiding metaphor; the goal is to solve the bifurcation problem, and the guiding metaphor will be an action-based view of society. The next couple of paragraphs will introduce the social metaphor and provide an intuitive sense of how Whitehead's project differs from a substance based metaphysics.

Instead of viewing reality in terms of enduring rock-like substances, Whitehead suggests that the ultimate stuff of our reality is akin to a flux of social transactions. Social transactions are the intentional acts that link private feelings to our public roles as children, parents, students, teachers, and so on. Social roles are clustered into organizations and these organizations often endure beyond the life span of individuals, but these structured entities are simply the public manifestation of ongoing social transactions performed with varying degrees of creativity by private agents. In Whitehead's philosophy, these privately nurtured public transactions are the ultimate 'bricks' of both 'private' and 'public' realms, not vice versa. Relational processes, not substances, are the bed-rock of reality.

As a metaphysical metaphor, 'society' introduces a relational dimension that is lacking in the material and mental atoms of the Cartesian-Newtonian Cosmos.[8] From a substance perspective society is simply an aggregate of people; just as atoms make up a molecule, and molecules make up a cell, so people are the 'substance' that make up society. Whitehead foreshadowed current sociological theories of society and rejected this substance viewpoint. In fact, Whitehead's vision helped shape the 'social action' theory of Talcott Parsons, whose theory of society introduced generations of sociologists to a process standpoint in which 'unit acts', not people, are the fundamental 'bricks' of society.[9] And these unit acts are not isolated events; rather they form relational wholes – systems of action. These systems in turn function as normative environments, reinforcing the routines of everyday life that impart to society its defining character. Thus unit acts are shaped by and reproduce the institutional structures of society. Institutions, whether communist or capitalist, matriarchal or patriarchal, are all defined by institutionalized routines. In emphasizing

---

[7] As the combined social and formal focus of my presentation constitutes a novel approach to Whitehead's metaphysical system, there is no single interpretive school of thought that I can appeal to for support; hence the footnotes rely heavily on Whitehead and less so on secondary sources.

[8] For an insightful analysis of how Whitehead's relational thinking departed from the atomistic theories that dominated late nineteenth century physics and psychology, see Capek (1964).

[9] At the time when Whitehead was formulating his ideas, sociology was at a pre-paradigm state of scientific development, and supposedly 'scientific' views of society were controversial. Whitehead was not attracted to the Hegelian idealistic view of social change, and was equally repelled by any materialistic reductionism. The sociological synthesis that facilitated the twentieth century institutionalization of sociological thought in the United States was itself inspired in part by Whitehead's metaphysical system. As noted elsewhere, Talcott Parsons initiated his classic synthesis by embracing a Whitehead inspired philosophical foundation, and using that framework of ideas to combine Idealist and Positivist strains of sociological thought (Dawson 1991). Thus the social metaphor that I employ is to some extent a product of Whitehead's influence, as the concept of 'society' with its sociological connotations has gained a purchase on popular thinking through the work of Parsons and his sociological successors (i.e., in responding to Parsons' theory of society, conservative and radical, micro and macro streams of theoretical thought were redefined, and the Parsonian notion of social structure as a factor in shaping social action became part of the 'taken for granted' furniture of sociological theory).

this point, Parsons was articulating Whitehead's insight that social transactions, not enduring substance, are the foundation of our social Cosmos.

In developing his metaphysical vision of the evolving Cosmos, Whitehead contrasted the dynamic basis of his transactional view of history with the static basis implicit in the Newtonian vision. For Whitehead the social analogy can be extended all the way down to the fundamental energy fields and quantum events of Nature. From the Newtonian standpoint cosmic history is composed of substances which can be abstracted from time and space and treated as states made up of static particulars. The dynamic view makes transactions its particulars, with history as a concatenation of transactions (see figure 1).[10] The concatenation of transactions is a public inheritance that future transactions sustain and build upon. A social transaction in the private phase of its career is a transaction in the making; in its public phase it is a determinate transaction. The great river of cosmic history is constrained and energized by social transactions and subjectivity and society are simply different aspects of a single dynamic transactional process.

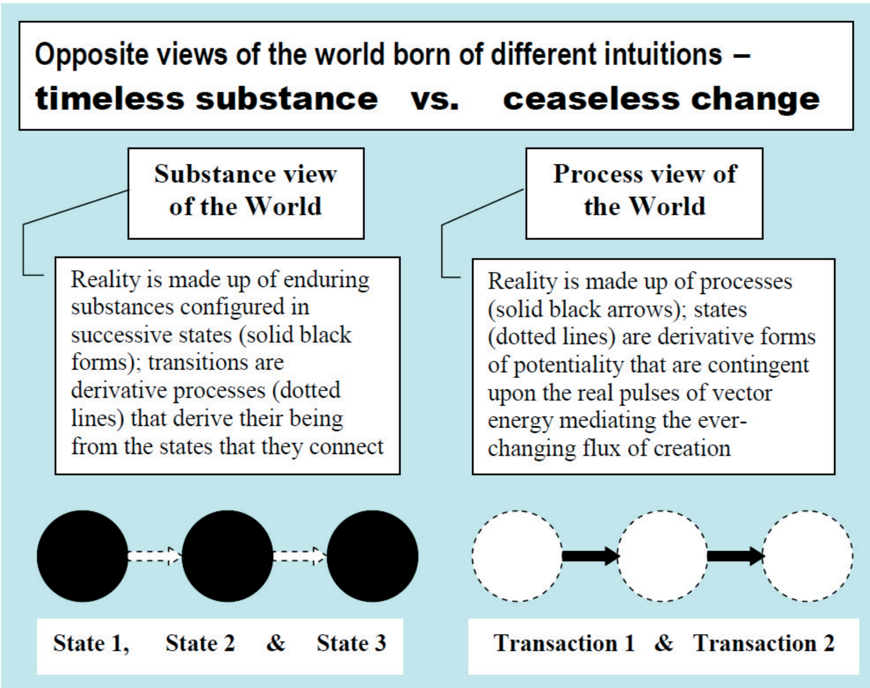


Figure 3.1 Substance versus Process visions of the changing state of the World.

In explaining his transactional view of history Whitehead differentiated between two types of processes, *concrescence* and *transition*, with the former representing the subjective evaluative phase that culminates in a determinate transaction, and the latter process representing the role played by fulfilled transactions as objective recursive ground of future transactions.[11] Thus in transition the subject's action is transformed

[10] It was expressed thus by Whitehead in one of his lectures: "Two notions of history, static and dynamic. The former considers things as occupants of space and time, the particular is the static thing, considered in abstraction from time and space. The other alternative is to view particulars as transactions and history as a concatenation of transactions, "with space and time as certain aspects of it" (Emmet, 2003:19).
[11] Whitehead defines the transitional and concrescent modes of fluency as follows:

into an objective affordance for future action, while during the phase of concrescence a subject responds emotionally to the available affordances and decides to enter into a determinate transaction.[12] Transitional and concrescent processes are complementary elements in the cycle of action that defines Whitehead's universe. They are the two sides of the dynamic vision of 'history in the making', reflecting the complementary public and private sides of every transaction. The Many public acts that make up society are the objective horizon informing each new private intention, and each fulfilled intention adds its factual outcome to the Many acts that make up the objective horizon of future acts. Transition and concrescence thus complement each other as efficient and final causes in an ongoing cycle of activity defining the stream of history.

Thus, rather than view the Cosmos in dualistic terms, as composed of two types of particulars, matter and mind (giving rise to the bifurcation between, on the one hand, a world of changing distributions of enduring matter that includes human agents, and on the other hand, human agents whose subjective vision includes the world), Whitehead's metaphysical system presupposes one universal class of particulars: social transactions, composed of complementary processes of concrescence and transition. Our commonsense notion is that physical action is made up of the movement of enduring bits of matter; similarly, a substance based view of social action reduces social action to the activity of people. But Whitehead states otherwise. In order to translate his vision of process into a detailed conceptual framework that could be applied equally to various domains of specialized scientific knowledge, Whitehead created a new system of philosophical concepts.[13] This new language is the medium through which he sought to escape the conceptual presuppositions of the Aristotelian-Cartesian heritage and frame his solution to the bifurcation problem.

3.13 A monadic vision of the physical universe

Whitehead's metaphysical goal was to define a set of concepts that would translate his dynamic transactional view of history into a logical, systematic description of all that we experience.[14] This translation involved transposing modern science from its individualistic-atomistic Cartesian-Newtonian substance signature into a socially inspired version of the Heraclitean-Leibnizian process signature. By interpreting

> One kind is the fluency inherent in the constitution of the particular existent. This kind I have called 'concrescence'. The other kind is the fluency whereby the perishing of the process, on the completion of the particular existent, constitutes that existent as an original element in the constitutions of other particular existents elicited by repetitions of process. This kind I have called 'transition'. Concrescence moves towards its final cause, which is its subjective aim; transition is the vehicle of the efficient cause, which is the immortal past. (*PR*:210)

[12] Whitehead writes that when we see an external object, it operates as a symbol, communicating its significance to us, "because in the long course of adaptation of living organisms to their environment, nature taught their use. It developed us so that our projected sensation indicate in general those regions which are the seat of important organisms." (SME:57) This is the basis of what James Gibson referred to as *affordances*.

[13] An extensive literature has developed within disciplines of physics, chemistry, biology, psychology, sociology and theology, devoted to applying Whitehead's conceptual apparatus to analysing phenomena associated with each discipline, as reviewed in Weber & Desmond (2008). See also on-line bibliographies detailing this literature at the *Center for Process Studies;* www.ctr4process.org/ .

[14] In Whitehead's words, his "speculative philosophy" is an "endeavor to frame a coherent, logical, necessary system of general ideas in terms of which every element of our experience can be interpreted." (PR:3)

Heraclitus' river of change in terms of Leibniz's ultimate monadic atoms of existence, Whitehead hoped to find a way of relating the subjective qualities felt in Nature to the objective scientific facts of Nature. Leibniz's monads were a promising notion, for in contrast to Descartes' inert matter, Leibniz's 'monadic' atoms were modelled on the human soul. Each monad had the power to represent the universe within itself, had its own appetitive entelechy, and was related to the world in such a manner that they were able to causally influence other monads.

In order to make the Heraclitean-Leibnizian vision compatible with contemporary scientific knowledge of Mind and Nature, Whitehead borrowed concepts from Maxwell's physics and Bradley's philosophical psychology. Whitehead had launched his intellectual career at Cambridge, with a doctoral dissertation on the mathematical foundations of Maxwell's theory of electrical and magnetic fields.[15] As Whitehead later recalled, it was through his study of the new vector geometries and their application in physics that he came to appreciate that mathematical physics had translated "the saying of Heraclitus, 'All things flow', into its own language" (*PR*:309). Whitehead used Bradley's psychological concepts to flesh out and connect a revised vision of Leibnizian monads with his mathematical vision of the Heraclitean flux. The key analogy was between the physical force of vectors of energy and the emotional flow of bodily feelings. Thus, in Whitehead's words, the "creativity of the world is the throbbing emotion of the past hurling itself into a new transcendent fact. It is the flying dart . . . hurled beyond the bounds of the world" (*AI*:206).

In arriving at this conjunction of images Whitehead drew upon his experience as a mathematician and logician, whose initial analysis of vectors was grounded in a novel system of semantics that was experiential in its orientation.[16] Whitehead's philosophical reflections on the foundations of mathematics informed his

[15] Whitehead had begun to develop his vector vision of physics during the time of his Cambridge dissertation on Maxwell's 1873 *Treatise on Electricity and Magnetism*. Maxwell had argued that electromagnetic quantities could be expressed as vector functions of the electromagnetic field. What had vanished with Maxwell's concept of the physical world, Whitehead noted, was the notion of a passively enduring material substance, defined by primary individual attributes and their accidental adventures. On Maxwell see Harman (1998), and on the Maxwell-Whitehead connection, see Dawson (2008).

[16] Whitehead approached vectors from a standpoint that differed from the typical school book illustrations of vectors (forces, expressed in terms of magnitude and direction acting upon a material object) and differed also from set-theoretical notions of vectors taught at a university level (formulated in terms of a vector space and associated field of scalars). As I have argued elsewhere, Whitehead's notion of vectors was informed by a relational philosophical outlook that he had acquired at the onset of his mathematical career in the course of his analysis of Grassmann's theory of extension (Dawson 2008). This outlook, to quote Luca Vanzago, was based upon the rejection of "the Aristotelian notion of substance and its equivalence with the subject of predication", and the adoption of a relational perspective in which substance is no longer the element from which relations are constructed, but is itself "the effect of the interplay between the relations themselves" (Vanzago, 2003:183-4). Robert Valenza (2008) provides an informed view as to how the notion of vectors formulated by Whitehead in *UA* differs from the set theoretical formulations of modern algebra. One obvious distinction is the total absence of concern for the philosophical issues which preoccupy Whitehead. Following Grassmann, Whitehead emphasized a visual interpretation of algebra, in which the basic mathematical operations and associated definitions are given a semantic interpretation that draws upon visual examples. As Valenza has noted, visualization lends itself to expressing relations that are not apparent if defined in set theoretical terms. Reading a set theoretical formulation, like reading any other language, involves presenting a series of events in time; visual presentations, by contrast, enable relations to be grasped simultaneously (Valenza 2008:30-1).

interpretation of vectors and were ultimately translated into a relational view of the physical universe (Whitehead 1906). This relational outlook, first promoted by Leibniz, differed significantly from the dominant absolute view associated with Newton in that it explained both space and matter in terms of a single class of monadic entities.[17] The outcome of the Leibnizian standpoint was, in Whitehead's words, that "Properties of 'space' and the physical phenomena 'in space' become simply the properties of this single class of entities" (Whitehead 1906:82). In rejecting the Newtonian absolute container view of space, Whitehead imagined instead monadic entities constituting their own relationally structured 'space' through the history of their interactions.

Drawing upon Maxwell's field theory, Whitehead provided a physical interpretation of this monadic activity. Whitehead's monad's differed from Leibniz in that each monad's existence is exhausted in a single transaction with its monadic Universe. Like a cell in Conway's game of life, a monad receives an input, evaluates it through a process of concrescence and produces an output. But that is not the end of the story, for while the monad's subjective actuality is exhausted in this single transaction, the public consequences of the act are not. The output of one generation becomes a defining input for the next generation of monadic cells; that is, the output becomes the transitional vectors that hurl their emotional energy into the future initiating another cycle of concrescence and transition. As in human history, the dead live on in the values that they bequeath to their successors.

In applying this model to Maxwell's field theory, Whitehead identified his notion of 'transitional' process with the vector attributes of a vector space, and his notion of 'concrescence' with the scalar attributes of the associated field.[18] "In the language of

[17] The primary distinction between the Newtonian and Leibnizian approaches to conceptualizing the material world consists in the basic presuppositions. The Newtonian view takes as its basic presupposition that reality is made up of several independent factors - time, space, and autonomous atoms of matter – with time and space forming an absolute frame within which atoms assumes various states in the course of time. Alternatively, Whitehead follows on from Leibniz and proposes a relational model, in which the properties of physical phenomena and of space itself are shown to be simply properties of a single class of entities – relationally constituted, interdependent monads. Victor Lowe (the author of Whitehead's definitive biography), in seeking to discover the origins of Whitehead's relational views, asked Russell if Whitehead, in the early years of their collaboration, had ever embraced the absolute, Newtonian theory of space. Russell responded "No," and added, "I think he [Whitehead] was born a relativist" (Interview with Russell in 1965, quoted in Lowe 1975: 91). The influence of Grassmann on Whitehead's mathematical and philosophical outlook is discussed in Dawson (2008).

[18] Scalars and vectors are elements in the geometric algebra which was the subject of Whitehead's first book (*Universal Algebra*, 1898). From the standpoint of Felix Klein, geometry has two ingredients, elements (like points, lines, etc.), and a group of transformations (such as translations and rotations). These transformations act on the elements of the geometry and show how the elements would appear if an observer shifted their standpoint relative to the element (for example, a flat surface viewed from above is a plane, but from the side is a line). Every vector has its associated scalar quantity. A scalar is an element is this geometry which is the same to all observers; it measures 'intensity', such that zero intensity negates the vector element. Vector elements transform in different ways depending upon the standpoint of the observer. Likewise, in physics the scalar element is not changed by transformations: hence the scalar is used to represent a physical quantity that is invariant under rotations of the coordinate system (e.g. a number representing the 'weighted' element of the vector such as its mass or energy) while the vector represents an element in a coordinate space with length (i.e., magnitude) and direction (e.g. representing the magnitude and direction of a force). Following Maxwell, Whitehead used scalars and vectors to represent the charge, magnitude and direction of the particles constitutive of

physical science”, Whitehead states that concrescence begins with a vector form but “the ‘scalar’ form overwhelms the original ‘vector’ form: the origins become subordinate to the individual experience” (*PR*:212). Whitehead suggests that just as “scalar quantities are constructs derivative from vector quantities” so concrescence is rooted in the past acting as efficient cause (*PR*:212). The scalar aspects of a vector represent a ‘private’ subjective evaluation and emotional appropriation of the feelings passed on in the ‘public’ vector transition.[19] Thus transition and concrescence are complementary aspects of the larger process: there is a constant vector transition of the past objective Many into the scalar form of a new subjective One; this private One of conjunctive feelings is in turn transformed as the monadic subject perishes and its objective legacy becomes part of a new public field of disjunctive feeling – a new vector in a field of the Many.[20]

A prominent Whiteheadian, Charles Hartshorne, has suggested that the novel insight at the core of Whitehead’s philosophy is *Creativity* (Hartshorne,1963:18). Creativity refers to the many-to-one-to-many cycle of activity that has been in the background of my description of the public-private vector relationship.[21] By naming his universal of all universals ‘Creativity’ Whitehead signaled the notion that something novel and creative lies as the heart of the Universe. The constant pulsating dialectic, whereby fields of Many disjunctive feelings are unified as One conjunctive feeling, and are added as one more actual entity to the field of the Many is no mere mechanical process. While the Many vectors that make up ‘public’ fields of physical force do act in a determinate non-creative fashion, the concrescence process (whereby each monadic event ‘privately’ unifies those ‘vector’ feelings into a new event, a quantum of monadic force with its own ‘scalar’ intensity) is inherently creative. When each new monadic One contributes its vector force to the Many, it gives expression to a creative potentiality that provides purpose and direction to the evolution of the

---

an electromagnetic field, such that every point in a Euclidian space was associated with a specific vector representing electromagnetic force.

[19] Hence Whitehead writes: “If we substitute the term ‘energy’ for the concept of a quantitative emotional intensity, and the term ‘form of energy’ for the concept of ‘specific form of feeling’, and remember that in physics ‘vector’ means definite transmission from elsewhere, we see that this metaphysical description of the simplest elements in the constitution of actual entities agrees absolutely with the general principles according to which the notions of modern physics are framed. The ‘datum’ in metaphysics is the basis of the vector-theory in physics; the quantitative satisfaction in metaphysics is the basis of the scalar localization of energy in physics; the ‘sensa’ in metaphysics are the basis of the diversity of specific forms under which energy clothes itself” (*PR*:116).

[20] Porter (1975) provides a useful commentary on the scalar aspect of Whitehead’s vector process. He notes that: “Every phase of the process is both objective and subjective. The antecedent world is presented as objective data for the novel occasion, but each datum is presented with the subjective intensity of its own satisfaction. It is a "weighted" condition”. This subjective intensity or ‘weight’ of a vector is determined during the process of concrescence when different vector prehensions arouse aesthetic contrasts that are conceptually configured as various possible unities. The final ‘satisfaction’ of the occasion is a selection from these possibilities based upon maximizing the subjective ‘richness’ or ‘weight’ of the occasion. These subjective intensities are thus a function of the total situation from the standpoint of the subjective aim of the occasion, as determined during the course of concrescence. See also Porter (1979) - “When we get "inside" an event, says Whitehead, the vector character of causation is overtaken by the scalar character of contrasts (*PR* 212/323).”

[21] Whitehead called ‘Creativity’, "the universal of universals", both the ultimate matter of fact, being as such, as well as the ultimate principle of novelty, of a unique entity. According to Whitehead, “‘Creativity’, ‘many’, ‘one’ are the ultimate notions involved in the meaning of the synonymous terms ‘thing’, ‘being’, ‘entity’.” (*PR*:21)

Universe. In conceptualizing the creative dimension of this process Whitehead looked for analogies in the field of psychology.

3.14 From Physics to Psychology: bodily feelings as a paradigm for vector force

While Whitehead's vector interpretation of feeling was inspired in part by a mathematical interpretation of Maxwell's fields of force, it was also deeply indebted to the philosophical psychology current in England at the end of the nineteenth century. In interpreting the creative processes constitutive of Nature, Whitehead assumed that we are part of Nature and that our subjectivity is disclosing generic characteristics common to all of creation. In adopting this standpoint, Whitehead drew upon the work of the English Idealist philosopher and psychologist, F. H. Bradley. According to Bradley's analysis, each ego is aesthetically connected with the world at a visceral level of emotional feelings that are more fundamental than visual or intellectual cognitions. Whitehead gave a realist twist to Bradley's idealist categories of thought, and interpreted the ego's subjective bodily feeling of connection with objects directly felt as a causal chain of direct physical connections, analogous to Maxwell's vectors.[22] Thus, even if we don't see the stone that we stub our toe against, we feel the energy embodied in our collision with the rock as an emotional force. In this encounter the felt object is analogous to a vector form of energy while the subject's live emotional response to the object is akin to the scalar quantification of energy.

For Whitehead, Bradley's crucial insight involved the role of those visceral bodily feelings that carry an emotional weight and prompt intuitive reactions (such as painful feelings associated with a stubbed toe). Adopting a vector interpretation of bodily feelings enabled Whitehead to envisage a way of escaping from the solipsism of a private psychological field made up solely of 'secondary qualities'. Whitehead argued that if the psychological field was an autonomous, purely private domain measurement would be meaningless and science impossible. Instead, he argued that our "private psychological world is really our intuition of a certain relationship of the body to the rest of the world", and that this "relation is governed by the states of my body" (Emmet 2003:12). What this relationship expresses is a common system of extensive connections, more fundamental than our visual perceptions, that is reproduced by the bodily transactions that directly connects the body with its physical situation and mediates a flow of emotional energy. "Feelings", Whitehead writes, "are 'vectors'; for they feel what is *there* and transform it into what is *here*" (PR:87 [W's italics]).

[22] In drawing upon Bradley's concept of emotional feelings (prehensions) to interpret the electromagnetic events that make up Maxwell's flux, Whitehead appropriates Bradley's system of psychological concepts for his metaphysical monads. Whitehead notes that: "In accordance with this doctrine of Bradley's, I analyse a feeling (or *prehension*) into the '*datum*' which is Bradley's 'object before me', into the '*subjective form*' which is Bradley's 'living emotion', and into the '*subject*', which is Bradley's 'me'" (*AI*: 268-9). Whitehead uses Bradley's notion of a *subjective form* to suggest how objects prehended in the Many make an emotional claim upon the subject, and that the intensities and contrasts between these competing influences guide the process of subjective integration of the Many public feelings into One privately felt world (concrescence), and borrows from Bradley the notion of the '*satisfaction*' as the termination of this creative integrative process.

Whitehead emphasized that the vector connections between the objective Many of the public datum and the subjective One is not an external connection, but an inclusive, internal connection, where the positive prehensions of a datum preserve it "as part of the final complex object which 'satisfies' the process of self-formation and thereby completes the occasion" (*AI*:271). By interpreting vectors as prehensions Whitehead incorporated the world of external actualities into the subjective realm of internal relations.[23] All entities, from complex cellular creatures down to the sub-atomic entities studied by physicists, have a subjective, perspectival connection to 'things felt' - a real connection to the felt world. This "object-to-subject structure of experience", Whitehead states, is the basis of "the vector structure of nature" (*AI*:219). These vector 'feelings' of extensive connection are Whitehead's alternative to the Aristotelian notion of substance. They are the vector transactions that define the basic societal context of the Universe. As Whitehead notes, "for our epoch, extensive connection with its various characteristics is the fundamental organic relationship whereby the physical world is properly described as a community" (*PR*: 288). Vector prehensions thus sustain both a public macro realm and a micro subjective realm. As Whitehead writes:

> The sole concrete facts . . . are prehensions; and every prehension has its public side and its private side. Its public side is constituted by the complex datum prehended; and its private side is constituted by the subjective form through which a private quality is imposed on the public domain. . . . Prehensions have public careers, but they are born privately (*PR*:290).

Thus for Whitehead the physical universe is not an aggregate of autonomous particles of externally related substance, but a community of internally related monadic occasions sustained by the emotional intensity aroused by ongoing transmission of vector feelings from generation to generation; it is these concatenated chains of prehended feelings that make up the extensive community of physical occasions constitutive of the Universe. "For Descartes", Whitehead notes, "the primary attribute of physical bodies is extension; for the philosophy of organism the primary relationship of physical occasions is extensive connection. This ultimate relationship is sui generis, and cannot be defined or explained. But its formal properties can be stated" (*PR*: 288). Whitehead spells out these formal properties in his mereotopological theory of extension.[24]

3.15 Whitehead's prehensive 'bricks' of physical feeling and the relational structure of the consequent 'brick-wall' universe.

[23] In philosophical terms an internal relation is one that is vital to the identity of the subject; the subject would be a different entity if its internal relations were different. External relations, on the other hand, do not alter the essential identity of the subject. For example, a person's identity as husband or wife is internally related to their partner's corresponding spousal identity, while the colour of their partner's hair might be regarded as externally related to the colour of their own hair, as it will not effect the identity of the latter.

[24] This theory of extension is developed informally in Part IV of *PR*. Whitehead's system of proposed axioms have been critically developed by a number of mathematicians and logicians, and is central to the school associated with Anthony Cohn (Cohn & Varzi, 2003). As noted in the previous chapter, Cohn's work has been an attempt to develop a logical language for computational systems working with spatial information, in which regions are the basic entities. The inspiration for this work goes back to Whitehead's theory of extensive connection (Gotts, Gooday & Cohn 1996:58).

Charles Hartshorne has argued that Whitehead was the first philosopher to "embody modern relational logic in a fairly complete metaphysical system" (Hartshorne 1950:33). The mathematical specification of the structured system of potential monadic connections constitutive of Whitehead's Universe is a way of representing the systematic structure of relations that arises from simple prehensions. The relationship in a simple physical feeling between the prehended and prehending monadic entity defines a real and particular actuality that Whitehead refers to as a 'nexus'. According to Whitehead, actual entities "involve each other by reason of their prehensions of each other", and this mutual prehension can cause a common subjective form of feeling to arise within a group of actual entities: any such "particular fact of togetherness among actual entities is called a 'nexus' (plural form is written nexūs)" (*PR*:20).

John Lango has argued that nexūs disclose the systematic structure of Whitehead's cosmos. This structure, at its most elementary level, can be explicated in terms of a system made up of nexūs where "any two actual entities in a nexus are linked by the ancestral of the relation feels or is felt by", and the resultant structure of simple physical feelings can be expressed in mathematical terms as a partially ordered set of relations that are "irreflexive, asymmetric, and transitive" (Lango 2000). Lango argues that Whitehead's universe is a 'nexal set' of actual entities and that this set continues to grow through ancestral relations of descent whereby physical feelings give rise to new occasions. Claus Ringel notes that for Whitehead "a nexus is a set of actual entities which form a unit, for example via coincidence or via temporal succession [and that] there always is a corresponding extensive quantum, thus a region" (Ringel 2008: fn.4, p.145). Simple physical feelings thus give rise to systematic spatial and temporal relations that define extensive regions of our physical Universe.

Whitehead notes that "all our physical relationships are made up of such simple physical feelings, as their *atomic bricks*" (*PR*:238 [my italics]). These bricks are joined together to form an extensive continuum of bricks, locked together in a relational nexus analogous to a brick-wall universe. Extensive connection is the most elementary of the simple vector feelings at the base of all physical relationships.[25] These 'bricks' are the feelings that are re-enacted in each new subject whereby the cause is transferred into effect; the result is an extensive continuum of uniform connections, a public *brick wall universe* made up of simple physical feelings that serves as the datum for each new monadic occasion (see figure 3.2).[26]

[25] Claus Ringel (2008) notes that for Whitehead "a nexus is a set of actual entities which form a unit, for example via coincidence or via temporal succession [and that] there always is a corresponding extensive quantum, thus a region" (Ringel 2008: fn.4, p.145). The term is another example of Whitehead's tendency to generalize and extend ideas derived from mathematics and logic and put them to work in his metaphysical system.

[26] While I owe the brick wall analogy to Lambert, Whitehead does make use of the brick metaphor. For example, he writes: "It is now obvious that blind prehensions, physical and mental, are the ultimate bricks of the physical universe" (*PR:*308). These 'bricks' are bound together both in each actuality (by the subjective unity of the aim governing their genesis to a final concrescence), and beyond that actuality (by the way in which, as prehensive elements of a superject, they become part of the objective datum prehended by a later subject – "thus objectifying the earlier subject for the later subject"). In this way prehensions are the medium of both internal and in some sense of external relations (*PR:*308-9). Although the brick wall image of the universe sits uneasily with the Heraclitean image of flux, I am stressing it as it is an image that expresses the ultimate systematic structure of the Universe. It is the systematic order of potential relations that underpins monadic activity; this system of extensive

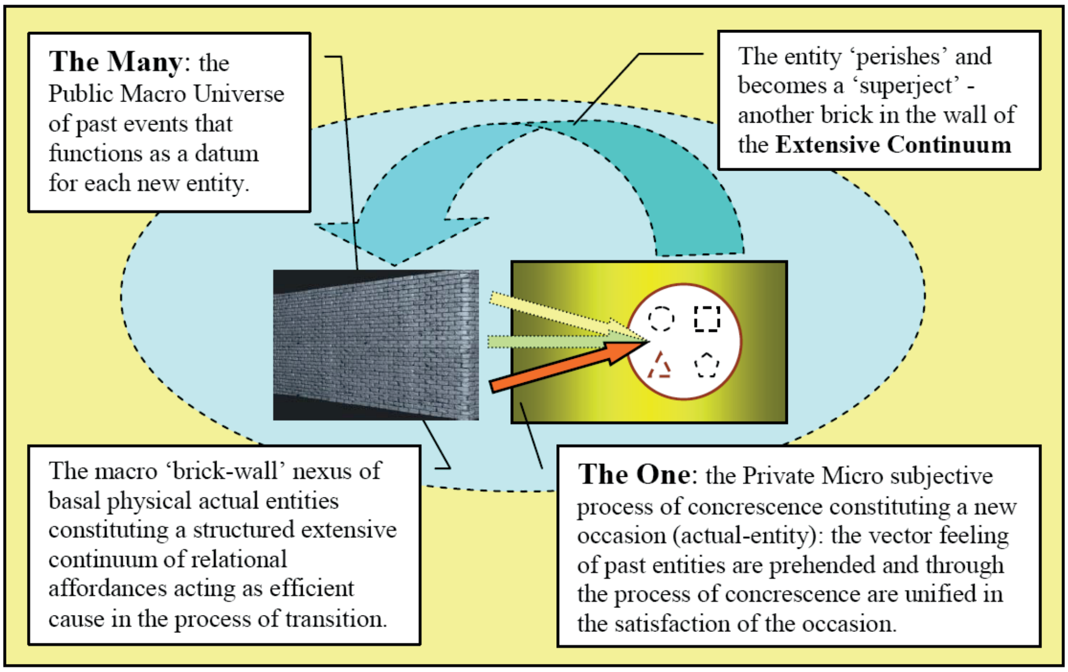


Figure 3.2 The brick-wall universe: The Many become One and are increased by One.

3.16 From simple brick-wall to rich mosaic: the evolution of societies

The simple prehensive 'bricks' of the universe can participate in more complex monadic events, giving rise to various nexal sub-structures in the universe. As a result, the brick-wall universe is transformed into a rich mosaic, made up of various kinds of serially, or non-serially related sets of 'bricks', embodying distinct qualitative emotional forms representing more or less complex kinds of structures. In order to explain how this transformation takes place Whitehead supplements his notion of *simple physical feelings* with the notion of *propositional feelings*. Simple physical feelings mediate the blind conformity at the base of the brick-wall structure of the Universe; propositional feelings build upon this structure and introduce novel possibilities that lure creation towards new forms of attainment. Whitehead's resolution to the problem of the bifurcated universe revolves around the way that simple physical feelings enter into the formation of propositional feelings. Human commonsense is grounded in simple physical feelings and while the development of propositions to some extent enables commonsense to transcend its origins, it does retain an essential reference to that systematic physical ground. It is this linkage that enables Whitehead to overcome the solipsistic problem associated with the Cartesian

---

connection is the ultimate societal substratum unifying Whitehead's universe. In contrast to the Cartesian-Newtonian vision of the universe, time and space are not ultimate presuppositions; rather, "time and space are characteristics of nature which presuppose the scheme of extension. But extension does not in itself determine the special facts which are true respecting physical time and physical space" (*PR*:289). In other words, for Whitehead our system of time and space relations is not an ultimate metaphysical postulate (as for example, Euclidian geometry was for Newton). Instead, geometrical relations are emergent characteristics of the society of monads that evolved in our Cosmic epoch. They reflect the customary form of elementary monadic transactions, or in other words, the basic social habits that define our epoch.

mind. In order to appreciate the nature of this linkage it is necessary to consider simple physical feelings in more detail.

Whitehead's hypothetical simple physical feelings are both an act of causation and the most primitive mode of perception, objectifying and 'representing' the perceived causal datum. In comparison to Descartes' atoms of substance, the irreducible basal monads of Whitehead's system are complex relationally constituted entities. Simple physical feelings have only one monadic entity as their initial datum (this is the monad in its transitional role as a 'superject' in the public realm of the Many). The initial datum (the superject) is the cause of the feeling, the feeling is the effect, and the monadic subject entertaining the feeling is the "actual entity conditioned by the effect" (*PR*:236). The content of the feeling is determined by 'objectification' of the form of the initial monadic entity, as a feeling for the second subject. This objectification provides the second subject with its 'perspective' on the first monadic subject. Hence as a causal factor, the superject acts as an initial datum, and this has the effect of bringing into being a new physical feeling and corresponding subject, and this subject in turn objectifies the first entity and feels it perspectivally as an objective datum. Note that in an elementary sense, *representation is a fundamental aspect of all processes*. In Whitehead's words "the initial datum is the actual entity perceived, the objective datum is the 'perspective' under which that actual entity is perceived, and the subject of the simple physical feeling is the perceiver" (*PR*:236).

Where a nexus involves more than two entities a further phase in the process of concrescence can occur. For a subject to feel a nexal form of togetherness the simple physical feelings of the many actualities prehended in the nexus must be 'transmuted' into one single physical feeling; this takes place via the conceptual prehension of some qualitative characteristic in the simple feelings that are shared by the superjects prehended in the initial datum. This occurs when the physical feelings felt in the members of a nexus share an ordered form of feeling that, through contrasts with the corresponding *conceptual feelings* (to be explained shortly), is able to be transmuted into a simplified datum that can then mediate a new perspectival objectification of the datum, in which the entire nexus is felt as a single entity through the transmuted feeling. Through transmutation, the individual bricks of the wall are grouped into a related nexal set that is prehended as a felt unity. As Whitehead notes:

> Transmutation is the way in which the actual world is felt as a community, and is so felt in virtue of its prevalent order. For it arises by reason of the analogies between the various members of the prehended nexus, and eliminates their differences. Apart from transmutation our feeble intellectual operations would fail to penetrate into the dominant characteristics of things (*PR*:251).

In Whitehead's relational world, transactions and the public and private realms that they connect are the concrete elements of every experience.[27] Each prehension "is referent to an external world, and in this sense will be said to have a 'vector character'; it involves emotion, and purpose and valuation, and causation" (*PR*:19). Actual entities (or 'occasions') are the final realities of existence in which we experience the 'subjective form' of prehensions ("Private Matters of Fact"); nexūs are

[27] As emphasized by Whitehead, the three interrelated notions of actual entity, prehension, and nexus are at the heart of his effort to conceptualize the "most concrete elements in our experience" (*PR*:18). All else is derivative abstraction. Hence, according to Whitehead, the "ultimate facts of immediate actual experience are actual entities, prehensions, and nexūs" (*PR*:20).

"Public Matters of Fact" constituted by the relationally connected inheritance bequeathed by past actual entities (*PR*:22). The public realm is defined by societies.[28] Transmuted simple physical feelings are the basis of this social order. This societal dimension of Whitehead's cosmos will be the focus of the next section.

3.2 Whitehead's communal universe and the creative role of habit

The metaphysical appropriation and generalization of the notion of society and its shared forms of order (dominant habits or 'customs') represents one of Whitehead's most original and important contributions to the process tradition.[29] In Whitehead's universe, monads have no enduring existence; only societies endure, and it is through societies that our monadic moments of experience are part of enduring identities. The self identity of a society is defined by the common subjective form immanent in all members of a society. In human society we refer to these 'defining characteristics' as social routines, habits, norms or customs. I shall follow the example of Lewis (1991) and use the notion of custom or social habit as synonyms for the generic 'defining characteristics' of all grades of society.

Referring to the significance of routines, Whitehead stated that, "it is the beginning of wisdom to understand that social life is founded on routine", and that without routine "civilization vanishes" (*AI*:111). Routine or custom is the basis of the Universe's evolutionary thrust towards societies that support greater degrees of creative freedom. Customs, for Whitehead are instantiated in emotionally charged vector events in the continuity between "the subjective form of the immediate past occasion and the subjective form of its primary prehension in the origination of the new occasion" (*AI*:213). For Whitehead, there is no dualistic divide between Nature and Culture, Matter and Mind. The way our world is structured places massive constraint on how we experience the world; these constraints are an expression of socially embodied habits. Without habits, the fluent world of process would be featureless chaos. Through habitual action the Cosmos acquires structure. Hence, in the Whitehead paradigm, *process* and *habitual activity*, rather than *material* and *mental substances*, form the basis of an explanation of representation, rationality and commonsense.

---

[28] The extensive continuum is a nexus encompassing all the actual entities of our epoch. It is also a society, since it is a shared form of order. However, within this continuum, sub-societies arise, and members of those societies can have different forms of order yet still interact. Whitehead's formal definition of 'social order' and 'society' is as follows:

> A nexus enjoys 'social order' when
>
> > (i) there is a common element of form illustrated in the definiteness of each of its included entities, and
> > (ii) this common element of form arises in each member of the nexus by reason of the conditions imposed upon it by its prehension of some other members of the nexus, and
> > (iii) these prehension impose that condition of reproduction by reason of their inclusion of positive feelings involving that common form.
>
> Such a nexus is called a '*society*', and the '*common form*' is the '*defining characteristic*' of that society (*AI*:236; *PR*:34).

[29] Not all Whiteheadian philosophers accord the same importance to the social dimension; often the formation of the monad is discussed in isolation from its social context. For a discussion of the generic metaphysical significance of notions of society and custom in Whitehead's philosophy (and related technical terms), see Lewis (1991).

Giving a central place to habit in explaining our capacity for making sense of our world has a respectable philosophical history. The great Scottish philosopher, David Hume, when confronted by the transitory evidence of experience, asked how we can draw conclusions about cause and effect. His daring answer was an appeal, not to divine Reason, but to "Custom or Habit" (Hume 1748: §36). Hume's notion of Reason can lead to cultural relativism - to the solipsistic dilemma of people locked in self-contained worlds of meaning – but by interpreting Reason in terms of habit he had rescued the realm of ideas from the Gods. Hume's Scottish contemporary, Thomas Reid took this process a step further; he took habits out of the mind and made them a property of practical interactions with the world.[30] Drawing upon Reid's insights American pragmatists such as Peirce, James and Dewey transformed Hume's insight that Reason was grounded in habit, into a far reaching evolutionary doctrine concerning the role of habit in the evolution of life. Whitehead took this a step further and applied it to the Cosmos as a whole.

By formulating a unified metaphysical vision of the cosmos, Whitehead was able to re-construct the Aristotelian-Cartesian notions of representation, knowledge and rationality on the basis of habit. Whitehead's customs or habits are grounded in the physical world; at the base of habits are real processes defining real patterns of activity.[31] If customs or routines are interpreted in Whitehead's metaphysical sense they explain not only the patterns of causality exhibited by Nature but, at the other end of the spectrum of social complexity, can be used to interpret human reason.

[30] According to Reid's theory of commonsense, the abstract ideas and principles that enable us to make sense of the world have their origins in our interactions with the world. For Reid, the ideas that are our cultural legacy have their origins and retain their meaning because of their use-value in mediating people's practical interaction with the world. Ideas derive their significance from their part in indicating recipes for action. They would be empty of meaning if they did not indicate real potentialities. For Reid, to acquire a concept is to acquire a disposition to interact with our world in a particular way. If a concept, such as eating, drinking, walking and talking, is to be a satisfying concept then the action it promotes must be fitted to the world. Concepts, from this standpoint, rather than signify a transcendent barrier between ourselves and the world, signify patterns of activity that enable us to contact our world. They are dispositions that bring about a satisfaction or fulfilment.

The US philosopher, David Banach, makes the connection between concepts and dispositions in his insightful critique of the Aristotelian-Cartesian roots of the analytic tradition. In presenting his pragmatist inspired approach to representation and knowing Banach argues that concepts be viewed as "dispositions to connect modes of interaction through joint activation" (Banach, 1987:79). The notion of habit and concept are, in this context, interchangeable; both refer to a pattern of interaction with the world and its potential attendant satisfaction. The acquisition of concepts is synonymous with the acquisition of the capacity to realize potentialities that our world affords. As Banach writes, "We, as a system of concepts [habits], act, through the expression of the potentials that this [organism-environment interaction] system has as a whole" (ibid).

Thus, for example, the 'meaning' of the concept "mother" is learnt by the child in the context of habitual patterns of interaction and attendant satisfactions. The concept embodies a complex of dispositions all of which signify potential patterns of interaction. The power to 'represent' does not lie passively in the token but is instead the agent's capacity to perceptively anticipate a pattern of interaction and some attendant satisfaction. Words play a role in indicating these patterns of activity, but in themselves they are empty of significance until a monadic agent interprets them in the context of its actual situation. Language, from this perspective, is not Descartes' filing cabinet of divinely bestowed ideas. Rather, from a process perspective, language is a historical legacy that comes to life when it connects a monadic subject to their world, not in some abstract sense, but by indicating habitual patterns of action that can potentially be realized in the present situation, bringing about a fulfilling interaction with the world.

[31] Humes' customs, understood in this sense, take on a far wider significance and, as Whitehead argued, become the basis for a realist interpretation of causality (*AI*:215).

Societies are the store-house of habit that mediate the evolution of the cosmos; more complex habits mediate more complex forms of subjective response in a spiralling cycle of increasing sensitivity to more diverse potentialities. However, these higher social modes of feeling are internally related to and dependent upon lower level societies and their characteristic modes of feeling.

Whitehead's societies are founded in the mutual immanence of their member occasions. The vector character of experience means that all perished occasions are immanent in future occasions, but this immanence only becomes physically forceful when a group of perished occasions are spatially and temporally contiguous and share in some common form. Such a *nexus* of occasions acts like a human crowd, when it shares a common emotion, such that the emotion is magnified and becomes a more forceful factor in the formation of future events. The capacity of a societal group to impose this common element of form on (most of) its members is the basis of all societies and of the social order that they exhibit. "The members of the society", Whitehead states, "are alike because, by reason of their common character, they impose on other members of the society the conditions which lead to that likeness." (*AI*:237) The 'common form' is a defining characteristic has been inherited by each member from other members that belonged to the datum of its concrescence (*PR*:34).

The Whiteheadian cosmos has evolved to its present state primarily through the influence of habits or routines. The order in our universe is a function of viable forms of sociality, passed on by habits; natural selection is based upon the success of habits in mediating the survival of societies. For Whitehead, socially reinforced habitual patterns of activity, not divinely fixed laws of nature have shaped the evolution of our cosmos. Simple societies defined by simple customs constitute highly predictable environments, but these environments form habitats which support the evolution of more specialized societies. As in the case of learning a musical instrument, simple habits form foundations upon which higher order habits can form. The Cosmic orchestration thus feeds upon itself, and multiplies like variations on a melody, with societies and societies of societies, forming nested hierarchies; sub-universes that support diverse forms of value-seeking activity that reproduces different social forms of order and value. It is these habitually tuned social activities that structure our Universe, and it is these self-same patterns of activity and satisfactions that mediate and shape human experience.[32]

Thus, although a vector flux of transactions is foundational from the process standpoint, thanks to emergent societal structures and the social sensibilities mediated by these structures our experience acquires a content and value that is tied to enduring identities. We cannot be indifferent to the social sensibilities that define our relationships to our situation as it is our sensibility of these social objects that have given rise to who and what we are, at all levels of our being and becoming. Habits, and the role that they play in the creation of tightly co-ordinated societies, are a key to this sensitivity. It is the trained ear that is sensitive to musical subtleties, and as every

[32] The sociality of existence is also an important theme in Dewey's philosophy. Frisina, for example, has argued that "Dewey's understanding of experience presumes a pluralistic environment that is constituted by an indefinite number of transactional sub-processes. Experience literally is an organism's sensitivity to this sociality. Knowing is a specialized form of this social activity." (Frisina, 2002:120)

musician will attest, it is the slow process of acquiring habitual skills through practice and more practice that nourishes this sensitivity. As Isabelle Stengers has argued, Whitehead's reformulation of process philosophy helps us to see habits, not as a sign of human weaknesses but as magnificent achievements. She writes:

> In a way laws of nature rely on robust habits, and so do our living bodies and our most complex behaviours. For Whitehead, as well as for William James and the whole Anglo-American tradition, starting from Hume, habits are not something to criticize. They are achievements indeed (Stengers, 2005).

Hence, the 'laws of Nature', for Whitehead, are not evidence of a pre-determined order of necessity; rather, they reflect a history of social transactions – very stable transactions manifesting the mutual immanence of the order that they impose on each other. Hence Whitehead's statement that: "There are no natural laws; there only temporary habits of nature" (Whitehead, quoted in Price 1954:363). Each social monad is an occasion in a systematic self-structuring society of occasions. Monads acquire their historical relevance through belonging to a society and societies survive and persist only to the extent that new generations of monads sustain the continuity of their existence. Physics provides massive evidence of this habitual continuity, and one of Whitehead's chief metaphysical preoccupations was to specify its nature. Bradley provided Whitehead with a way of transforming Maxwell's vectors into subjects, but by adapting the notion of habit, Whitehead was able to provide a social explanation for the enduring physical 'substances' of our universe. Further examination of simple physical societies will help elucidate Whitehead's complex metaphysical explanation of the role of 'habit' in the evolution of the cosmos.

3.21 Strain feelings and elementary strain societies

The most universal manifestations of the common social forms or customs characterizing our cosmic epoch are found in the nexūs that Whitehead describes as *strains*. Whitehead's notion of strain feelings has its origins in the nineteenth century analysis of physical stresses and strains in matter (a 'strain' as the energetic bonds holding the particles of a fibre together). In Maxwell's electro-magnetic fields, strains became lines of force running through the 'ether'; Maxwell's model, Whitehead states, provided the basis for a new vision, in which matter represented entanglements in the ether – knots – that impose "stresses and strains throughout the whole jelly-like ether", such that agitations of ordinary matter are transmitted (in various forms of energy) "through the ether as agitation of the stresses and strains" (Whitehead 1938:136-138). Whitehead appropriated this notion of strains and transformed it into one of 'strain feelings'; and these vectors of physical feeling became the energetic-emotional communal fabric of his social Universe. Hence Whitehead writes:

> Every reality in the real world promotes feeling and is felt. Nothing is 'merely private' or an inert factor in the public sphere. The feelings that are of 'dominating importance' in the present epoch are 'strain feelings'" (*PR*:310).

Strain feelings have a crucial significance in Whitehead's metaphysical system as they define a systematic physical order of relations in which all higher order modes of feeling, including perception and intellectual analysis, are grounded. His redefinition of strains is thus fundamental to resolving problems associated with Descartes' bifurcated universe. Whitehead's 'strain feelings' are vector feelings that have been socialized, or in Whitehead's technical terminology 'transmuted', such that they pass on their mode of feeling directly to each other, and together make up a concrete social nexus of directly connected monadic entities. Strain feelings are the most fundamental social entity to arise from the transmission of simple physical feelings and the most elementary of the observable physical structures in our universe; they are the energetic-emotional fibres that internally connect the members of Whitehead's organic universe. Various families of strains have evolved, each with its characteristic forms of behavior. However, each strain exhibits some generic properties, in that all are made up of a connected series of physical vectors, defining an extensionally related social body (manifesting geometrical properties), such as the magnetic fibres or lines of force associated with our planet (see figure 3.3). Hence Whitehead writes that:

> . . . the growth of ordered physical complexity is dependent on the growth of ordered relationships among strains. Fundamental equations in mathematical physics, such as Maxwell's electromagnetic equations, are expressions of the ordering of strains throughout the physical universe. (PR:311)

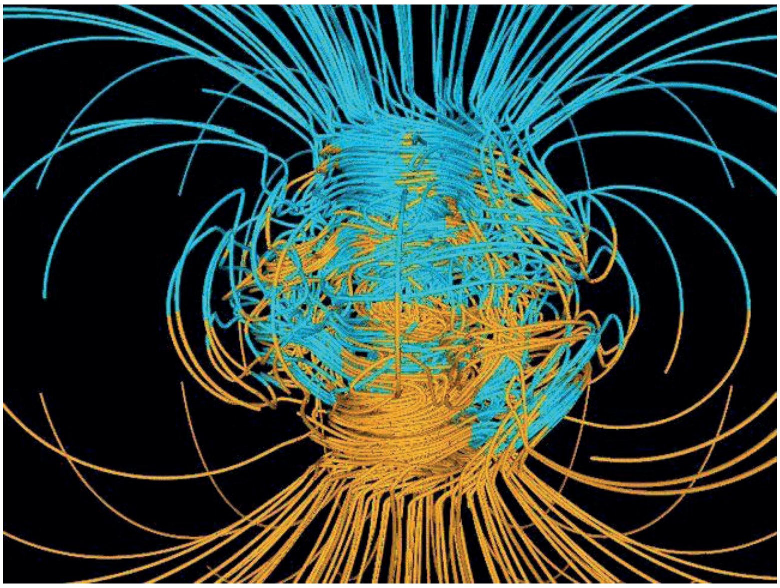


Figure 3.3 The Earth's Magnetic Field – a dynamic but relatively stable strain structure[33]

Thus the 'brick-wall' universe of prehensive relations is a social nexus made up of strain feelings. This nexus derives its precise structure of mathematical relations from the extensive connections mediated by strains; strains define the "systematic order of extensiveness which characterizes the cosmic epoch in which we live. *The societies of enduring objects – electrons, protons, molecules, material bodies – at once sustain that order and arise out of it*" (*PR:*326). Thus the systematic order exhibited by societies of strains provides the ground for the macro communities that structure our physical universe. These macro communities are the enduring objects that form the horizon of our conscious experience; they are collective modes of attainment that mediate everything with enduring identity and value. In Whitehead words:

> The real actual things that endure are all societies. They are not actual occasions [i.e., not monads, since these are constantly perishing] . . . The Universe achieves its values by reason of its coordination into societies of

[33] Graphic by Gary A. Glatzmaier (UCSC) from a computer model of the magnetic fields generated by the Earth's core; blue lines show vector forces flowing inward and yellow lines depict the flow of force directed outward. This image is available at http://apod.nasa.gov/apod/ap021125.html .

> societies, and into societies of societies of societies. . . . Thus an army is a society of regiments, and regiments are societies of men, and men are societies of cells and of blood, and of bones, together with the dominant society of personal human experience, and cells are societies of smaller physical entities such as protons, and so on and so on (*AI*:237-239).

In order to understand how the internal relations constitutive of the process of concrescence selectively reproduce the particular 'characteristic forms' of a society Whitehead extends his theory of concrescence to explain how feelings are selectively unified by each monad; that is, he describes the way in which physical prehensions are qualitatively individuated, contrasted with conceptually prehended patterns of potential satisfaction, valued and selectively unified in the light of the best available goal state. The process of concrescence is a means to the desired end. This subjective process revolves around the formation of a subjective aim; the monad's subjective aim arises from the conceptual valuation and selective unification of feelings afforded by the initial datum of transitional feelings. The subjective aim is thus a conceptually prehended perspectival vision of unity that must be translated into physical fact through negative prehensions that eliminate all qualitative internal connections except those that satisfy the aim.[34] In order to explain this selective process Whitehead introduces his notion of a universal relational form of potentiality, which he refers to as an 'eternal object'. Eternal objects play an essential role in the private concrescence of each occasion and in the operation of the public customs that mediate social influences.

3.22 Sensa - qualitative universals that individuate strains

Whitehead stated that "blind prehensions, physical and mental, are the ultimate bricks of the physical universe" (*PR:*308). These bricks come in two species, physical and conceptual, and both play an essential part in Whitehead's explanation of the role played by customs in creating the rich social structure of our Cosmos. In simple terms, the physical bricks are real physical prehensions, uniting a transitional superject with a concrescing subject. Conceptual bricks, or prehended *eternal objects*, are abstract models of these potential relations.[35] It is through the conceptual bricks that the monadic subject is able to create the plan that guides their selective prehension of

[34] In using the term prehension to signify the multiplicity of vectors contributing their form of feeling to the new monad, Whitehead introduced the notion of 'negative prehension', to signal the selective process that takes place within the monad. The subjective form is made up of the vivid contrasts between the feelings prehended, and not all are compatible for joint exemplification.

[35] The relationship between eternal objects and the physical entities in which they are instantiated is a generalization of the mathematical ideas worked out by Whitehead in *UA* regarding mathematical manifolds and concrete things. According to Whitehead, when some concrete set of objects possess a common qualitative property (such as sound) such that the relations between the members of the set can be described in terms of relational variants of this property (such as when musical notes A, B and C are differentiated in terms of pitch, loudness, etc.), the same concrete scheme of theing can be represented in purely abstract terms as a manifold of elements, where the "relation of a thing in a scheme to the corresponding element of the corresponding manifold is that of a subject of which the element can be predicated" (*UA*:14). *His notion of the manifold–scheme relationship is most readily appreciated in terms of an abstract geometrical manifold that is structurally isomorphic to the corresponding concrete scheme of spatial regions.* Just as an eternal object can be the characteristic form of a society, so a "manifold corresponding to a scheme is the manifold of the determining property of the scheme" (*UA*:14). The various senses in which Whitehead uses an eternal object can be related to the general notion of abstract substitutive schemes, as spelt out in terms of manifolds in *UA*.

physical bricks; the conceptual design arrived at in private lies behind the physical actual entity constructed in the public domain. This guiding function can be illustrated with respect to strains.

Eternal objects, like prehensions have their public and private sides. Prehensions have their public side in the objective form of the datum prehended and their private side in the subjective feeling of the prehending subject; likewise, eternal objects have a public and private dimension. The public side corresponds to a general form of objective potentiality, while the private side is a subjective evaluation of that potentiality with respect to a given aim in a specific context. Whitehead thus refers to two 'species' of eternal objects; on the one hand an eternal object can refer to a universal that is a 'characteristic' form of some relational mode of extensive connection between actual entities (a general type), but it can also refer to a private form of definiteness, experienced as a subjective valuation of the relevance of a qualitative potential felt by an actual entity (*PR*:290-2).

Thus, in depicting a world of determinate prehensions, Whitehead viewed their particularity as the exemplification of universal patterns of potentiality latent in the basic structure of Creativity.[36] Every fact is the embodiment of an abstract potential, such that "things which are temporal arise by their participation the things which are eternal" (*PR*:40). The world of concrete fact is made up of realized patterns of relational connection, while the abstract realm of eternal objects is made up of patterns of potential relational connection. Hence, alongside the physical extensive continuum is an ideal extensive continuum; both dimensions are grounded in and reflect the same monadic process. How eternal objects play their role in Whitehead's system can be illustrated with respect to the function of sensa in strain feelings.

The *qualitative sensa* disclosed in strain feelings are the most elementary form of an eternal object.[37] Sensa individuate the different types of strain feeling, such as various frequencies of electromagnetic radiation, and manifest themselves subjectively in qualitative feelings associated with those different frequencies. The basic element of strain feelings are the prehensive object-to-subject connections between simple physical feelings. The prehensive connections mediated by strain feelings are individuated qualitatively in terms of the feeling of types of sensa. Whitehead envisaged the nature of these simple eternal objects to be something analogous to simple qualitative feelings of colour such as 'redness'. Another simple form of eternal object that plays a key role in Whitehead's vision of concrescence is 'pattern'. Pattern is felt as a contrast between sensa, as the "realization of the pattern is through the realization of this contrast" (*PR*:115). Sensa, when functioning as eternal objects of the subjective species, guide the process of concrescence, and when functioning as

[36] Whitehead first lays out a vision of eternal objects in *SMW*, chs. 10 & 11. His views are further developed in subsequent works, and have been a fertile field of dispute ever since, but the basic notions, as stated in *PR*, are that: 'The eternal objects are the pure potentials of the universe; and the actual entities differ in their realization of potentials' (Whitehead, 1978: 149).

[37] Whitehead writes: "[E]ternal objects which will be classified under the name 'sensa' constitute the lowest category of eternal objects. . . . Sensa are necessary as components in any actual entity, relevant in the realization of the higher grades. But a sensum does not, for its own realization, require any eternal object of a lower grade, though it does involve the potentiality of pattern and does gain access of intensity from some realization of status in some realized pattern. . . . Thus each sensum shares the characteristic common to all eternal objects, that it introduces the notion of the logical variable, in both forms, the unselective 'any' and the selective 'some'." (*PR*:114)

eternal objects of the objective species they convey a reference to the physical feelings prehended in the objective datum (*PR*:310).

In their dual role, qualitative sensa mediate the selective passage of feeling content passed along the strain-fibres of our physical universe. As a realized value they give each extensive region its determinate particularity. The ordered connections between these regions give rise to the systematic forms of order exhibited by strain societies, and these systems of extensive connections define the geometrical order of the universe. The characteristic forms of order that individuate Whitehead's societies can therefore be specified in terms of eternal objects; *customs are socially embodied eternal objects*. Whitehead's philosophy thereby reverses the prejudices of Aristotelian metaphysics; qualitative feelings of sensa are primary, and the enduring objects are the societies that emerge out of selective prehension of these sensa. Likewise, Whitehead reverses the prejudice of Platonism; for Plato the world of Eternal Forms was fully real and the world of appearances enjoyed a derivative reality. For Whitehead, eternal objects are mere potentiality, and become fully real only when realized in a physical actual entity.

3.23 The role of contrasts between physical and conceptual prehensions in strain feelings

The events that make up a living society are not wholly defined in terms of inherited customs. Life, above all, means novelty. The primary characteristic of life, Whitehead states, is "the capture of intensity"; this capture is dictated by the present, not the past (PR:105). Hence:

> Explanation by 'tradition' is merely another phraseology for explanation by 'efficient cause'. We require explanation by final cause. Thus a single occasion is alive when the subjective aim which determines its process of concrescence has introduced a novelty of definiteness not to be found in the inherited data of its primary phase. The novelty is introduced conceptually and disturbs the inherited 'responsive' adjustment of subjective forms (*PR*:104).

In order to explain the selective realization of patterns of potentiality prehended in a datum Whitehead suggested that each monad be viewed as a dipolar entity, with contrasting physical and mental modes of subjective feeling, each with their particular species of prehended object. Thus, the factual physical prehensions of a monadic occasion constitute a physical pole in the subject's feeling of it world, while the feeling of potential patterns of connection are mediated by the conceptual prehension of eternal objects. This prehension of patterns of potential relational connection constitutes the mental pole of the monadic occasion. This mental pole plays a vital part in explaining the way in which 'final causation' or purpose enters into the formation of a new monad.

Conceptual prehensions complement physical prehensions and provide a mechanism for relevant novelty to enter into the subjective aim of an occasion. Two related principles drive the subjective aim towards its end: maximizing the richness of aesthetic contrast and eliminating logical inconsistency. Conformal prehensions that simply reproduced the feelings of the past would result in a Universe characterized by a dull uniformity of feeling. Things are otherwise, Whitehead argued, because in addition to conceptually prehending the form of the eternal object manifested in the physical feeling, the subject also prehends novel but related eternal objects in the graded array of forms disclosed in the structured realm of ideal possibilities. Whitehead explains the availability of this graded array of eternal objects by introducing a primordial monad that he refers to as *God*.[38] In terms of each monad's subjective aim, God's key metaphysical function is grounding a domain of relevant subjectively felt potentialities of value. "God's purpose in the creative advance", Whitehead writes, "is the evocation of intensities" (*PR*:105).

In the context of this discussion, the notion of God need not concern us. Like the hardware and operating system of a computer program, the notion of God remains in the background of Whitehead's explanation of how societies and their constituent feelings evolve (except for when he deals with his notion of human civilization). What is important with respect to the subjective aim of the newly concrescing actual entity is that there is a factor that makes available a graded realm of pure eternal objects and that the concrescing subject's conceptual prehension of this domain, in conjunction with its physical prehension of the public domain of factual actual entities, triggers feelings of the evaluative relevance of specific eternal objects.[39] Such a

---

[38] I will not enter into the controversies surrounding the nature of Whitehead's God, which often relate to perceived theological implications of Whitehead's concept. However it is worth dealing briefly with the metaphysical functions of God.
It is in God's nature to save the perished past, just as it is to initiate the new moment of becoming. Consequently God is both the primordial instance of concrescence and the primordial instance of transition. Thus God is both ultimate ground for the physical realm of attained fact and ultimate ground for the conceptual realm of pure potentiality – the alpha and omega, so to speak of all things. As Alpha, God's 'consequent nature' is the initial condition that enables a novel monadic physical datum to function as efficient cause of a new physical prehension. As Whitehead states, "God and the actual world jointly constitute the character of the creativity of the initial phase of the novel concrescence" (*PR*: 245).As Omega, God's 'primordial nature' then plays a role in establishing a corresponding conceptual connection with a datum of ideal possibility and final causes. In Whitehead's words "The primary 'lure for feeling' [final causation] is the subject's prehension of the primordial nature of God. Conceptual feelings are generated, and by integration with physical feelings a subsequent phase . . . supervenes" (*PR*:189).
Thus God's consequent nature, as an existential ground of potentiality for the superjective nature of perished monadic subjects, provides the ultimate ground of potentiality that awakens the desire that is the seed of all Creativity. Likewise, God's primordial nature is the existential condition for the availability of the infinite domain of possible value disclosed subjectively through prehension of the realm of eternal objects. Note that the world is an essential precondition for the functioning of God's consequent nature; likewise, God's primordial nature is essential to the world's capacity to feel value. Hence God and the World are equally primordial in the constitution of the Universe. Without the World, God would be the empty vacuity of infinite possibility lacking any capacity for realization, and without God, the World would be meaningless attainment of actuality, lacking in purpose or value. These complementary aspects of God and the World are the focus of Part V in *PR;* they also are the focal point of the two essays that constitute Whitehead's philosophical last word "Mathematics and the Good", and "Immortality" (Whitehead, 1948 a & b).

[39] The subjective aim arises through the conceptual prehension of the eternal objects immanent in the initial physical prehensions *within the context* of the structured continuum of related eternal objects prehended conceptually in the primordial nature of God. The subjective aim is a vision of a possible path to attaining a satisfying unity of feeling through the selective elimination of some of the monads prehended in the datum. Hence the subjective aim acts as a final cause, guiding the subsequent process of selective elimination of physical vectors. This final cause is not an inflexible blue print, but is more an evolving assemblage of potential ways of linking together the structure of prehensive 'bricks' to obtain a satisfying outcome; this may entail successive valuations of available options and consequent modifications of the initial subjective aim. How the initial aim is modified in the course of the concrescence of a monad is primarily a function of the transmutation of the initial simple nexus into more complex forms of social unity present but not perceived in the initial objectification of the datum.

conceptual prehension of a functionally related form is referred to as a *reversion*. Reversions suggest novel ways for the datum to be unified; as such reversions involve the conceptual prehension of forms of potentiality that may offer a novel path to an enrichment of feeling.

As a result of reversions, an element of freedom and novelty can enter into each occasion. Through the realm of eternal objects the new subject prehends mutually exclusive possibilities latent in its situation, evaluates them, and makes determinate choices. Thus, each conceptual prehension values the felt eternal object (as a pattern of relational potentiality) 'up' or 'down' in importance, such that the eternal object becomes more or less important in the subjective aim guiding the process of concrescence towards its satisfaction (*PR*:240-1). The evaluation of eternal objects is thus a process of evaluating concepts specifying plans of action. In strain feelings the mode of evaluation is primarily aesthetic, and the plans take a very simple form.

In a strain feeling a reversion may take the form of a polar contrast to the form prehended in the physical prehension – the reverted form is one of simple negation. Instead of a positive charge of intense feeling there is its negation. The result is a contrast between the sensa prehended in the datum as the object of the prehensive feeling, and the negated subjective feeling of this form prehended in the new occasion.[40] Whitehead assumes that the prehensions of simple contrasts of this type can provide an aesthetically richer mode of subjective feeling and that this potentiality becomes the guiding eternal object in the subjective aim of the monad. He writes that "the admission of the selected elements in the lure, as felt contraries, primarily generates purpose; it then issues in satisfaction; and satisfaction qualifies the efficient causation" (*PR*:188). Whitehead argues that generating contraries in successive occasions is the basis of the *vibratory characteristic of all strain feelings*. He also proposes that a felt contrary of this type is the germ of consciousness.

Thus the domain of possibilities latent in the conceptual realm of ideal potentiality has a graded relevance to any concrescent subject and every process of concrescence begins with a prehension of strain feelings and the corresponding feeling of the most simple of eternal objects. A simple eternal object is irreducible to any more elementary forms of definiteness; however a group of these simple eternal objects, if abstractly related to each other in a determinate way, can form a complex eternal object. Complex eternal objects in their pure form represent abstract patterns of potentiality and can be joined together to form increasingly complex conceptual representations of possibilities latent in the brick-wall structure of the Universe. Granville Henry has argued that in this guiding role, "eternal objects can be understood as algorithms that (partially) describe the concrescence of the entity from initial causes to its final consummation" (Henry 1993:16).[41]

[40] "Thus in successive occasions of an enduring object in which the inheritance is governed by this complex physical purpose, the reverted conceptual feeling is transmitted into the next occasion as physical feeling, and the pattern of the original physical feeling now reappears as the datum in the reverted conceptual feeling. Thus along the route of the life-history there is a chain of contrasts in the physical feelings of the successive occasions. This chain is inherited as a vivid contrast of physical feelings, and in each occasion there is the physical feeling with its primary valuation in contrast with the reverted conceptual feeling. . . . In this way the association of endurance with rhythm and physical vibration is to be explained. They arise out of the condition for intensity and stability." (*PR*:279)

[41] Whitehead's notion of concrescence appears to owe a good deal to his reflections upon the foundations of logic. Logic is a specialized instance of a more general process of concrescence founded

Though Henry's analogy has its limitations, it does throw some light on the role of eternal objects as templates for selectively combining simple physical bricks of feeling into some specific relationally constituted form.[42] How these 'programs' are utilized depends upon the actual physical situations and the aims of the subject. Thus eternal objects in abstraction from the physical datum do not determine the meaning of a situation for a subject, but the contrast between physical *and* conceptual prehensions provide a foundation for the attribution of meaning to the possibilities that a monad prehends, and thus form the semantic ground of more complex forms of judgment.

3.24 Sensa and the complex semantic ground of higher order modes of feeling

If an eternal object could be prehended in total abstraction of any real-world context, it would be devoid of significance however, when prehended conceptually in the context of a concrete situation with its pressing emotional priorities, eternal objects acquire meaning. Each conceptual prehension discloses a dual semantic reference corresponding to the public and private sides of a prehension. In their role as public universals, eternal objects are the mathematical Platonic forms, and concern the world as an abstract universal, consisting of extensively structured medium of potentiality. In their subjective role eternal objects are elements in the definiteness of a subjective form of feeling, experienced as an emotion, an intensity, pleasure or pain (*PR*:290-2). Thus as public universals, eternal objects are individuated into objective classes of extensionally defined entities; while eternal objects as private modes of feeling have an intensional reference to specific forms of feeling.[43] The semantic content of eternal objects can therefore be defined in two ways, corresponding to the public and private sides of a prehension. On the one hand there is the eternal object as object of a conceptual prehension, interpreted in terms of the background provided by the public vector-side of a physically prehended datum; on the other hand there is the eternal object as a particular subjective mode of feeling, interpreted in the context of the private scalar-side of a physically prehended datum. Whitehead's formal theory of extension has its focus on the public role of eternal objects, while his theory of concrescence focuses on the private role of eternal objects.[44]

---

on aesthetic intermingling of feelings aimed at the elimination of inconsistency (see *MT*:82-6). "Inconsistency is that fact that the two states of things . . . cannot exist together (*MT*:72-3). Logic is simply a more specialized instance of a more general process of concrescence founded on aesthetic intermingling of feelings aimed at the elimination of inconsistency (see section 3.25, also *MT*:82-6).

[42] While the program-eternal object analogy is valid with respect to defining complex potential pathways towards prehensive unification (through the elimination of logical inconsistency), it undervalues the creative role played by eternal objects as aesthetic 'lures for feeling'; clearly logical consistency constitutes a far narrower mode of contrast than is in fact operative in human experience where, for example, feelings of importance may based on other qualitative factors aside from feelings of logical consistency.

[43] Whitehead discusses the significance of intensional and extensional forms of meaning with respect to logic in his essay "Indication, classes, numbers, validations" (1934:227-229). The essay is interesting if read in the light of related passages from his metaphysical writings, in that he was clearly thinking about logic in terms derived from his metaphysical system (and vice versa). At the time that Whitehead was writing modern intensional logics had not been developed, hence he did not have some of the formal logical tools available to Lambert.

[44] Whitehead divided his metaphysical systems into two theories to cover these respective domains. Book IV of *PR* is his theory of extension; it is concerned with the ultimate divisibility of the datum, understood as a set of potential regions defined extensionally. Book III of *PR* is his theory of

The public and private sides of eternal objects have a bearing upon Whitehead's approach to differentiating the intensional and extensional basis of semantics. For Whitehead, the intensional discrimination of something as a factor in our experience means that we experience the thing in the context of the situation in which it is presented. To grasp something purely in the mode of extension involves abstracting the thing from both its situation and all the potential functions that the thing affords. The thing becomes the bare logical subject indicated by the symbol that defines its individuality as an objectively unique 'thing'.[45] This relationship of indication still refers to the same thing, but the thing has been stripped of all intensional qualities; the symbol remains merely to indicate a determinate logical subject of a possible "unspecified variety of functionings" (Whitehead 1934:227).[46] In a given occasion of experience there may be no specific relationship of indication. If, for example, a tree becomes a factor in our experience it may be in the context of pleasant afternoon, blue sky, sun shining, green grass, birds whistling, and so forth, with these various elements all functioning intensionally in the constitution of the meaning of the composite situation. In this context the intensions 'blue sky', 'green grass', 'tree on the horizon', etc., all uniquely determine the composite intension formed by these various elements. However, if an angry bull enters our field of vision, most of these intensional properties will vanish as the situation is abstractly simplified to two distinct entities; 'that', the dangerous bull, and 'it', the tree as a safe refuge. Logic, Whitehead argues, takes this process of abstraction to an extremity, such that the only individuality that remains is akin to that "which is involved in counting" (*ibid*:228).

Whitehead's distinction between extension and intension in logic can be related to the semantic role of eternal objects in human vision. An occasion of visual perception can be explained in terms of two primary modes of feeling – causal efficacy and presentational immediacy; the former dominate in simple physical feelings, while the latter involve further transmutations that diminish the role of physical prehensions and enhance the importance of conceptual prehensions. In a complex process of concrescence, involving reversions and transmutations, the concrescent occasion begins with feelings in the mode of causal efficacy, and via transmutations, culminates with a feeling in the mode of presentational immediacy. This process is depicted in Figure 4.

---

concrescence, understood as the genetic analysis of subjective forms of enjoyment, and the way that these contribute to an ultimate satisfaction. Both are entwined as Public-Private dimensions of the process of Creativity. Whitehead viewed logic as a form of analysis dealing with classes of entities defined in purely extensional terms; that is, the intensional attributes of entities were stripped from them and the bare thing operated as a mere possibility for an "unspecified variety of functionings" (Whitehead 1934:227). These are the ultimate regions of Whitehead's extensive continuum, stripped of their qualitative particularity.

[45] The notion of indication is vital to Whitehead's metaphysical system. I expand upon its significance in a subsequent section, however in the context of narrow logical considerations, Whitehead defines indication as "the unique determination of a thing by the specification of some of its relationships to the functionings of some human body, and of some human intellect [where the selected relationships] possess a sufficient stability so that they can be repeated in different occasions of experience" (Whitehead 1934:227).

[46] There is an element of circularity in Whitehead's definition as he states that 'intension', is taken here to refer to some property or mode of composition "which is not among those considered in pure logic". He acknowledges that the boundaries of logic are conventional but asserts that intensional meanings have no place in logic (Whitehead 134:229).

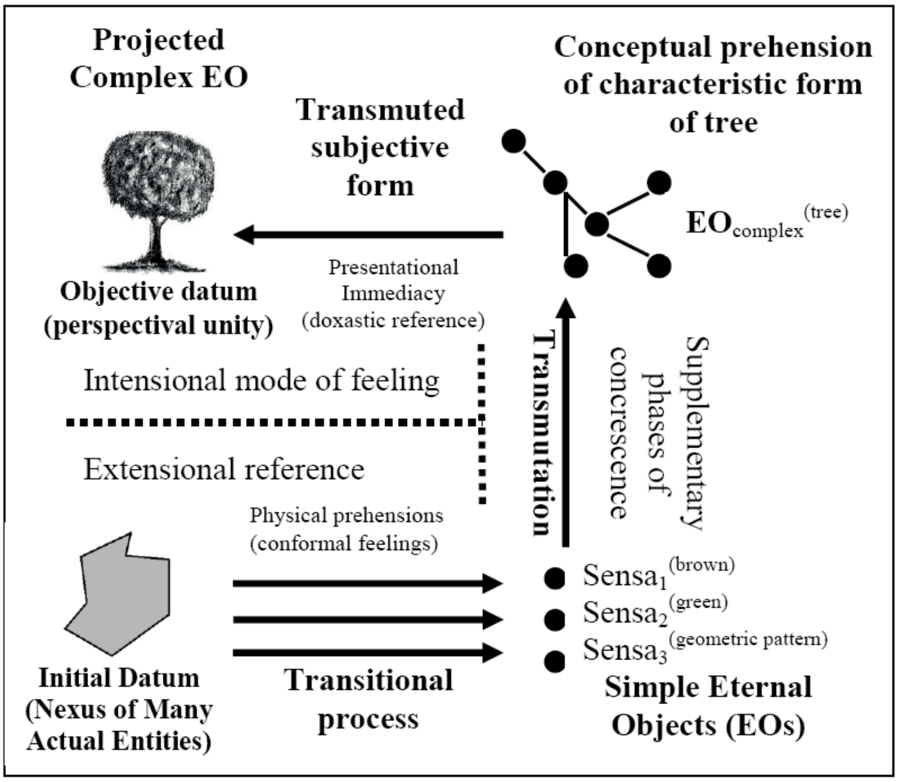


Figure 3.4 The role of eternal objects as a medium of semantic reference in the process of concrescence

The above diagram provides a simple depiction of the concrescence of a complex occasion; one that originates in simple sensa transmitted via causal efficacy from feelings that originate in the nexal set of regional connections (constitutive of a tree) and progresses through various transmutations to terminate in the perception of a tree (via feelings of presentational immediacy). Whitehead notes that *presentational immediacy* "takes the form of the creative imputation of the complex eternal object, ingredient in the bodily efficacy, onto some contemporary focal region felt in the strain feeling" (*PR*:316). Vectors of strain feelings mediate the transmission of sensa from that factual tree nexus to the human body. These feelings act as the datum for the transitional processes whereby the initial strain feelings are transmuted (via the dominant customs of the bodily-neuronal society), into the prehensive feeling of larger regions of brown and green along with specific geometric patterns. These are felt as patterns of potentiality in the initial datum, and guide the initial transmutation of that datum into a more differentiated objective datum. Subsequent supplementary cycles of transmutation may take place (not represented), mediated by the conceptual prehension of more complex objective eternal objects, such as leaves, branches and tree trunk, with their felt extensional reference to the physically prehended datum of actual entities. In addition, these entities will be prehended as a subjective form of possibility, with their sensuous and emotional appeal. Conceptual prehension of the complex eternal object instantiating the form of a tree might provoke a final recursive cycle of transmutation and unification producing a feeling identified extensionally with the initial datum but subjectively transmuted into the form of feeling that the subject experiences as the presence of a tree.

This unified perspective of the tree (in the form of 'presentational immediacy') will terminate the process of concrescence but will in turn become a superject qualifying a subsequent occasion in the subject's stream of consciousness. As such it will form part of the societal datum of potentiality functioning transitionally to influence the formation of a further monadic occasion in the personal society to which the human agent belongs. The pattern of intensional qualities embodied in the complex eternal object representing the tree, has a direct bodily tie (via vectors of physical feeling) to the extensional regions indicated by the perceived entity. As such, when the subject changes their bodily location with respect to the tree, further prehensions of sensa, and associated processes will tend to extensionally validate or invalidate the previous projected image. For example, it could become apparent that the tree was simply a two-dimensional cut-out.[47]

A complex eternal object 'tree' in a perceptive act thus has a reference to both the extensional and intensional properties of prehensions, in that they have an extensional reference to the eternal object in the prehended datum, and an intensional reference to a mode of feeling in the subjective form.[48] Both forms of semantic reference are necessary to define the concrete form of value realized in the physical prehensions of a monadic occasion. The satisfaction thus entails a unification of the subjective and objective species of eternal object mediating the evaluation of physical and conceptual prehensions.[49] This dual reference becomes extremely important in higher grades of feeling, such as visual perception, imaginative visualization, symbolic reference, and other forms of human consciousness. This dual reference is central to Whitehead's notion of propositional feelings which, as previously noted, are at the heart of all more complex forms of feeling.

---

[47] This example was inspired by some of Husserl's analyses, and many useful analogies could be drawn between Whitehead's work and Husserl's phenomenological studies, as has been suggested by a number of authors stretching back to Wolfe Mays and Herbert Spiegelberg. However in the interests of terminological parsimony I have decided not to discuss these connections.

[48] Whitehead divided his metaphysical systems into two theories to cover these respective domains. Book IV of *PR* is his theory of extension; it is concerned with the ultimate divisibility of the datum, understood as a set of potential regions defined extensionally. Book III of *PR* is his theory of concrescence, understood as the genetic analysis of subjective forms of enjoyment, and the way that these contribute to an ultimate satisfaction. Both are entwined as Public-Private dimensions of the process of Creativity. Whitehead viewed logic as a form of analysis dealing with classes of entities defined in purely extensional terms; that is, the intensional attributes of entities were stripped from them and the bare thing operated as a mere possibility for an "unspecified variety of functionings" (Whitehead 1934). These are the ultimate regions of Whitehead's extensive continuum, stripped of their qualitative particularity.

[49] In envisaging the compatibility of particular physical prehensions Whitehead used his notion of eternal objects to signify the qualitative Universal (in the philosophical sense) embodied in the particular. Some universals are jointly compatible are others are not. For example, a prehended region may be characterized by a particular shade of the Universal 'redness', and another prehended region felt qualitatively as a shade of green. These two regions are not compatible for unification in a single region exhibiting the same universal. However, some universals, such as colours, may be compatible with other universals, such as the 'degree of brightness'. In arriving at a conceptual prehension of the eternal object that expresses these compatible prehensions the incompatible forms of feeling are negatively prehended. The resultant eternal object will then serve as a guide in the elimination of incompatible physical prehensions. Hence in arriving at a satisfaction, many transactions may be eliminated by way of negative prehension, until the monad arrives at a single, unified and coherent sum of 'positive prehensions'.

3.3 Higher order societies: habits, propositions and belief

When considering the monadic members of a simple physical society such as constituted by strain feelings, the overwhelming evidence is that of uniform habits; in higher order societies the trend is away from uniformity toward a wider repertoire of behavioral responses. In both non-living physical societies, such as strains, and in higher types of living societies, such as people, habits play a dominating role. In both cases these involve the integration of conceptual and physical prehensions, but in the case of physical societies the purposive aim of each new occasion integrates itself with an element of the "realized definiteness of the physical datum of the physical prehension" and the result is conformity of habit that Whitehead describes as 'physical purpose', or 'pure instinct' (*PR*:184). Pure instinct (i.e., causal efficacy) is the lowest form of custom or habit and is solely a function of the elementary conformal mode of perception manifested in strain feelings (*SME*:79). In other words the eternal object guiding the conceptual prehension of the characteristic form of the society, when it becomes part of the integrated satisfaction, sinks back into the general physical feeling of the prehended society. This simple form of habit has mediated the expansion of the physical universe, leading Whitehead to note that:

> The most successful examples of community life exist when pure instinct reigns supreme. These examples occur only in the inorganic world; among societies of active molecules forming rocks, planets, solar systems, star clusters (*SME*:82).

Causal efficacy, or pure instinct, is a significant factor in human experience, as among other things it is the basis of our perception of time, "in the derivation of state from state", and of physical causality in the "immediate present moment conforming itself to the settled environment of the immediate past" (*SME*:35; 41). However, in living societies more complex forms of habit arise by building upon the 'pure instinct' embodied in causal efficacy and introducing a factor of novelty. This factor discloses itself in the public datum of a concrescing entity as, in Whitehead's specialized terminology, a *proposition*.

Propositions, in Whitehead's broad definition of the term, are at the heart of all higher order forms of habit or custom. As in the case of simple physical societies, living societies involve both physical and conceptual prehensions, however in the case of living societies it is not conformity, but the "flash of novelty" introduced by reversions that dominate. A proposition arises when this novel eternal object is integrated with a physical prehension originating in the physical pole in a manner that creates a hybrid entity where the eternal object conveyed indicates possibilities that transcend those conveyed by the physical feelings alone. Whitehead states that the main function of propositions is to act as lures for feeling. They accomplish this function by being introduced into the physical universe as real internal components of the actual occasions forming the datum of a new occasion (*PR*:185). "Such entities", Whitehead writes, "are the tales that might be told about particular actualities. Such entities are neither actual entities, nor eternal objects, nor feelings. They are propositions" (*PR*:256).

According to Whitehead, propositions play a vital part in most societies, primarily as a lure for feeling. Via propositions societies of occasions can pursue a transcendent goal: the 'satisfaction' of individual occasions become part of something bigger than that moment, as in the excitement and anticipation of the hunt. Through propositions societies become more than a simple medium of efficient forces, and are able to transmit transcendent goals that energize and give purpose and direction to our activity. "A propositional feeling", Whitehead states, "is a lure to creative emergence in the transcendent future" (*PR*:263).

Whitehead's notion of a proposition ("otherwise termed a *theory*") differs from its conventional logical sense in that in his metaphysical system "the primary function of theories [propositions] is as a lure for feeling, thereby providing immediacy of enjoyment and purpose"; hence in his scheme the logical applications of propositions are of secondary importance (*PR*:184).[50] Whitehead does not deny their logical function, he simply regards their function as lures as being of wider practical significance; the narrow approach to propositions he suggests, has arisen because unfortunately "theories, under their name of propositions, have been handed over to the logicians, who have countenanced the doctrine that their one function is to be judged as to their truth or falsehood" (*PR*:184).

Propositions are central to the way that customs are passed on from the public to the private domain; a prehended proposition can act to modify the subjective aim of an occasion. That is why Whitehead calls them 'lures for feeling'. For the most part, propositions play their persuasive role in societies that lack consciousness, and the subjective forms of feeling through which propositions are prehended are dominated by feelings of aesthetic evaluation, rather than judgements of truth or falsehood.

The role of propositions as lures for feeling can be illustrated with respect to the previously discussed perception of a tree (figure 3.4). The unified object prehended in the subjective mode of presentational immediacy will function, in its superjective form, as a proposition. The datum initiating and constraining the process of concrescence is made up of two vital societal nexi; on the one hand there is the tree, and on the other hand the person perceiving the tree. Just as the tree exists as an enduring object via the constant reproduction of its member occasions, so the mind of the person persists as a locus of identity through the reproduction of its member occasions. The proposition representing the tree is thus an item in that part of the perceiving subject's datum that is disclosing their past state. Our memories are therefore part of the brick-wall universe, and a key role of memories is the transmission of propositions.

In the interests of simplicity I avoided discussing the notion of propositions when describing the role of eternal objects in the semantics of perceiving a situation; however the presence of a proposition is crucial to all higher-order forms of

[50] It should be noted that in first-order logic a *theory* is a consistent set of sentences (the axioms) that describe a structure, meaning that there are no contradictions within the set of sentences. While there may be more than one structure that 'models' or satisfies a theory, in practical applications there is usually one particular structure which the theory is intended to interpret. For example, the axioms specifying Peano arithmetic are intended as an interpretation of the natural numbers and their operations. A particular interpretation will 'satisfy' or validate a theory in the same way that a particular interpretation will satisfy a sentence. For example, if I assert the proposition that "there exists an *x*, and *x* is called Bob and *x* is a murderer", then this proposition will be true (satisfied) if, in the context (domain) of this assertion, there is an entity called 'Bob' who is also a murderer. In an analogous sense, it could be said that Whitehead's propositions are intended as an interpretation of the datum to which they refer, and if valid, the datum will satisfy the proposition.

perception. Visually objectifying the tree involves assigning meaning to a region of our immediate environment. That is, the complex eternal object is predicating a prehended feeling of a region of our datum, and that predication is being felt as an objectification in the mode of 'presentation immediacy'. However, presentational immediacy by itself is barren, "because we do not directly connect the qualitative presentation of other things with any intrinsic characters of those things" (*SME*:24). In order to get beyond the 'mere appearance' of things, presentational immediacy must be related to causal efficacy. This temporal connection links the qualitative spatial presence to causal consequences; in Whitehead's language, this linkage involves a mode of feeling which he refers to as *symbolic reference*.

In practice, vision in the mode of presentational immediacy is always tied to causal efficacy in a mixed mode of perception that Whitehead refers to as *symbolic reference*. Through symbolic reference the use value of an object becomes reflexively incorporated into its visual presentation. For many species of organisms, such visual objectification has an important survival function, as for example, when a rampaging bull charges a hunter, the hunter's identification of a tree as a place of refuge might save their life. In circumstances such as these, Whitehead emphasizes that the primary way in which the propositions is experienced is as the 'entertainment of a feeling', rather than as the object of a judgment. He writes that typically, "Horror, relief, purpose, are primarily feelings involving the entertainment of propositions" (*PR*:188). In my example, with the thundering hooves of the buffalo behind him, the fleeing subject would prehend the objectification of the tree with feelings of relief and purpose. Thus, as Anthony Steinbock has stressed:

> Far from being lifeless, listless, existing in some pristine void of pure thought, propositions are primordially affective. They function as the religious vocation, the seduction for the enjoyment of aesthetic pleasure, the tantalizing tease, the pitch of the salesman, the marriage "proposal," the "come on," the "proposition" of the lover (Steinbock 1979).

Hence, in the words of Whitehead, when we see an external object, it operates as a symbol, communicating its significance to us, "because, in the long course of adaptation of living organisms to their environment, nature taught their use. It developed us so that our projected sensations indicate in general those regions which are the seat of important organisms" (*SME*:57). Whether through perception or imagination, propositions mediate this process by directing our attention to some situation of potential interest. The way in which an object symbolically communicates use-value is the basis of what James Gibson referred to as visual *affordances* (e.g., when visually grasping the physical presence of a chair the objectified presence communicates directly to the perceiver that this thing affords a place to sit). For Whitehead it is *habit*, in the mode of *reflex action* that mediates this communication of information.

Since thought hinders reflex action, and the rapid anticipatory responses that reflex action facilitates, the survival of "more developed types of living communities requires the successful emergence of sense-perception to delineate successfully causal efficacy in the external environment" (*SME*:82). Thus although reflex action may initially be fine-tuned by a combination of perception in the mode of causal efficacy and presentational immediacy (for example, when an infant is using tactile bodily modes of feeling to discern proximity of objects and using these to fine tune their visual perceptive sensibilities of position), the role of simple physical feelings in the mode of causal efficacy recedes as the visual response becomes reflexive and habitual. A semantics of success drives the process of developing the complex bodily societies that mediate vision, as the survival of such societies is dependent upon it acting as a datum that facilitates "decisions in the immediate occasion [that] should have the closest relevance to the concurrent happenings among contemporary occasions"; hence, reliable habits evolve and these are of "supreme importance for the survival of the enduring object [person]" (*PR*:318). Hence, for utilitarian reasons the habits that underpin vision rely upon propositions that help pick out, in a parsimonious fashion, significant affordances in the subject's habitat. As Whitehead notes:

> In every possible way, the more advanced organisms simplify their experience so as to emphasize those nexus with some element of tightness of systematic structure (*PR*: 319).

For Whitehead, the habits underpinning vision entail an automatic assertion that the proposition *is* true (visually we see the tree), while if grasped imaginatively the proposition is an idea of a potentiality that *may* be true for a given situation. The potential fallibility of propositions is particularly evident in imaginative propositions where the factual domain or datum that can physically satisfy the proposition is not immediately at hand. Nevertheless, the proposition is still systematically related to that hypothetical datum and its veracity or otherwise will determine the truth of the proposition. For example, if I am lying in bed at night and imagine that there is a tree in my backyard, there is a determinate structure of simple physical feelings that define a causal, spatio-temporal pathway of vector prehensions that link me to the imagined tree. This logically structured system of relations that directly links us to all that has occurred and indirectly links us to what is occurring contemporaneously or in the future, is termed by Whitehead an *indicative system*.

3.31 The crucial role of indicative systems in Whitehead's notion of a proposition

A fundamental aspect of visual experience is perspective. Perspective binds our field of vision into a structured unity, and it enables us to mentally 'switch perspectives' – to stand in each other's shoes. Mentally we can run through a journey both in a spatial and temporal sense, and others hearing a description of the journey can do likewise. Whether a proposition is a recipe for a meal, or a plan of battle, implicit in the ideas is a scheme of temporal, spatial and causal relations that link an initial situation to an outcome. Whitehead's notion of an indicative system provides an explanation of this implicit 'commonsense' perspectival background implicit in any proposition. The key criteria of a proposition for Whitehead, is the contrast between the eternal object and the particular subject of the concept; 'tree' alone is not a proposition, but 'that tree' is. For every proposition its *particulars must be indicated*.[51] He writes that:

> because the proposition concerns just those particulars and no others. Thus the indication belongs to the proposition; namely, 'Those particulars *as thus indicated* in such-and-such a predicative pattern' constitutes the proposition. Apart from the indication there is no proposition because there are no determinate particulars. (*PR*:194)

[51] See also my earlier discussion of indication in section 3.21. Whitehead seems to be implying that each subjective monadic prehension of a datum involves something analogous to the extensional reference which Whitehead refers to with respect to the abstractly stripped down extensional object of his formal logic.

In Whitehead's metaphysical system, the system of extensive connection at the root of our physical universe is an indicative system, such that the datum "reveals the mutual implication of all actualities" which means "that the actuality's own membership in the universal community of actualities is revealed to it by the indicative scheme of its own datum" (Nobo 1986:375). In other words, each entity grasps the whole community in itself; or, as Whitehead stated, through this systematic schema each actuality "sees itself as mirrored in other things" (*SMW* 207). Each new monadic entity is joined uniquely to the brick-wall universe and derives its relational identity from this social bond. Jorge Nobo has argued persuasively that this indicative schema (the extensive continuum of actual or potential relations prehended by each actual entity in its datum) is foundational to Whitehead's metaphysical system. It is the common ground of simple physical societies and the unified cosmic community that they constitute. It is the congruence between this systematic form of potentiality and the realized Universe of simple actualities that gives each monad's subjective experience, alongside its intensional qualitative sense, its determinate extensional reference to past and future events. "There is no solipsism of the present moment", Nobo states, "because the actual and potential history of the world is constitutive of the experiencing subject" (Nobo 1986:375).[52]

An indicative schema, in Nobo's sense, is re-enacted in each simple physical occasion. Whitehead provides a logical specification of a simple indicative system, proving that for "such an indicative system it is possible to give a precise indication of the members of the nexus from the standpoint of any one of them" (*PR*:195). If, for example, our Universe is a nuclear family, knowledge of the relationships between the members of the family from the standpoint of any one individual will allow the person with that knowledge to describe the family from the standpoint of any member of the family.[53] More complex types of indicative system can be defined (as for example,

[52] Nobo emphasizes that the efficient transitional and final concrescent phases of process involve "the contrast between the determinate *position* and the determinate *definiteness* of an actuality: the former is the *relational* aspect, whereas the latter is the *qualitative* aspect, of the actuality's essence" (Nobo 1986: 31). Because of the relational aspect the entity retains, in its repetition, the indicative spatial and temporal location of its original becoming. Earley (2008) has argued that Whitehead's metaphysical system of relations presupposes something analogous, in formal terms, to the intensional logic developed by Mertz (1996).

[53] For example, assuming it is a traditional Anglo-nuclear family, if Bob is the father of Jack and Sue, and husband of Mary; then it is possible to nominate the corresponding relational ties from the standpoint of any of the nominated individuals. Whitehead provides a more formal example:

> A nexus exhibits an 'indicative system' of dual relations among its members when:
>
> (I) one, and only one, relation of the system relates each pair of its members; and
>
> (II) these relations are instances of a general principle; and
>
> (III) the relation (in the system) between any member A and any other member B does not also relate A and a member of the nexus other than B;
>
> (IV) the relations (in the system) between A and B and between A and C suffice to define the relation (in the system) between B and C, where A, B, and C are any three members of the nexus. (*PR*:194-5)

Whitehead writes: "It will be noticed that in a nexus with an indicative system of relations, the subjective aspect of experience can be eliminated from propositions involved. For a knowledge of B and C and D as from A yields a proposition concerning C and D as from B. Thus the prevalent notion, that a particular subject of experience can, in the nature of the case, never be eliminated from the experienced fact, is quite untrue." (PR:195) projective geometry), but they illustrate the same principle.[54] The notion that an indicative system is embedded in all societies is not meant to assert that the complex feelings of the whole universe, past and future, enter into each occasion. Rather, Whitehead is asserting a bare, minimal structure of extensive spatio-temporal connections that play an indicative role linking the form felt here to a location there and then. This structure underpins what Steinbock has described as Whitehead's "crucial insight"; namely, "that a proposition 'contains' two subjects"; the logical subject and the percipient subject (Steinbock 1979).

For Whitehead, *every proposition* "presupposes some *general nexus* with an *indicative relational system*"; and this nexus "includes its locus of judging subjects and also its logical subjects" (*PR*:195). For Whitehead, the logical subjects are not parts of a sentence, nor are they concepts to compare with other concepts; rather, they are 'particulars' in the metaphysical sense of being "particular facts in a potential pattern" (*PR*:194). In any proposition these particulars must be indicated; the proposition applies to those particulars and no others. Hence the relationship of indication is essential to a proposition, as it ties together the percipient and the predicated logical subject the proposition. "Thus", Whitehead concludes, "in a proposition certain characteristics are presupposed for the judging subject and for the logical subjects" (*PR*:195).

Whitehead notes that what he has been referring to as 'indication' is essentially Aristotle's category of position; it is a way of acknowledging that each propositions meaning is supported by particular situations. Take the example of a judging subject entertaining the proposition 'the tree growing on the moon'. According to Whitehead, the proposition's meaning (and its truth or falsehood) will depend upon the context in which the judging subject prehends the proposition and the tacit reference that they associate with it. Language is vague and hence various meanings are possible. However one possibility is that the proposition means '*it* is a tree AND *it* is growing AND *it* is on the moon'.[55] Thus understood, the proposition presupposes:

(i) a definite set of actual entities called the 'logical subjects of the proposition' (the indicated nexus of actual entities referred to by the '*it*' which includes at least three implied logical subjects, the tree, the moon, and a lunar environment in which growing is possible);[56]

(ii) and a definite set of eternal objects called the 'predicates of the proposition' (eternal objects 'treeish', 'growing' and 'on the moon');

(iii) such that the predicates define a potentiality of relatedness for the logical subjects (no logical contradiction or inconsistency in predicating an *it* in the way specified, namely, that of a treeish entity growing in a lunar environment);

(iv) where the predicates form one complex eternal object or complex predicate (the 'tree growing on the moon');

---

[54] Whitehead's reflections on indicative systems of this type dates back to his early work on projective geometry where the systematic projective relationship between the sub-regions of a region are discussed in the context of Grassmann's calculus of extension and his later utilization of projective geometry as the basis of perspectivally unified monadic universe (see *UA*:pp.224ff., also Whitehead 1906).

[55] This example is a modification of the 'Socrates is mortal' proposition discussed by Whitehead (*PR*:264-5).

[56] Whitehead notes that it is pure convention to assume that a proposition has only one logical subject (PR:265).

(v) such that "for the 'singular' proposition there is the potentiality of this complex predicate finding realization in the nexus of reaction between the logical subjects, with assigned stations in the pattern for the various logical subjects" (i.e., that the complex predicate [iv] could be exemplified in the nexus of actual entities [i]);
(vi) and that a percipient or judging subject positively prehends the proposition and feels it as a lure to entertain that particular indicated possibility (*PR*:186; the examples illustrating each point are my own).

Thus Whitehead stresses that the actual world "in some *systematic sense*, enters into each proposition" through indicative feelings; propositions always tell a story that asserts some facts about an indicated potential pattern (*PR*:194). This applies to all types of propositions, including the two main types: 'perceptive feelings' (presentational immediacy) and 'imaginative feelings'. The distinction between these two types of feeling is based upon the way that the symbol and the logical subject of the symbol are related. Propositions felt in the mode of perceptive feelings involve a conjunction of indicative feeling and symbolic recognition; that is, the indicative feeling from which the logical subjects are derived and the physical recognition from which the predicative pattern are derived are the same feeling. Seeing a tree involves a coincidence of predicative symbol and indicative system (the qualitative sensa mediating the intensional qualitative properties and composition predicating the object that I see out my window [i.e., the intensions 'green', 'brown', 'tree-shape' that express the eternal objects that in conjunction make up the complex eternal object that functions as the predication 'tree'] and the chain of simple physical feelings emanating from the extensional region indicating the logical subject predicated by the proposition). When propositions are disclosed in imaginative feelings the indicative feeling and physical recognition differ (*PR*:261-3). The latter case is most readily illustrated with respect to language; the words on the computer screen mediate symbolic recognition of the predicate 'tree outside the window' but there is no physical identity between the symbol and its indicated logical subject (in other words, this symbol 'tree' is not a real tree).

In perceptive feelings the role of presentational immediacy subsumes to some extent the referential role of the symbol, while propositions felt in the mode of imaginative feeling involve a stronger form of symbolic reference. Symbolic reference is the mode of intellectual feeling exemplified in normal human consciousness. In mediating the use of language, symbolic reference draws upon conceptual prehensions and builds selectively upon more primitive causal and perceptive feelings to effect further massive simplification of environments. People do this through a further transmutation of the perspectival objectification in the datum and its corresponding subjective form "chiefly by concentrating on a certain selection of sense-perceptions, such as words for example" (*SME*:82-83).

3.32 Words alone do not constitute a proposition

With words there can be a chain of derivation, from symbol to symbol, whereby the ultimate ancestral derivation of the meaning of the symbol is lost in the past. In this case derivative symbols operate by habit, whereby a symbol is associated with meanings directly elicited by the symbol and its associations. Whitehead writes that:

> A word has a symbolic association with its own history, its other meanings, and with its general status in current literature. Thus a word gathers emotional signification from its emotional history in the past; and this is transferred symbolically to its meaning in present use (SME:84).

In part because words can transcend their origins, some analytic philosophers have treated language as an autonomous system of meaning; however, Whitehead rejected this tendency. Though language acquires its own self-referential autonomy, words alone do not make a proposition. They must be tied by a system of indication to a world, and equally, to a subject that is able to entertain the proposition. This is the background that underpins human commonsense abilities to make sense of symbolic communications. Hence, the phrase 'Socrates is mortal', though it may appear to suggest an independent fact, is for Whitehead entirely dependent for its meaning upon the existence of a general nexus that includes a judging subject and the systematic background of a world. Without this general nexus the words would refer not to particulars but to universals and Socrates, the historically significant philosopher that once lived in Athens, would be entirely irrelevant.

According to Whitehead, human language is infused with vagueness, such that propositions harbour various possible shades of meaning that depend upon the nexal locus of its judging and logical subjects. For example, for students sitting through a philosophy tutorial analysing syllogistic logic, the proposition 'Socrates is mortal' will have a very different meaning than the same words uttered by a dejected Plato after witnessing the death of his master. Whitehead notes that each potential meaning defines its own locus of subjects; "and only for such subjects is there the possibility of a judgment whose content is that proposition" (*PR*:196;264-5).

For Whitehead the key to transforming empty symbols into a proposition is the physical system of extensive connections that is the foundation of our bodily being and the corresponding abstract system of potentiality. This system provides the foundation of other indicative systems, including those that refer to purely imaginary entities or to imaginary constructs. Whether or not Socrates was an actual historical figure, the chain of indications that give meaning to the words 'Socrates is mortal', "starts from my conscious intuition of particular elements in the universe" and follows "a chain of indications which finally lead you via spatial – temporal indication and books to Socrates". Some of these links in the indicative system are inferential and hypothetical, as an actuality can only be indicated in relation to other actualities, however "the method of indication is such that I can describe how various elements indicated to me are indicated to you" (Emmet 2003:3 [citing Whitehead]).

Thus, Whitehead redefines propositions in terms of his metaphysical system. He concludes that there are four main types of entities in the world, two primary and two hybrid types. "The primary types are actual entities and pure potentials (eternal objects); the hybrid types are feelings and propositions (theories). Feelings are the 'real' components of actual entities. Propositions are only realizable as one sort of 'objective' datum for feelings" (PR:188-9). In all forms of transmutation, from causal efficacy to presentational immediacy to symbolic reference, mental and physical operations are intertwined, issuing in publicity and derived from publicity. This is apparent in the operation of propositions in human perceptions of all types, whether mediated by vision, the imagination, or language.

Throughout all modes of feeling: "The vector character of prehension is fundamental" (*PR*:317). As an ingredient in a datum, a proposition only becomes meaningful when it is positively prehended by a concrescing subject, to become a lure for feeling guiding the multi-fold datum of the primary phase of concrescence toward the "unity of the final satisfaction of feeling" (*PR*:185). In other words, while eternal objects in causal efficacy are simply communicating the uniform general character of the form of feeling that defines the simple physical entity, eternal objects via propositions are able to embed more complex templates for potential action within a physical datum that can act as routines, guiding the subject toward a potential satisfaction.

3.33 Belief, disbelief and suspended judgement

The role played by propositions as the focus of conscious affirmative judgments is spelt out by Whitehead in terms of similar elements and processes as those used to define perceptive feelings and imaginary feelings - which is to be expected - as according to Whitehead, "a direct affirmative intuitive judgment is a very sophisticated case of conscious perception" (*PR*:273). While propositions, by definition, are true or false in terms of their indicated situation, the making of such judgments is a function of a relatively rare type of intellectual feeling that people exercise sparingly. When they do, they utilize an inferential process to try and convert suspended judgements into feelings of belief or disbelief. I will not go into detail about the various transmutations and contrasts involved in the concrescence of such an occasion of feeling, but will describe the main intellectual feelings of satisfaction that are involved at the termination of this process; namely, belief, disbelief and suspended judgement.

All propositions arise from a special type of integration between a physical and conceptual feeling, where the datum of the physical feeling is an actual entity or nexus of entities, and the datum of the conceptual feeling is an eternal object. During the subjective process of concrescence the subject integrates into an objective datum the physical feeling of a determinate set of actual entities. This is the nexus of entities in the datum indicated by their felt physical relation to the subject of the feelings. As noted previously, these actual entities are the logical subjects of the proposition. The eternal object prehended in the proposition signifies some possibility about these logical subjects. In this role, the eternal object is the 'predicative pattern' qualifying the logical subjects with some specific form of definiteness. Whitehead illustrated the way these relationships enter into an occasion of judgement using the example of Socrates.

One of the various meanings of the proposition 'Socrates is mortal' could be that the Athenian philosopher of that name (spoken of by Plato and referred to in various historical records) did in fact live and was in fact mortal. The logical subjects in this case are the historical situation in Athens that includes amongst its actual entities the singular person of Socrates, his school, including the person of Plato, all participating in a systematic scheme in which mortality is realizable. The predicative pattern, 'Socrates is mortal', according to this interpretation does not refer to a cat of that name, but to the logical subject indicated by the systematic scheme that I have described. Judging the truth or falsehood of this statement would depend upon the relationship between the nexus objectified as the logical subject of the proposition and the eternal object that acts as the predicate pattern qualifying the logical subject. Since these logical subjects can not be directly experienced, judgement would depend on inductive arguments based upon the weight of available evidence. Depending on the evidence, the judgment could be one of belief, disbelief or suspended judgment (*PR*:264-5).

The same feelings characterize judgments in formal logic, except that in those circumstances the process of judgment depends not on gathering evidence, but on demonstrable self-evidence of the identity between the eternal object asserted in the predicate on the one hand, and the systematic schema of relations in the objectified nexus, on the other. Whitehead suggested that this 'mingling of forms' is the real foundation of both symbolic logic and algebra (Whitehead 1934:238-9). The three types of feeling provoked by the 'mingling' of and resultant contrasts between the patterns in both the predicate and the indicated nexus of a proposition are:

(i) feelings in the 'yes-form' (identical),
(ii) feelings in the 'no-form' (incompatible diversity), and
(iii) feelings in the 'suspense-form' (neither self-evident identity or incompatibility) (*PR*:270).

These judgements concerning the identity of the patterns in the nexus and the predicate are illustrated by Whitehead with respect to formal logic as follows:

(i) 'if p&q then q', is a 'yes-form' (identity of the mingled patterns);
(ii) 'p & not p', is a 'no-form' (incompatible diversity);
(iii) while 'p & q' is a 'suspense-form' (neither identical, nor incompatible) (Whitehead 1934:238).

3.34 The main function of intellectual feelings is to concentrate attention on potential actions, thereby modifying subjective aim.

In my analysis of Whitehead's metaphysical system, I have traced the concepts used to define the basic social processes constitutive of the Universe, from simple strain feelings through to feelings of judgment. This has involved discussion of various types of societies and their associated forms of feeling. In the most general terms, Whitehead's system involves just two types of society; non-living societies and living societies. The former is dominated by the instinctive physical purposes communicated by simple physical feelings such as are constitutive of strain societies; the outcome is the lawful order of nature. Living societies involve the emergence of hybrid feelings, in the form of propositions. All propositional feelings, whether perceptive, imaginative, intuitive, and so on, are species of 'intellectual feeling', in the broad sense that they involve the mental pole of the monad responding to propositions (*PR*:266). Thus in general terms, the foundational bricks of Whitehead's cosmos are made up of transactions defined by physical purpose, while the complexity that emerges in the organization of this brick-work edifice is the result of propositional 'bricks' of potentiality, in the form of intellectual feelings, supervening upon and giving direction to the physical potentialities latent in the simple physical purposes of the strain-societies.

As an element in a datum, propositions capture novel possibilities and make them available as a lure for future action; however, this practical function remains tied to the indicative schema and emotional energy mediated by simple physical feelings.

Both simple physical purpose, as communicated by the dominant customs in strain societies, and intellectual feelings, as communicated by the types of customs built around propositions, build upon comparisons involving physical and conceptual prehensions. These comparative feelings are the shared basis of all societies. However, in living societies novel potentialities are conceptually prehended, and societies evolve which capture, through their social habits, more flexible subjective aims. The higher level forms of intellectual feelings, such as are mediated by intuitive judgments, heighten the plasticity of the subjective aim by focusing our conscious attention on specific potential outcomes. Hence Whitehead emphasises that:

> Intellectual feelings, in their primary function, are [the medium of] concentration of attention involving increase of importance. This concentration of attention also introduces the criticism of physical purposes, which is the intellectual judgment of truth or falsehood. But intellectual feelings are not to be understood unless it be remembered that they already find at work 'physical purposes' more primitive than themselves. Consciousness follows, and does not precede, the entry of the conceptual prehensions of the relevant universals (*PR*:273).

Hence Whitehead's philosophy brings the intellect down to earth. Following the path pioneered by Reid, James and others, the intellect is not a God-like faculty apprehending a realm of ultimate truths, but is entirely woven into the historical fabric of our physical evolution. Its primary purpose is not to contemplate truth but to help social organisms focus upon goals that will promote their survival and enhance the flow of satisfactions that they experience in their ongoing transactions with the world. As Frisina has argued, Whitehead views intellectual activity as "an attempt to grasp and transform the inter-connective causal patterns that enable us to project satisfying expectations" (Frisina 2002: 137). In this role, intellectual feelings promote flexibility of subjective aim. Propositions become part of social narratives that bind a society together, and give both continuity of purpose, and a potential for creative modification of purpose to future members of that society. In this role, propositions become a crucial element in human history.[57]

Beliefs, for the most part, serve as the unquestioned background of routine action that bind human communal life together; they are taken-for-granted frameworks of action, and rarely subject to critical reflection. Taken-for-granted beliefs are for the most part enacted in a non-discursive fashion through the practical actions of everyday life, as people follow the lures provided by habits, routines, rituals, and reflex actions that structure social action. The promise of satisfaction promoted by the idea of a coffee break, involves a proposition functioning primarily, not as the object of a judgment, but as the object of emotional attention which is influencing my subjective aim; it triggers and guides the physical enactment of a routine that leads to the anticipated satisfaction. Likewise the primary role of intellectual judgments is that of creatively influencing our subjective aim; beliefs are a means to that end. Thus Whitehead writes that:

> . . . the main function of intellectual feelings is neither belief, nor disbelief, nor even suspension of judgment. The main function of these feelings is to heighten the emotional intensity accompanying the valuation in the conceptual feelings involved, and in the mere physical purpose which are more primitive than any intellectual feelings. They perform this function by the sharp-cut way in which they limit abstract

[57] See Porter's illuminating analysis as to how Whitehead's metaphysical system can be used to frame the various elements that make up human history (Porter 1975; 1979).

> valuation to express possibilities relevant to definite logical subjects. In so far as these logical subjects, by reason of other prehension, are topics of interest, the proposition becomes a lure for the conditioning of creative action. In other words, its prehension effects a modification of the subjective aim (*PR*:273).

3.35 The evolutionary potential of systems of formal logic

It is worth noting that while Whitehead did argue that the chief role of propositions in the Cosmos was that of directing and channelling subjective appetites, he did predict a glorious future for algebra and symbolic logic precisely because it would come to play a vital part in disclosing aesthetically and emotionally powerful patterns of potential activity. Thus, while the realm of consciously entertained propositions is a small island in the universe of non-conscious propositions that sustain the habitual or reflex actions which all living societies rely upon, he was keenly aware of the benefits of mechanizing logical systems of thought, and believed that human civilization was at the onset of an era that would witness the flourishing of symbolic logic as a tool for analysis and creative reflection across the full spectrum of human concerns.[58] Although his development of a formal calculus fitted to deal with complex human concerns remained rudimentary, he retained his faith that one day logical engines would transform our lives by expressing patterns of potentiality beyond those associated with bare abstractions of space, time and quantity (his metaphysical system can be viewed as the preliminary outlines of such a system). This is evident in the following ringing declaration of faith in the creative potential of algebra and symbolic logic:

> When in the distant future the subject has expanded, so as to examine patterns depending on connections other than those of space, number, and quantity -- when this expansion has occurred, I suggest that Symbolic Logic, that is to say, the symbolic examination of pattern with the use of real variables, will become the foundation of aesthetics. From that stage it will proceed to conquer ethics and theology (Whitehead, 1948:99).

Whitehead's appreciation of the creative potential of mathematics and symbolic logic is one of the strengths of his metaphysical system. It should be noted that a strong thread of logical consistency runs through his complex vision of Creativity. Whether he is speaking of strain feelings, perceptive feelings or feelings of intuitive judgment, the same public-private transaction structure is in evidence, and the movement towards the attainment of a unified satisfaction is one of selective elimination of inconsistent feelings, based upon self-evident contrasts between an objectified nexus of entities in the public datum and the form of feeling entertained by the subject. Both sides of this metaphysical equation are subject to modifications through various transmutations, but the same formula still applies. At the heart of this metaphysical process is a dynamic vision of habit.

[58] Whitehead's stated his belief in the introduction of his first major work, that "a science is gradually being created, which by reason of its fundamental character has relation to almost every event, phenomenal or intellectual, which can occur", and that the ideal of mathematics should be the creation of a calculus that facilitates reasoning "in connection with every province of thought, or of external experience" where human thought attains some precision (Whitehead, 1898:viii). Hence his career began with an examination into the potential of Boolean and Grassmannian algebras as "engines for the investigation of the possibilities of thought and reasoning" (*ibid*), and culminated in an endeavor to extend mathematical insights and logical rigor into the field of metaphysics.

3.36 Rethinking the nature of knowledge in terms of process grounded habits

In the Cartesian universe a gulf separated mind and matter; on the one hand there is the autonomous mind surveying the universe of ideas disclosed in the mind of God, and on the other hand the atoms that blindly determine the fate of the physical universe. In Whitehead's universe this gulf does not exist. Perceptive feelings begin at the foundation of things and concepts become available through the slow sedimentation of habits in structures of social action. Habits define societies, and societies mediate between the infinite potentialities latent in God, and the practical potentials available in our subjective worlds.[59]

Thus, central to the emergence of the simple physical societies at the base of our brick-wall universe are simple habits. More complex societies build upon this systematic societal ground to constitute more complex relational structures, using novel habits that simplify the increasingly diverse patterns of relational connections so as to individuate and objectify relevant contrasts in the social environment. Societies drive this evolutionary process by defining the conditions of their own success via the transmission of habits that work. Habits bind both sides of the private public equation together, with the society as datum providing the subjective aim with various templates for guiding the process of unification, and the subject as private form of feeling culminating in a public expression of that habit.

Higher forms of mental creativity are also building upon habit; language, as we have seen, is grounded in habits of body and mind without which our mental freedom and creativity would be impossible. Logic, likewise, is a convention based activity that has its necessary reference to a systematic background that is defined by habit. The human mind is a stream of processes ordered by an ever-changing set of habits and associated concepts. Representations, rather than being pictures held up before the mind's eye, or words that capture the form of the world in their own autonomous subject-predicate structure, are recipes for action that relate patterns of potential activity to indicated situations, disclosing potential satisfactions. Representations (as programs for action) mediate our expectations, and when our expectations are satisfied our perceptual anticipation is fulfilled. Such recipes are first and foremost oriented toward aesthetic ends, and are unconsciously adopted in the course of addressing feelings of pleasure and pain, rather than as conscious strategies to achieving logical satisfaction. Knowledge, from this perspective, is rooted in habitual forms of expectation; reason, or rationality, is a special form of habit mediated self-reflection that enables us to conceptually transform the inter-connective causal patterns sustained by habit and the satisfactions that habits mediate.

Possibly the most profound philosophical implication of Whitehead's analysis, is our need to reconsider the nature of language, and the basis of its functionality in terms of Whitehead's reconstructed notion of habit. Whitehead's analysis brings to the fore the role of the systematic backgrounds that encompasses the 'general nexus' of subjects implicated in symbolic reference. For Whitehead, the commonsense utility of

[59] According to Whitehead, the telos running through all of creation is that of evoking greater intensities of feeling, and societies are the medium that facilitate this process. Hence Whitehead writes: "God's purpose in the creative advance is the evocation of intensities. The evocation of societies is purely subsidiary to this absolute end" (PR:105).

language is grounded in the habitual patterns of activity that shape the Cosmic system of societies. Our body is part of that system and connects us with it. Reason is a specialized instance of a universal form of activity; it is not separated from Nature, but is a specialized expression of potentialities latent in all of Nature that retains its practical referential relation to that ground. In everyday commonsense usage of language that background is indicated in various ways, and the indicated systematic background constrains the vagueness of words. In the flow of a conversation, people are aware of and play to these habit-defined horizons of meaning. Hence, the mathematical ideal of rational exactness is not a glimpse into Plato's eternal realm, nor is scientific disclosure of the Laws of Nature giving us a glimpse of the divine blue-print for Creation. Rather, both mathematics and science are humanly constituted understandings that reflect the habitual basis of seeking satisfaction in our interaction with our world. As Warren Frisina concluded, Whitehead and Dewey challenge us to shift our basic metaphors "away from those that present knowledge as the static replication of forms and toward those that present knowledge as a way of forming relationships with the things that make up our world" (Frisina, 2002:234).

3.37 The resolution of the bifurcation problem: Sociality, rather than individuality, as the basis of mind.

I have now described the basic elements of Whitehead's metaphysical system. Clearly, the basic building blocks of his metaphysical system are not the unchanging atomic units of existence posited by Parmenides; rather, they are habits acting upon the vectors of 'physical feeling' that awaken desire and motivate action within a frothing plenum of short-lived monads. The existence of a monad is mediated by perspectivally unified feelings of past feelings. Guiding this perspectival connection to the world are habits; habits are emotionally catching patterns of activity. Through habits the primordial chaotic flux of noise becomes a community - a basal community that marches to a simple rhythm and melody. From these simple social foundations a rich social Universe of music evolves. Whether described in terms of transition and concrescence, vector and scalar processes, efficient and final causes, or public datums and private feelings, Whitehead's vision reflects the same presupposition; namely, that all creative effort is ultimately social effort. "Each task of creation", Whitehead stated, "is a social effort, employing the whole universe" (*PR*:223).

Thus, the history of Whitehead's brick-wall Cosmos is the history of its social transactions, its topology is defined by habit, and the bricks that it is built from are vectors of feeling. The Cartesian bifurcation problem, when viewed in terms of Whitehead's prehensive feeling, is dissolved. Though object-subject duality still exists, it is simply an expression of the public-private relation at the heart of every social transaction. Each transaction begins with an objective side in the public input of past events; this input is selectively unified in a private process mediated by a new subject. This subject terminates with a unified 'satisfaction' and in so doing becomes a 'superject', an object in the public domain with determinate relations to other objects. As a superject prehended in an initial datum, the subject's '*self' is a determinate physical subject in a factual world* (One physical actual entity added to the Many); however, for the agent unifying that world into a new perspectival unity, *the world becomes a unified conceptual object in the mind of the subject* (the Many become a conceptual One). Thus, the bifurcation paradox is resolved; the bodily subject in the physical room and the phenomenal room in the mental subject are simply the public and privates sides of Creativity.

From this perspective, the enduring human mind belongs to both the public and private sides of Creativity. As an enduring object, our subjective identity is a nexus of socially related events which every new moment of conscious thought draws upon and perpetuates. Human memories are recursively 'stored' in the public datum of past events but live in the private realm of subjective feeling. Thus for Whitehead, even our individual memories are a social inheritance, and all modes of feeling, from strains feelings through to imaginative propositions, are social accomplishments. The social domain is both the starting point and termination of each subjective act. Subjective acts are social transactions; they have their origins in society and their termination in society. 'Self', in so far as it is social mode of feeling mediated by prehensive activity of this type, is not a substance but an integrated mode of activity. Our personal unity, in Rescher's words, is "a unified manifold of actual and potential process . . . a structured system of processes, a cohesive and (relatively) stable center of agency" (Rescher, 1991: 85)

Thus, rather than view the ego as an enduring identity, and public actualities as mere thoughts in the mind of the subject, Whitehead makes the public-private-public flow of transactions the basis of continuity, and makes emergent public 'social structures' act as a constraint on and affordance for private processes. Whitehead noted that in "this inversion we have the final contrast between a philosophy of substance and a [process] philosophy of organism" (*PR*:151). Whereas Cartesian philosophy viewed thought as the 'output' of a subject, Whitehead viewed the subject (or 'superject' in his terminology) as the output of thought. Hence, "Descartes in his own philosophy conceives the thinker as creating the occasional thought. The philosophy of organism inverts the order, and conceives the thought as a constituent operation in the creation of the occasional thinker" (*PR*:151). And since these thoughts are the product of a societal effort, it is society, not the individual subject that creates the thinker. Not "I think, therefore I am", but "It thinks, therefore it is", where the 'it' in question is an expression of a social universe. Whitehead's metaphysical system provides a rational understanding of the 'it' in question and an affirmation of the insight that inspired his great philosophical efforts, as expressed in the lines from Shelley quoted at the beginning of this analysis:

> The everlasting universe of Things
> Flows through the Mind . . .

3.4 Whitehead's metaphysical system as a step towards a 'broad' vision of the role of formal logic and mathematical abstraction in human civilization

I began this chapter by citing Dewey's positive assessment of Whitehead's metaphysical system. In fairness to Dewey it should be noted that he strongly disapproved of Whitehead's tendency to join metaphysical speculation with the mathematical ideal of abstract systematization. Dewey believed that in attempting to formulate an abstract mathematical model of monadic connectedness Whitehead was substituting abstract logical connectedness for concrete existential temporal connectedness. The result was a spiralling concern with the formal architectural intricacies of his metaphysical system, in a dialectic which was mathematically, rather than experientially grounded. This way of thinking, Dewey argued, tends to repeat the process of abstraction and reification which Whitehead criticised so effectively in his analysis of the Cartesian-Newtonian scientific paradigm (Dewey 1951:657-8). It is a weakness, Dewey argues, that reflects the undue influence of Whitehead's mathematical background.

This is not a criticism that I accept. The use that Whitehead makes of mathematical processes of abstraction in formulating his metaphysical system reflects his vision of human history. Mathematical patterns have a fundamental place in our cosmos, and human history has been marked by a process of increasing awareness of this latent mathematical dimension. Whitehead's metaphysical system stresses the aesthetic importance of mathematical patterns during the process of concrescence; as eternal objects they relate sensa into a broader scheme of potentiality which in turn provide the necessary conditions for the development of higher forms of life. All societies, from strains upward, are defined by such patterns. While mathematical patterns alone do not give rise to sentient organisms they have played a key role – particularly in supporting those intuitive forms of measurement experienced in feelings of representational immediacy. Higher forms of intuition mean a correspondingly more important role for mathematical patterns, and this has been most evident with the emergence of human forms of sentience. Ideally, speculative philosophy should combine discursive, intuitive formulations of its main points, together with systematic, formal expressions of the same insights. Logic disciplines the imagination; it also provides it with wings to fly.

From the dawn of human civilizations, when humans first began to use tokens for counting and for representing the geometrical relations between stars, the creation of physical symbol systems has been an essential vehicle for enhancing human awareness of mathematical pattern. In Barry Allen's (2004) sense of an *artifactual ecology*, physical symbol systems are the work of artisans, shaping the social ecology that serves as a datum of human awareness. From this standpoint, computational systems are part of a technological continuum, reflecting the increasingly vital role that physical devices have played in representing complex eternal objects. These complex eternal objects transcend the mathematical patterns found in Nature; they involve novel patterns in propositions that represent patterns of potentiality relevant to the formation of human civilizations. In this role computational systems are an extension of natural language and other forms of symbolic representation through which we humans have constituted imaginative worlds which express ultimate intuitions concerning life and its value; every AI system is both an expression of such a vision and an artefact designed to serve the ends of the vision that it implements. As such computational systems are an integral part of the social environment mediating the potentialities of human consciousness.

Thus, Whitehead's metaphysical system enables us to go beyond the sterile opposition of the Cartesian-Newtonian's bifurcated vision of two distinct classes of material substance. The consequences of the bifurcated view of reality have been apparent in computer science, no more so than in polarized positions regarding the effort to develop a 'strong AI' akin to human consciousness. From the standpoint of Whitehead's philosophy, the entire debate is an instance of the 'fallacy of misplaced concreteness'. In Whitehead's philosophy, intelligence is a mode of feeling supported by a complex social environment. From a Whiteheadian standpoint, Classical AI embodies a fundamental misunderstanding of the nature of human intelligence. This does not mean, of course, that computing cannot be the foundation of very powerful information processing systems. It means simply that from a Whiteheadian standpoint, the systems of measurement and calculation that underpin scientific models, and current machine based computational implementations of these models, lack a variable element of subjective intensity that is essential to human consciousness. Indeed, as I shall briefly outline, computers are engineered specifically to eliminate the possible confounding effects that such an element could create.

3.5 A process metaphysical view of computer science and computer engineering

Whitehead's pan-experiential solution to the bifurcation problem makes mind and body co-dependent evolutionary partners in the development of consciousness. What this means in terms of a philosophy of mind has not been fully worked out, however Max Velmans (2009) has made a significant contribution to developing what could be regarded as a Whiteheadian philosophy of consciousness.[60] Velmans recognizes the validity of both a scientific third person perspective, and a phenomenological first person perspective on the phenomena of consciousness, and stresses that extension and location in phenomenal space are experienced realities that ground our scientific model of the world. However, his philosophy also stresses that this reality is observer dependent, and is a reflection of the evolution of the perceptual organs of the human organism. Thus it is a semi-realism. Like Whitehead's philosophy, it brings to the fore considerations that stand in the way of the Strong AI dream of producing a conscious machine. However this does not mean that machines cannot simulate some aspects of human consciousness.

In discussing the computational modelling of human behaviour, Bernd Schmidt (2000) has drawn the distinction between replication and modelling. He defines a replica as "an identical copy of the original" (2000:15). An artificial diamond, for example, would exactly replicate the physical structure and properties of a natural

---

[60] Velmans (2009) refers to his philosophy as 'reflexive monism'. In dealing with various contemporary solutions to the mind-body problem he argues that:

> Reflexive monism suggests a way of understanding these relationships that neither splits the universe into two incommensurable mental and physical substances nor requires consciousness to be anything other than it seems. It neither splits consciousness from matter nor reduces it to a state of the brain. Instead, it suggests a seamless, psychophysical universe, of which we are an integral part, which can be known in two fundamentally different ways. Whether one adopts the perspective of an "external observer" or a "subject", the embedding surround, interacting with brain-based perceptual and cognitive systems provides the supporting vehicle for one's conscious view, and what we normally think of as the phenomenal "physical world" constitutes that view. Nor does reflexive monism ultimately separate the observer from the observed. In a reflexive universe, humans are differentiated parts of an embedding wholeness (the universe itself) that, reflexively, have a conscious view of both that embedding surround and the differentiated parts they think of as themselves. (Velmans 2008:33 [on-line edition])

Although Velmans' work appears in the *Handbook of Whiteheadian Process Thought* (Weber 2008) Velmans does not directly engage with and build upon Whitehead's metaphysical system. In 1989, George Lucas reviewed a range of problems relating to the philosophy of mind and remarked that although relevant contributions are scattered through the literature, "the full dimensions of a Whiteheadian philosophy of mind have yet to be worked out" (Lucas 1989:145). That would still appear to be the situation.

diamond. From that standpoint, an artificial mind would be indistinguishable from the real thing; such a computer would pass the famous Turing test with flying colours. A model, on the other hand, is defined by Schmidt as "an abbreviated depiction of an excerpt of reality based on abstraction and idealization" (*ibid*). Such a model may functionally simulate natural phenomena, but they need not capture every aspect of the reality that they model. An aircraft, for example, models the flight of a bird, and a flight simulator models the behaviour of an aircraft. Social AI, from this standpoint, is an attempt to model aspects of human social action. All representations of human action, whether via art, sculpture, poetry, literature, theatre, film, scientific theory or computational simulation, is an abstract model, in this sense. From a Whiteheadian standpoint, a good model enables us to understand and predict how, in certain circumstances, people will behave, but these models should not be mistaken for the real thing.

Adopting a similar position, Henry and Valenza (2008) have argued that humans and current computational machines process information in fundamentally different ways, and that though many aspects of human activity may be modelled, and in some cases performed more effectively by machines, the mathematical and engineering techniques used to achieve this capacity mean that machines are a fundamentally different type of process, subject to different limitations and potentialities. The reasoning behind this conclusion is related to the engineering imperatives that guide the construction of computers.

Henry and Valenza's argument against the possibility of strong AI focuses on the role of measurement in computational systems. The basis of the isomorphic relationship between the computer as physical system and the computer as mathematical system is measurement: the physical system is designed to conform precisely, via specific types of measurements, to the mathematical system. Henry and Valenza, drawing on the work of Whitehead, argue that any system of measurement that objectifies actuality in sharp patterns, and removes the causally efficacious role of the subject's evaluative function from the picture is an abstraction from the reality of existence. Whitehead's metaphysical system is an attempt to specify how the sharp forms of perception and logical analysis that science utilizes, rest upon a vague, non-quantitative, but causally efficacious domain of subjective feelings. The matrix of subjective feelings mediates an evaluative process; ultimately, what is really satisfying prevails. The computer's form of satisfaction, by way of contrast, is a mathematical 'satisfaction' that is abstracted from the world of causally efficacious feelings.

Henry and Valenza's arguments rest upon the nature of the Turing machine as a paradigmatic expression of the computational process. Present day computational systems, as developed by computer scientists and engineers, are based upon symbolic programming techniques, and mechanical devices that run these programs. As outlined in chapter one, programming is a development of formal logic and mathematics, and finds its foundational expression in the abstract construction referred to as a 'Turing Machine'. As demonstrated by Turing, his hypothetical machine maps directly onto a mathematical system, and can be specified entirely in terms of a simple system of mathematical sets and relations that specify certain discrete states or discrete operations related to those states (Henry and Valenza 2008:194). Turing showed that in principle his hypothetical machine is able to compute any mathematical algorithm that can be solved using a finite (i.e. effective) computational procedure. Hence, in mathematical terms, Turing's hypothetical machine defines a state space of possible operations that a real machine must exactly replicate. This need for an isomorphic relation – such that a precise correspondence obtains between both systems - has special significance for the physical engineering of a real computational machine.

In attempting to build a mechanical computational system, the computer engineer inverts the usual scientific relationship between a mathematical model and the actual physical system that it replicates. For example, in physics the aim is to develop a mathematical model that can associate a mathematical space with a corresponding physical system, so that a "mathematical transformation in the associated space can be used to predict or, in some cases, even control the behaviour of the system in question" (Henry and Valenza 2008:195). Such a scientific prediction is possible only when there is a correspondence between the mathematical model and the physical world, as ascertained by measurement. Empirical measurement, as employed by Newton in developing his law of gravity, is fundamental to science. Measurement is the basis of determining the empirical accuracy of a mathematical model; predicted outcomes in the hypothetical world must be measured against actual outcomes in the real world, and the mathematical system designed so that it corresponds with the behaviour of the physical system. In physics, this attempt to map a real physical system onto a mathematical system will always be constrained by limitations relating to the parametric and instrumentation used to obtain measurements. Thus, no matter how much a model is fine tuned to accurately map a physical system it will always be an imprecise approximation of the reality. The computer is an inversion of this relationship.

Rather than proceed from a physical reality to a mathematical model (with the evolution of states of the mathematical model designed to match, as determined by measurement, the evolution of the states of the physical system), a computer proceeds from a mathematical model to a physical system (with the evolution of states of the physical system designed to match, in a manner determined by measurement, the evolution of the states of the mathematical model). Take, for example, the role of a transistor in controlling a capacitor, as employed in the operation of computer memory. In practice, memory capacitors may exhibit decay or excess in their charge, but a measurement procedure is used so that any deficiency or excess charge within certain wide parameters is ignored. What is important, if the physical machine is to successfully replicate the mathematical structure of the Turing Machine, is that its capacitors represents the binary values of 0 or 1, such that the transition of the transformation of the physical machine from state to state can be driven by the digital logic embedded in the machine so as to exactly match the mathematical transformations executed by the Turing Machine.

From the standpoint of computer science, the states of interest in the physical computer are those that have been engineered to correspond exactly to states in the mathematical space defined by the Turing Machine. Hence the computer engineer's task is not to develop a mathematical model that approximates the functioning of a mechanical system but the inverse; to develop a machine that exactly simulates the functioning of Turing's mathematical construction.[61] Henry and Valenza sum up this peculiar inversion of the relationship between physics and computer science, by noting that "the computer represents the given mathematical space, rather than the other way around", and that:

> In consequence of the discreteness of its associated abstract state space, the only physical states of interest for a computer qua computer are discrete in two dimensions of measurement: time and descriptive content (the attribute measured) (Henry and Valenza 2008:197).

Thus, computers are systems based upon causally efficacious powers and associated supplementary prehensive processes that have been engineered to remove reversions, and intrinsic aesthetic contrasts as a source of creative contingency; to the extent that machines can function 'creatively', it is through the application of mathematical algorithms in self-modifying loops of efficient causation. That is, machines are engineered to reflect the mechanistic bias of the Cartesian model of Nature. The subjective, qualitative realm is not unified with the objective world of extended matter; rather, the subjective (as the functionally efficacious 'creative' factor operative in the concrescence of an occasion of process) is constrained by the design of the datum from playing any significant role in the operation of the computer. Hence, computers have no intrinsic internal dimension of qualitative evaluation analogous to human aesthetic sensibility. In humans, the creative contingency associated with reversion and aesthetic contrasts has been socially enhanced and magnified, resulting in a datum that supports conscious flights of imaginative fancy, with anticipated satisfactions sometimes leading to novel forms of fulfilment. The bifurcation of reality, that Whitehead attempted to overcome through his metaphysical system, provided no explanation as to how the human organism could evolve from a simple atomic structure. Whitehead's vision, on the other hand, suggests a mechanism whereby objectivity and subjectivity are woven together as the functional basis of evolving forms of feeling, habitual patterns of social sensitivity that sustain and drive more complex patterns of purposeful activity.

Hence, Henry and Valenza argue that the inference may be drawn from Whitehead's metaphysical system that: "a computer system, as an isomorphic replica of a given mathematical system via a system of measurements, is nothing more than a set of measurements and must therefore be noncreative and nonconscious" (Henry and Valenza 2008:200). Thus, in the terms employed by Schmidt (2000), although computers may model human creativity, they cannot replicate human creativity. Unless, that is, they are engineered on the basis of different principles than those embodied in the Turing Machine. Lambert's Post-Classical paradigm, from this perspective, is not an attempt to replicate human commonsense; rather, it is an attempt to build a better model, one that mathematically simulates aspects of the human creative processes that had been misunderstood and neglected in the Classical model.

3.6 From Whitehead to Lambert: a Post-Classical path to engineering commonsense

---

[61] An appropriate physical description of the computer will therefore focus on the machine's capacity for discrete representation, both in terms of the evolution of states at discrete time intervals, and in terms of the measurements that define these states.

Lambert's notion of 'engineering commonsense' was meant to be provocative, and from the previous remarks it is clear that Post-Classical AI is an attempt to model rather than replicate human intelligence. Specifically, Post-Classical AI is a model that draws upon the potentialities of physical and mathematical systems to engineer a computational system that simulates the operation of key elements of human commonsense. The key distinction between Classical and Post-Classical AI in this endeavor is the set of abstractions that guide the modeling process. Classical AI was grounded in certain abstractions that formed a hierarchy of conceptual dependency, the nature of which was outlined in the previous chapter. Post-Classical AI is based upon a different set of premises that have framed a different notion of commonsense. Whitehead's philosophy is one formulation of this alternative *process* model of commonsense.

In order to fulfill the promise made at the start of this chapter I will now sketch the barest outline of how Whitehead's ideas can be used to frame Lambert's Post-Classical AI. The key to Post-Classical AI is the machine agent's use of physically instantiated patterns to predicate indicated logical subjects, thereby simulating the role of propositions and providing a basis for a machine version of 'belief'. I will detail how this is accomplished in subsequent chapters, however in elementary terms it is a process version of the chain of conceptual dependency outlined at the onset of this chapter, where an ontology of *what there is* → (grounds) *what is represented* → (grounds) *what is known.*

In translating this conceptual dependency into a machine language, Lambert:

(i) first provides a formal specification of Whitehead's brick-wall universe that incorporates all entities - including agents - as fragments of process (*what there is*);

(ii) he then suggests that enduring objects can be individuated as recurrent patterns of causal influence, and shows how these physical patterns in the brick-wall universe can be objectified using simple or complex customs (*what is represented*);

(iii) lastly, based upon indicative systems that pattern-match customs against reality, he shows that his formal system of representation can be used by an agent to form beliefs about the world (*what is known*).

The formal techniques used to give effect to each stage in this process vision are then employed in the formulation of a multi-agent based model for an artificial mind.

In the decade and a half subsequent to the writing of *EMWC* Lambert has had the opportunity to further develop and implement his Post-Classical vision using multi-agent systems of increasing complexity and capacity. This has involved going beyond the initial proposal for engineering commonsense and considering questions related to integrating 'social machinery' with human social organizations. He has done this without overt reference to the arsenal of Whiteheadian concepts that this chapter has introduced, so it remains an open question as to whether my Whiteheadian framework promotes a deeper appreciation of the potential of Post-Classical AI. Does process philosophy provide a useful canvas against which to address the many philosophical and practical questions raised by Lambert's endeavor to engineer 'social machines'? In order to address this question I will open the *EMWC* lego® box, and attempt to show how Whitehead's philosophy can be used to throw light on Lambert's strategy for engineering a Post-Classical mind.

Chapter 4

The Formal Process ontological foundations of Post-Classical AI

Chapter 4 – Section Headings

4.0 Introduction: from philosophical to computational ontology

> [It] is a misconception to suppose that process philosophy, siding with becoming, rejects being. Rather, it is a doctrine of being in becoming, permanence in the novel; by contrast, philosophies of being are doctrines of becoming in being, novelty in the permanent (Hartshorne, 1965).[1]

In the previous chapter, I outlined a Whiteheadian system of metaphysics that provided a discursively formulated model of the functional basis of human commonsense. I concluded by suggesting that Lambert's Post-Classical project can be viewed as a mathematical translation of key elements of Whitehead's metaphysical model, whereby an ontology specifying what exists grounds notions of representation and belief, providing the basis for a social approach for engineering machine intelligence. The aim of the next four chapters is to examine in detail Lambert's process version of the hierarchy of conceptual dependency, and trace how Lambert converts a metaphysical conceptual vision into a Post-Classical mathematical strategy for engineering commonsense.

Lambert's aim in Part II of *EMWC* is to frame a computational approach to artificial intelligence within a general *principia intellectus* (philosophical mode of understanding) that will avoid the rationalism, solipsism and linguistic realism of Cartesian AI, and "deliver an alternative functional architecture for commonsense machines" (*EMWC:*133). Underpinning this strategy is a notion of how mathematical structures and models relate to the real world. It will become apparent that the 'real world' that Whitehead describes in terms of his brick-wall universe of prehensions is intimately tied to the structure and functional role of the computational ontology upon which Lambert builds his Post-Classical strategy.

The focus of this chapter is on Lambert's formalization of the metaphysical notion of process. Lambert conceptualizes the notion of existence in terms of process, and then builds upon a formal theory of processes to develop process-based interpretations of

[1] Quoted by Lambert (*EMWC*:133)

representation and belief; these concepts are then form the basis of a proposal for a Post-Classical artificial mind. This Post-Classical architecture bridges the divide between the Situated AI dictum that the 'world is its own best model', and the Classical AI notion of symbolic models (abstract 'worlds') that represent some aspect of the real world in a simplified form. The divide between Situated and Classical AI reflect the bifurcated conundrum at the heart of Cartesian paradigm; a philosophical divide that gave rise to the behaviorist view in which organisms are shaped by the physical world (the world is the environment imposing the conditions of selection that determine the evolution of the physical agent), and a cognitivist vision in which the world is shaped by the mind of the agent (where the agent is creating models of its world and plans of action which have no direct dependence upon the world). Lambert's Post-Classical strategy can be viewed as an endeavor to bridge this divide.

As outlined in Part II of *EMWC*, Lambert's strategy for engineering commonsense rests upon a concept of process; process is then used to explain representation; representation to explain beliefs; and beliefs to explain mind. As a consequence post-classicism is advanced in four steps, each of which provides a mathematical interpretation of the following subjects in turn: existence, representation, belief and mind (chapters 6, 7, 8 and 9 in *EMWC* respectively, as summarized in chapters 4, 5, 6 and 7 of this work). In this chapter, I will firstly make a terminological distinction between the metaphysical and computational notions of ontology, focusing on the mathematical role played by computational ontologies as models for interpreting their worlds. I will then consider the semantic challenge of linking a formal structure to the domain that it purports to describe and argue that Whitehead's metaphysical understanding of the extensive continuum provides a basis for linking formal concepts to the world of appearances, and that it is therefore a useful conceptualization for guiding the development of a formal ontology. I will then describe the metaphysical ontology developed by Lambert in *EMWC* and relate it to Whitehead's 'brick-wall universe'.

4.1 Metaphysics and computational ontology

In Lambert's Post-Classical paradigm, metaphysics has both a broad guiding role, in a meta-theoretical sense, and a narrow ontological role. In its broad role, process metaphysics suggests a different set of presuppositions to frame our understanding of mind, world and commonsense. In this sense process metaphysics is providing an alternative to the Cartesian vision that has guided the Classical AI approach to modeling intelligence. Lambert also employs process metaphysics in a more restricted sense, which is closely related to the computational notion of ontology. It is this latter, specialized notion of ontology that is central to this chapter; the focus is on Lambert's effort to specify the ultimate metaphysical dimensions of existence that ground his Post-Classical notions of representation and belief. However, before detailing Lambet's approach to this task, I will first clarify the distinction between the philosophical and computational meanings of ontology.[2]

The computational notion of ontology, as employed by engineers engaged in knowledge representation (KR), has a specialized meaning that is related, but not

[2] This analysis will draw upon the work of Gruber (1995; 2008), Guarino (1998), Zúñiga (2001) and Smith (2003b).

identical to the philosophical notion of ontology. For example (and to foreshadow developments that will be outlined in subsequent chapters), Lambert's current ontology differentiates five layers of representation, beginning with what he refers to as the metaphysical layer, ascending through environmental, functional and cognitive layers, to the social layer. The framework is oriented by a process vision, but each layer has its own *conceptualization*, *ontology* and *theory*. The notions of conceptualization, ontology and theory have a specific meaning when used in this computational context, quite distinct from, for example, the philosophical sense of ontology, or the sociological sense of theory.

Ontology, in modern philosophical systems (such as Whitehead's process philosophy or Edmund Husserl's phenomenological metaphysics), is an attempt to describe the class (or classes) of entities necessary for a complete and exhaustive description of the Being and Becoming of all that goes on in the universe. Metaphysics and ontology are more or less synonymous in this philosophical sense; both refer to the attempt to develop a minimal system of concepts adequate for a truthful description and classification of all that exists (Smith 2003b). Philosophers have sometimes followed the example of Husserl and divided their metaphysical ontology into formal and material systems, with formal ontology specifying the abstract universals common to all the phenomenal domains of existence (such as Husserl's mereology of parts and wholes), while material (or regional) ontologies will specify the properties of specific phenomenal domains (such as Husserl's distinction between material ontologies made up of material, ego-alter, and symbolic relations).[3] In sum, the formal ontology details those universal characteristics common to all phenomena, while material ontologies differentiate the phenomenal realm into domains on the basis of some distinctive characteristic form. Thus, ontology for the philosopher is an attempt to provide a comprehensive answer to the question: "what exists?" Ontology, as usually employed in computer science has a different meaning.

In computer science, the terms *conceptualization*, *ontology* and *theory* are interrelated. A *conceptualization* specifies the key concepts necessary to describe a domain of interest and in this sense is similar to the philosophical meaning of metaphysics or ontology; the key issue for a conceptualization is the semantic challenge of slicing up reality into a system of meaningful and consistent concepts that adequately capture the facts about the domain of interest.[4] A computational o*ntology* translates the conceptualization of a domain into a formal logical (mathematical) structure. This structure consists of an underlying set (a Universe) together with formal interpretations (a vocabulary) of the function, constant and relation symbols that determine what aspect of the Universe the ontology investigates. To provide a simple example, if the Universe is the Smith clan, then the underlying

[3] Body, ego and *geist* (as used by Husserl to define his material ontologies) can be translated into the the Whiteheadian domains dominated by causal efficacy (pure instinct resulting in conformity to past behaviour), presentational immediacy (propositions conveyed by reflex habit directly from the objects signified), and symbolic reference (propositions based on customary meanings conveyed about objects indirectly signified by language). Note that Whitehead's theory of extension is an effort to specify the structure of potential relations implicit in the datum of a concrescing entity, and hence grounds the 'transmuted' societal structures that have been objectified in the datum (i.e., the prehensive ground perspectivally unified in the process of obtaining the satisfaction that defines the particular subject/superject).

[4] For a comparison of computational conceptualizations, that includes Lambert's approach, see Lambert, Saulwick, et.al. (2008)

set is all people with the name Smith, the vocabulary might include only constants (Tom, Bob, Sue, etc.) and relations, (brother of, sister of, etc.). Depending upon the vocabulary a different structure is obtained, and a different interpretation of the domain will be produced (i.e., different specifications of the kinship structure will generate different valid sentences about the domain).

Any structure will generate a set of valid sentences concerning the domain, and this set of sentences is referred to as the *theory* of that structure. When the structure is a model of an ontological domain, a set of sentences that completely describe the domain may be referred to as axioms. This understanding of ontology is strongly indebted to mathematical model theory.[5] Gloria Zúñiga (2001) has proposed two definitions that relate closely to Lambert's terminological distinction between conceptualization, ontology and theory. She writes:

> An *information system* (IS) *ontology* is an axiomatic theory made explicit by means of a specific formal language. The IS ontology is designed for at least one specific and practical application. Consequently, it depicts the structure of a specific domain of objects, and it accounts for the intended meaning of a formal vocabulary of protocols that are employed by the agents of the domain under investigation (Zúñiga 2001).

Thus, in computer science, the notion of conceptualization is closer to the philosophical sense of metaphysics or ontology, in that conceptualization usually

---

[5] Model theory is a branch of mathematical logic that considers the semantic connection between structures and sentences. Model theory is an extension of Tarski's approach to the definition of truth (cf. Ch. 1, footnote 8). As Lambert (2009) observes, Tarski introduced the idea that the truth of a sentence can be expressed by a metalanguage statement, citing "'it is snowing' is a true statement if and only if it is snowing" as an illustration. As explained in my introduction to logic, in chapter one of this work, a model theoretical representation specifies truth in a particular interpretation. Hodges (2001) illustrates the general notion using the following analogy. The sentence, "he is killing all of them", could be true or false depending upon who 'he' is and what 'all of them' refer to. By assigning meaning to these terms, for example, by noting in the local newspaper that 'he' is Alfonso Arblaster of 35 The Crescent, Beetleford, and that 'them' are the pigeons in his loft, the sentence is defined as true for that interpretation (i.e., the interpretation *satisfies*, or in model theoretic terms, is a *model* of the sentence).

Model theory is concerned with the interplay between mathematical structures and sentences understood in a purely formal sense. Although Suppes (1960) demonstrates that there is a linkage between the notions of model and theory as applied in mathematics and the meaning of these terms in the empirical sciences these practical applications are not relevant to the purely formal concerns of model theory.

In formal terms, a structure involve a *domain*, *signature* and *assignment functions*. By using these elements to formally specify the syntax and semantics of a language it is possible to generate a structure. Structures assign meaning to sentences constructed in the formal language of that structure. A sentence may be true relative to one interpretation but not another, as for example, when a statement about the properties of numbers is interpreted using the set of real numbers as the structural metalanguage, as opposed to complex numbers.

When a structure is used to interpret the meaning of sentences in a formal language it is described as a *model* for that language. When a structure validates or satisfies a sentence or theory (set of sentences) then it is described as a model of that sentence or theory. For example, if the statement (sentence *S*) is true when interpreted in terms of the set of real numbers, then that particular interpretation *I* is a *model* of *S.* This is equivalent to saying that *I satisfies* S, or in symbols '$I \vDash S$'. As Hodges (2001) notes, to say "that *I* is a model of S is to say that S is *true in I*, and so we have the notion of *model-theoretic truth*, which is truth in a particular interpretation". Hodges (1996) provides the standard introduction for students with a mathematical background; Hedman (2004) is an introductory text aimed at students with a background in formal logic.

refers to a discursively formulated model of the things or processes constitutive of a specific empirical domain. However, unlike philosophical ontology, an IS conceptualization is also concerned with the way that language is used to refer to this domain. Clarifying relations of extension and intension, which for philosophy is an epistemological concern, can be significant for an IS conceptualization (as will become apparent in Lambert's ontology). In addition, Guarino makes the point that two IS ontologies may use different axiomatic theories, each specifying a distinct state of affairs, with each attempting to express the same shared conceptualization of the situation (Guarino 1998). Similarly, Lambert's formal 'metaphysical' layer of his KR ontology is but one possible way of formally capturing the situation specified by a conceptualization of a domain. In the words of Zúñiga:

> A conceptualization is the universe of discourse at work in every possible state of affairs for the particular domain (or domain space) of object that is targeted by the IS ontology (Zúñiga 2001).

The relationship between conceptualization, ontology and theory can be illustrated using a simple conceptualization, expressing an Anglo-Australian genealogical system that is implemented as an ontology using the Prolog (*Pro*gramming in *Log*ic) programming language.[6] Programming in Prolog consists essentially of describing a situation (represented in the form of a database) and asking questions about that situation (testing whether propositions are true in terms of the database). For example, a Prolog database may consist of a list of genealogical facts (as constants) and rules (relating to variables), that together explicitly or implicitly represent the relationships between family members, expressed in the formal syntax of Prolog. The ontology is thus implemented in terms of formal structure, consisting of a Universe that provides an interpretation of a vocabulary of constants, relations and functions; the interpretation assigns elements, functions or subsets of the Universe to corresponding constants, functions or relations and thereby generates the formal sentences (clauses, consisting of facts or rules) that make up the database.[7] When using Prolog, rules can

[6] For details on Prolog see Sterling & Shapiro (1994) and Bratko (2001).

[7] The following clauses and their natural language translations give some idea of the structure of such a program:

Prolog program (database facts)	**Ordinary Language**
father_of('Tom', 'Catherine').	Tom is the father of Catherine.
father_of('John', 'Mary').	John is the father of Mary.
father_of('John', 'Samuel').	John is the father of Samuel.
mother_of('Catherine', 'Mary')	Catherine is mother of Mary.
mother_of('Catherine', 'Samuel').	Catherine is mother of Samuel.
married_to('John', 'Catherine').	John is married to Catherine.
male('John').	John is male.
female('Catherine').	Catherine is female.
male('Samuel').	Samuel is male.
female('Mary').	Mary is female.
. . . etc.	. . . etc.

A query then consists of syntactically matching a proposition similarly expressed, for example, asserting that "Catherine is the mother of Samuel", against the propositions that make up the database. A more complex database will include rules (relating to variables) that enable queries to extract information that is only implicit in the facts directly expressed (as constants). For example, it would be possible to extend the above database with rules such as:

be incorporated into databases to define very complex domains; equally, they can be employed to frame very complex queries. In fact, anything expressible in predicate logic can also be expressed in Prolog. Thus, if a database is implemented in Prolog, its combination of facts and rules define a searchable domain space that specifies all the facts that are true for that world.

Hence in simple terms a computational ontology is a formal specification of a conceptualization of a domain. In implementing an *ontology* in a database, a Prolog program lists the specific facts and general rules used to describe a situation. These facts and rules together constitute the *axioms* that define a situation. The Prolog sentences specifying these facts and rules may also be viewed as a *theory*, in the formal logical sense, as they constitute a set of sentences, such that any sentence which follows from the set is also a member of the set. Thus the domain ontology is a formal approach to describing a domain; the ontology specifies a formal language that enables a situation within that general domain to be described completely by the Prolog program / theory. A situation, when described formally in Prolog syntax is a logical structure; that is, it is a theoretically defined interpretation (model) of a situation that specifies whether a query is meaningful, and if meaningful, whether it is true or false.[8]

As Barry Smith (2003b) has argued, information systems ontology has a pragmatic goal: it starts with conceptualizations, and goes from there to the formal description of a corresponding domains of objects; these domains of objects "are nothing more than nodes in or elements of closed world data models devised with specific practical

---

Prolog program (database rules)	**Ordinary Language**
grandmother_of(X, Z) **if** *mother_of(X, Y)* **and** *mother_of(Y, Z).*	X is the grandmother of Z if X is the mother of Y and Y is the mother of Z.
grandfather_of(X, Z) **if** *father_of(X, Y)* **and** *father_of(Y, Z).*	X is the grandfather of Z if X is the father of Y and Y is the father of Z.
sister_of(X, Y) **if** **. . .** *etc.*	X is the sister of Y if . . . etc.

Thus, for example, the above program could be executed by adding a query, such as:
Grandfather_of('Tom', 'Mary'),
[ie., "Is Tom the grandfather of Mary?"], which the program would evaluate as *true*.
For further details on the mathematical interpretation of Prolog, and examples of how Prolog queries of such a database, can be used to explore the structure of Whitehead's notions of concrescence, eternal objects, etc. see Henry (1993).

[8] Sowa (2000) notes that Tarski's semantic method of defining the truth of a sentence in terms of a metalanguage (i.e. model theory) can be applied to Prolog; where the model is equated with the database, matching a sentence with a relationship implied by the database validates the claim of the sentence, and a sentence not matching a relationship found in the database is assumed to be false. This interpretation presupposes that relationships not found in the database of facts are false (Sowa 2000:384). Viewed in terms of model theory, a situation when described formally in Prolog syntax is an interpretation. That is, the Prolog program (the facts and rules which together constitute the axioms that define the situation) is an interpretation. These axioms are the 'theory' or set of sentences that specify the database structure. A query in the form of a proposition is tested against this database. Henry notes that a complex query can be structured as a rule, which functions as a procedural definition (Henry 1993: 28-9). In model theoretical terms the database is an interpretation that *models* any query related to the situation. A sentence (query) may be true relative to one interpretation (axiomatically defined structure) but not another. If a query is true when interpreted in terms of the particular structure, then that particular interpretation is a *model* of that sentence (i.e., the interpretation satisfies the sentence, or supports the truth of that sentence).

purposes in mind" (Smith 2003b). This can lead to the solipsistic world models that Lambert criticized. Information systems in this sense are oriented to the 'world' implicit in a system of beliefs, rather than to a slice of the real world. As Smith points out, IS ontology is then not concerned with reality, so much as with well-behaved reality surrogates. This solipsistic stance is implicit in Gruber's (1995) suggestion that for AI systems what 'exists' is what can be represented in the formal system being used. Thus, from this Classical perspective the focus of ontology shifts from the actualities of the world to beliefs about the world. From this semantic standpoint, ontology refers to the objects, relations, and anything else needed to specify the workings of a model or possible world. This is certainly true for the Classical paradigm, which defined truth in terms of a solipsistic model of the world. However, in Lambert's case, the process philosophical presuppositions of Post-Classical AI suggest a more intimate relationship between mathematical model and the world; namely, that the mathematical model expressing the metaphysical dimension of reality – Lambert's formal version of Whitehead's 'brick-wall' universe - is joined empirically to the real form of the world.

4.2 A model-theoretic interpretation of Whitehead's process of concrescence

That a mathematical structure can have a deep connection to the realm of human experience is fundamental to Whitehead's claim to have overcome the Cartesian bifurcation problem. As described in the previous chapter, Whitehead's metaphysical system can be viewed as a systematic endeavour, building on the evidence of human experience, to explain the relationship between the realm of appearances and the objectifications that form the basis of scientific theories. This explanation grounds Whitehead's belief that his abstractly objectified theory of extensive relations is disclosing a systematic dimension of our Universe that is fundamental to our capacity to formulate rational, empirically based explanations. That is, Whitehead believed that his mathematically objectified system of extensive relations was expressing the real structure of potentialities that are actualized by particular actual entities and nexi of actual entities. Similarly, in Lambert's Post-Classical ontology, this system of extensive connections represents the structure of potential relations between actual process 'events'. In Whiteheadian terms, it is the ground-level society that constrains the more complex societal structures that may emerge within the datum of a concrescing occasion.

To use a Prolog analogy, the Whiteheadian *initial datum* of a concrescing occasion supplies a database of facts (nexi of actualities), rules (customs/propositions) and a query (initial subjective aim derived from God's primordial nature) that guides the subsequent process of concrescence; the process of concrescence is analogous to the computational process of unification, leading to a determinate satisfaction (superject/final form of the concrescing occasion).[9] In model-theoretical terms, the

[9] Henry (1993) advances a detailed formal model using predicate logic and Prolog, of Whitehead's notion of concrescence. It has been the subject of a critique by Lango (1994). I had intended to relate Henry's work to Lambert, but time and space have prevented significant comment. Note however, that Henry makes no attempt to formalize the datum of concrescence (the extensive continuum); his focus is on providing a formal interpretation of the phases of concrescence leading to a satisfaction. His guiding idea is that the subjective aim can be viewed as a Prolog query, that the aim is satisfied via a series of *if-then* conditional clauses that specify the relations of dependency between the hierarchy of transmuted

societal datum is an interpretation that can satisfy a proposition. The extensive continuum is the general form that constrains and enables all emergent forms of societal action. As argued previously, without the extensive continuum functioning as an indicative system, there would be no coherent basis for interpreting the logical subjects of propositions, whether in visual appearances or in symbolically denoted propositions. Lambert's ontology of processes plays a similar foundational role in facilitating machine intelligence.

Whitehead's faith that his metaphysical system had captured essential features of the human situation was based upon the methodology that had guided his process interpretation of mathematical and scientific concepts; he described this methodology as the *method of extensive abstraction*. His commitment to this method was an expression of his rejection of the Cartesian bifurcated cosmos. In model-theoretic terms, his method suggests why the perceived form of the world can be viewed as an interpretation that satisfies Whitehead's theory of extension. This congruence, as will become apparent, helps to explain why a process inspired theory of extension provides a useful ontological foundation for a machine-based approach to generating useful knowledge about the world.

4.21 Model theory as a bridge between Whitehead's mathematical model and the world of appearances

In my account of Whitehead's metaphysical system I suggested that he interpreted the physical prehensive bricks of his process philosophy as a relational mathematical structure. This intellectual objectification of the physical world is the outcome of complex processes of transmutation that involve massive simplification and abstraction from the multiplicity of diverse physical prehensions and conceptual possibilities that enter into an occasion. As Whitehead's metaphysical system makes clear, the "material universe is largely a concept of the imagination which rests on a slender basis of direct sense-presentation" (Whitehead 1917:199). Why should we privilege one such imaginative universe over another? Whitehead's answer to this question is framed in terms of his vision of reality. As discussed in the previous chapter, Whitehead envisages an evolutionary process that underwrites a 'semantics of success' which has selectively promoted efficacious habits, including those that guide and constrain the role of the imagination in vision. In order to justify his own preferred concept of the world, Whitehead developed a *method of extensive abstraction* that was intended to make manifest the principles underlying the reflex habits that shape visual objectifications: habits such as 'the search for convergence to simplicity'.[10] I will not enter into the detail of this method; however Janet Fitzgerald

---

feelings that can enter into a satisfaction. Thus, for example, an intellectual aim will only be satisfied if the necessary structure of comparative feelings – down to the basal physical feelings – is prehended.

[10] Palter, a leading authority on the mathematical aspects of Whitehead's work, argues that the formal method employed by Whitehead in his 'method of extensive abstraction' is replicating procedures employed instinctively in everyday visual perception. Thus, for example, Whitehead represents events using abstractive classes that are "perceived through the senses as extending over one another", using "general concepts of events and of extension between events [that] are clear enough to make the search for convergence to simplicity an 'instinctive procedure of habitual experience'". As Fitzgerald notes:

> This instinctive procedure is lifted to the level of reflective awareness by the method of extensive abstraction without any serious loss of contact with immediate sense experience. Abstractive classes 'guide thought' in directions determined by sense-perception but to destinations beyond all possibility of detection by sense perception. These destinations are the

(1979) has demonstrated that it involves drawing upon mathematical techniques in order to arrive at an algebraic mathematical model of the world of visual appearances.[11] In both its earlier guise as a systematically derived mereology, and its later form as a mereotopology, it is "a modern replacement of the static Newtonian world concept by a multi-relational mathematical concept of the world" that enables Whitehead to avoid "the bifurcation of nature which he so often and so rightly criticized" (Fitzgerald 1979:166; 160). This method, through its specification of his theory of extensive connection, underpins Whitehead's metaphysical system; it also provides an empirical justification and explanation of the abstract formalism at the base of Lambert's model of existence.[12]

According to Janet Fitzgerald the Method of Extensive Abstraction is a mathematical procedure for obtaining the basic entities and relations postulated by metaphysical system, via abstraction from visual perceptions. These abstractions are the instinctive procedures lifted to the level of reflective awareness; the method "with all its logical and mathematical devices" can be viewed as a "a model of the relationships apparent among natural entities" (ibid:164). Fitzgerald argues that Whitehead was not trying to construct geometry from sense perception, but rather, was using a mathematical model to act as a 'theoretical bridge' "between the perceptual and mathematical world" (Fitzgerald 1979:123).[13] In order to fulfil this bridging role Whitehead's mathematical model must possess the property of *isomorphism*, meaning that the structure of the mathematical model and the structure of the field of perception must

---

> abstractions of science, entities which are truly in nature, though they have no meaning in isolation from nature (Fitzgerald 1979:164).

[11] Fitzgerald outlines the method of extensive abstraction, using the example of points, as follows: Whitehead observes that there are no infinitely small sense data; anything we see has some finite extent. The world is made up of events possessing volume and duration. An event at first appears as an undivided whole, but under scrutiny it can be mentally divided into parts contained within the whole. Thus events may be contained within another like a set of Chinese boxes. One event overlaps or encloses other events. This relation of enclosure is used to define a point as a certain class of spatial objects which would be said to 'contain the point'. . . . A point is finally defined as the class of objects which enclose members of an enclosure series which possesses a punctual type of convergence. The route of convergence is called by Whitehead an abstractive set of events. An example of an abstractive set would be the set of spheres concentric to a certain point. Each abstractive class is composed of an infinite series of successively smaller events converging towards, but never reaching, a terminal event. By simplifying and extending certain relations in this process, Whitehead eventually arrives at the ideal simplicity of 'nature at an instant', and goes on to derive physical laws of relativity and universal gravitation. (1979:11-12)

[12] According to Fitzgerald, the key distinction between Whitehead's earlier method of extensive abstraction and his new theory of extension, is that the former conceived extensiveness as purely derived from the notions of extensive part and extensive whole, while the theory of extension replaces the notion of extension with that of extensive connection, focuses on the connections between wholes, and derives the part-whole relationship from the relationship of extensive connection. Fitzgerald comments that this did not mean rejecting his earlier method of extensive abstraction; rather, it signalled that "Whitehead had merely come to see that indispensable to the philosophy of science is an elucidation of concrete fact from which the sciences abstract. This is the task of metaphysics." (ibid:142)

[13] In Janet Fitzgerald's words:

> It was his [Whitehead's] belief that the scientific concepts of space and time are based upon relationships which are disclosed in the perception of nature. Thus, Whitehead wanted to bridge the gulf between the world of sense perception and the world of science by applying the logic of relations to the world of sensed space and sensed time. By so doing, he hoped to define the abstractions of science in terms of spatial and temporal events which are the actual objects of perception. . . . (Fitzgerald 1979:10)

be sufficiently similar to enable a one to one mapping between the entities and relations of both fields.[14] This isomorphism between model and the field of application (Nature) is the basis for judging the 'fit' of the model; the better the fit the greater the value of the mathematical model. To achieve this fit, Whitehead uses the methods of mathematics and logic to clarify the nature of perception, and thereby arrives at an idealized representation of the potential structure of the world that we perceive.[15] The outcome is the theory of extensive connection.

Whitehead's aim was to link the theoretical entities utilized in an abstract mathematical science, such as physics, back to the world of appearances. Typically, in a mathematically oriented discipline a possible realization of a theory would mean a 'set-theoretical entity of the appropriate logical type'.[16] Whitehead's method of extensive abstraction is an attempt to derive appropriate set-theoretical entities systematically beginning from the self-evidence of the domain of appearances. His method involves generalization from the structured form of perceived actuality to arrive at a structure of abstract potentiality conditioning all actualities.[17] The perceived world is thus serving as an interpretation or model of the theory; fragments of the actual world can thus serve as a model for a theory, such that, for example, a real life family, and its customary way of defining genealogical relations, could serve

---

[14] Fitzgerald notes that an "isomorphism displays similarity of structure between the two fields of comparison by setting up a one-to-one correspondence between the entities and relation of both fields". If the perceived entities possessed the same level of exactness as mathematical entities this correspondence could easily be demonstrated, however such is not the case. Hence, the challenge for Whitehead was to relate the inexact world of nature to the exact world of science.

[15] Whitehead achieves this by setting up correspondences between visual models and set-theoretical analogues of those models. The set-theoretical formulations are meant to disclose operations (in the mathematical sense) that underpin visual objectifications. The field of application for Whitehead's method of extensive abstraction is not simply the raw sense data of perception; as my analysis of his metaphysical system makes clear, visual perceptions includes construction arising from instinctive inference. Thus the imagination bestows upon Nature the exactness required to set up the isomorphism between abstract sets on one side and natural entities on the other. Nature has definite entities and relations, and these are potentially exact (as potentialities in the datum of the actual entity). During the process of perception, these potentialities are the object of processes (transmutations). Hence, in Fitzgerald's words:

> Nature is potentially exact and definite. The process of perception [concrescence] actualizes this potentiality. If therefore, we judge the method of extensive abstraction as an isomorphism between the mathematical entities and the natural entities as given in sense-perception, it is a classical algebraic model since its sets up a correspondence between natural perceived entities and geometric entities (Fitzgerald 1979:134).

[16] Fitzgerald provides and example from physics, in which a set-theoretical model is an axiomatized system involving 5 primitive notions defining a set of particles (P), an interval of real numbers corresponding to elapsed time (T), a position function (s) defined on the Cartesian product of P & T, a mass function (m) defined on the set of particles, a force function (f) defined on the Cartesian product of P, T and the set of positive integers. The possible realization of this theory would be the ordered quintuple P = <P, T, s, m, f>, and a model of classical particle mechanics is "any such ordered quintuple". Fitzgerald continues, that "the actual physical model in the physicist's sense is related to this set-theoretical sense of models. We simply let the set of particles be the set of planetary bodies (in the case of solar system)." (ibid:125)

[17] Fitzgerald notes that the standard definition of a model given in most mathematical logic texts is based upon Tarski's approach, according to which "a model is a non-linguistic entity in which a theory is satisfied." (ibid:124) Or in Tarski's words:

> A possible realization in which all valid sentences of a theory **T** are satisfied is called a model of **T**. (Tarski [1953, p.11], cited by Suppes 1960, p.287)

She argues that for Whitehead's theory of extensive connection the perceptual realm serves as the possible realization in which his theory of extensive connection is satisfied.

as a model for the theory (a set of sentences) that provides an axiomatic specification of family relations. Whitehead's approach, in deriving a formal system from commonsense modes of experiencing the world, was foreshadowing later developments in AI and ontology, as articulated by Patrick Hayes (1979; 1985) in his first and second Naïve Physics manifestos.[18] However, as was argued in my outline of Whitehead's metaphysical system, the semantic foundations of our commonsense capabilities involve a complex interaction between bodily indicative systems and habitual capabilities. Lambert's formalization of a brick-wall universe is the first step in capturing this complex foundation of commonsense.

4.3 Lambert's process ontology as a mathematical model of existence

Lambert begins his elucidation of a process metaphysical framework with a discussion of process and its relationship to the problem of change. The problem of change is introduced with reference to Aristotle's metaphysical system and his proposal that reality is based upon a foundation of objects that exist independently of observers. For Aristotle, the dynamic nature of the world is explained in terms of changes in the objects that make up the world. Each object has a permanent 'essence' and a transient 'accidental' aspect. (*EMWC:*134) This metaphysical perspective (an expression of linguistic realism) fosters ambiguity, for it allows the same object to differ, violating Leibniz's law that $(x = y) \Rightarrow (Q(x) \Leftrightarrow Q(y)$ (*EMWC:*134-5). Lambert argues that the alternative is:

> to insist that changed things are different things, thereby safeguarding Leibniz's law. Heraclitus is famed for proposing this idea around 480BC, and drawing the conclusion that one can never cross the same river twice. This provides a very fine grained individuation of things, but it can be extended without sacrificing Leibniz's law, by allowing temporal sums of these things to also qualify as things" (*EMWC:*136).

Lambert concludes that viewing things as temporal sums of temporary distinct things leads to an alternative ontology of 'processes'.[19] He therefore adopts an intuitive notion of process as the axiomatic foundation of his philosophy, and using three primitive notions - fragmentation, time and causal influence - specifies a formal theory that provides an axiomatically formulated framework for describing physical processes.

Lambert's aim is to create the axiomatic theoretical foundations for a mathematical description of the world that is consistent with a process metaphysics; what Whitehead (1906) attempted to do for the physical world, Lambert (1996) attempts to do for commonsense. Grounding commonsense in the real world is fundamental to Lambert's vision. Hence, creating a simplified mathematical model of existence is, for

---

[18] Very different notions of commonsense underlie Whitehead's Method, and the Naïve Physics approach, however both would appear to regard the world as disclosed in commonsense as a model of a formal theory. Hayes, as Barry Smith notes, asserts: "That there is nothing in the model theory of first-order logic which a priori prevents the real world being a model of an axiom system" (Hayes 1979, p.181; 1985, p. 10). Smith reviews some of these issues in a longer, early draft of Smith (2003b) available at his website.

[19] A far more thorough discussion of the historical antecedents and logical justification for adopting a process ontology can be found in D. W. Mertz, *Moderate Realism and its Logic* (1996). Mertz calls his ontology "instance structuralism", and argues that Whitehead's metaphysics of process is an example of such an ontology (1996:81).

Lambert, the first step in creating a computational model of an agent with a capacity for commonsense action. Thus, as with Whitehead (1906) Lambert avoids philosophical discussion.[20] Rather, he puts process philosophy to work as the conceptual frame for formulating the ontological ground of his system.

Thus, bracketing a host of complex philosophical questions, Lambert simply defines processes as the basic elements of his ontology and makes process serve as the basis for his Post-Classical classification of reality, as summed up in his Universe Tenet: "The universe is composed of temporally individuated processes, some which are fragments of others" (*EMWC:*136). Along with fragmentation and temporal individuation, Lambert specifies causation as the third element necessary for his formal definition of process. Informally, he defines process "as *any fragment of the causally evolving universe*" (*EMWC:*136). Since Hume, causation has been a problematic philosophical concept; however, in line with a number of process philosophers Lambert rejects a strong notion of causal determination in favor of a weaker relationship of *causal influence*. Hence Lambert formulates as his Causation Tenet: "Causal influences exist between processes" (*EMWC:*137).

4.31 The Synchronic and Diachronic Dimensions of Process

Lambert's intuitive definition of a process as "a fragment of the causally evolving universe" implies that each process entity has a *synchronic* and *diachronic* dimension; the former dealing with composition and the latter with evolution (*EMWC:*138). Lambert's further treatment of process is based on providing a formal axiomatisation of these two dimensions, using a modified version of Zermelo-Fraenkel Set Theory with the Axiom of Choice (abbreviated ZFC). Using the axioms of ZFC it is possible to envisage abstracted fragments of the universe as sets encompassed within further sets, with all sets qualifying as processes. These sets have a similar status in Lambert's universe of processes as Whitehead's nexal sets of actual entities; a societal nexus can be envisaged as a process, even if it is a macro nexus, made up of many nexūs of nexūs encompassing a span of human history. However, while ZFC holds the promise of providing a framework for individuating the universe it also has some shortcomings, including the *foundation problem*, the *overpopulation problem*, the *class problem* and the *intensionality problem* (*EMWC:*139). The first three refer to problems arising from the definition and membership of sets.[21] The last problem plays a central role in Lambert's re-interpretation of ZFC.

---

[20] This was a strategic decision on Whitehead's part, and did not represent a denial of the importance of metaphysical speculation; his 1906 paper was a mathematical work, not a philosophical treatise. As became apparent in Whitehead's later works, there was a productive tension between Whitehead's formal mathematical, and speculative metaphysical reflections (as reflected in the creative tension between Parts III and IV of *PR* – Whitehead's discursively formulated theory of prehensions and his formally envisaged theory of extension).

[21] Dealing with each in turn, the foundation problem refers to the ZFC's presupposition that every set is an abstraction from the empty set. How should this empty set be related to the real world? Second, the overpopulation problem refers to the over populated universe that arises from an ontology that bestows equality of existence on an ontology of all possible sets that arise from the sets of sets, and sets of sets of sets, and so on, that arise from the application of various ZFC axioms, such as the power set axiom. Hence the ontological framework of ZFC needs to be weakened to relegate many of the possible sets to the ontological status of potential fragments of the universe (or possible universe), rather that actually existing fragments of the universe. The third problem is related to the overpopulation problem and arises through ZFC's introduction of the notion of classes to define useful sets, as opposed to sets that are too abstract and far removed from the empty set (as illustrated by Russell's paradox). ZFC deals

Although ZFC is normally taken to provide an extensional account of reality (as set descriptions are taken to refer its members, with each set being totally defined by its membership), Lambert defines his version of ZFC as providing an intensional account of reality. This is because the way that the composition of the set is defined enters into its identity.[22] Lambert acknowledges that "set theory is extensional if set descriptions refer only to sets", but suggests that these descriptions "become intensional when taken to refer to fragments of the real world." (*EMWC:*142) This enables the one fragment of reality to be classified in several different ways, while avoiding the extensional presupposition that different set theoretical descriptions entail different objects. For example, using the set theoretical representation of processes, it would be possible to describe the tree outside my window in two different ways; firstly, as a fragment of process, composed of sub-processes making up the branches and the trunk {{stem, leaves} trunk}, and secondly, as a fragment of process composed of sub-processes {stem, leaves, trunk}. Rather than regard these two different assemblages of processes as two different trees, Lambert introduces his intensional definition, such that they can be viewed as two different classifications of the same tree (*EMWC:*143).

For the reasons outlined above, Lambert rejects ZFC as a means for understanding *process composition* throughout the universe, but proposes instead to axiomatise the intuitive notion of process by weakening the ZFC axioms. He uses first order logic to develop a framework in which the domain of discourse is restricted to processes, and defines a series of axioms. Firstly, process extensionality is defined by making processes identical on account of process composition. Secondly, the universe is introduced by defining process as any fragment of the causally evolving universe. In this way composition can be understood both as fragmentation (which delivers an extensional world) and as set membership (which delivers an intensional world); hence, fragmentation "exposes what there is" while set membership shows "how we select it" (*EMWC:*143-5).

Lambert then deals with the commonsense notion that processes can be united or separated into further processes. For example a fruit-bearing tree can be viewed as a single process or alternatively, the fruit and the tree can be viewed as separate processes. Lambert formulates these properties in his fusion and fission axioms

---

with these difficulties by introducing adhoc restriction upon its axioms. (*EMWC:*139-142) A succinct summary of Lambert's approach to and modification of ZFC is found in Lambert (2009).

[22] The notions of intension and extension have been referred to in previous chapters and they are important concepts for Post-Classical AI. In logic and mathematics intensional definitions give the meaning of a term by specifying the properties that are necessary and sufficient to belong to a particular set. An extensional definition of a set means listing all the members of that set, while an intensional definition means specifying the properties that make a member eligible for membership of the set. In the real world, different intensional definitions can make important distinctions even though they denote the same object. They may, for example, refer to issues of context, such as those associated with the 'morning star' / 'evening star' distinction. While the distinction between extension and intension has a long history, its modern usage is heavily indebted to Carnap and his successors, who have sought to build the distinction into their formal language. Formal systems in which intensional features can be represented are generally referred to as *intensional logics*. Mathematical and logical sets are intensional if their identity depends not only on their extension (i.e. their members) but also on the way members are presented or defined. For example, from the intensional point of view the set of all non-negative real numbers and the set of all squares of real numbers are not identical, even though they have the same extension.

(specifying the joining and separating of processes) (*EMWC:*145). Formally, these joining and separating processes can be represented by the corresponding mathematical operations.[23] Following from this formalization, "any process which is a component of another process must also be both a component and a fragment of the universe." (*EMWC:*149) This, however, does not prevent nonsensical propositions such as having a universe made up of fragments of nothing. Hence Lambert introduces a foundation axiom, that asserts, "the component of any process are its atomic fragments" (*EMWC:*150).

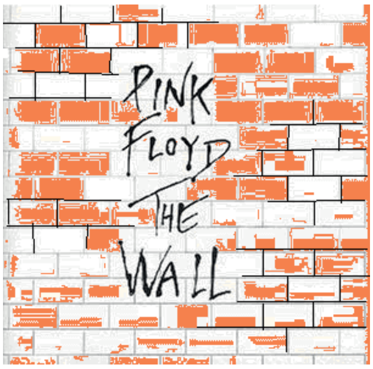


Fig. 4.1 Lambert's vision of a Process Universe: each atomic fragment - just another brick in the wall?

Drawing upon a 'brick wall' universe analogy, Lambert notes that:

> The effect of the Foundation Axiom is to completely collapse the synchronic theory of processes into the Boolean algebra theory it envelops, by interpreting the component relation in terms of atoms for that Boolean algebra. In attempting to visualize the outcome, a useful analogy is to think of it as a theory for describing a brick wall. Under this interpretation, Ω (universe) denotes the brick wall, each of the individual bricks are the components or atoms of the wall, and any collection of whole bricks, including the empty collection, constitutes a fragment of the wall. Each individual brick is both a component and a fragment of the wall, with that individual brick being the only component of that fragment. The fusion and fission operations then permit us to classify various fragment of the wall on the basis of the bricks they contain. (*EMWC:*150)

---

[23] Composition of processes is signified by the symbol $\in$ and the symbol $\equiv$ used to signify process identity. The symbol $\leq$ is used to signify that processes are fragments of the evolving universe. The symbol Ω is used to signify a universal process, called the *universe*. This symbol is also used to denote *everything* and the *world*. The symbol for complement is $\mathcal{C}$, and the complement of everything ($\mathcal{C}\,\Omega$) is nothing, denoted by $\perp$. The variables x and y are fragments of the universe Ω. Fusion of x and y is denoted $x + y$ (the union of x & y), while fission of x and y is denoted by either the inner fission product x with y ($x \bullet y$) or the outer fission product ($x - y$), where inner fission product denotes the greatest process enveloped by both of the given processes x and y (akin to the intersection of x & y), while the outer fission product consists of the remainder when the inner fission product is removed (akin to the complement of the intersection of x & y). A flavour of the resultant system is suggested by the extensionality, fragmentation and universe axioms:

1. $\forall_{x,y}\,[x \equiv y \Leftrightarrow \forall_z\,[z \in x \Leftrightarrow z \in y]]$ - extensionality axiom
2. $x \leq y$ *iff* $\forall_z\,[z \in x \Rightarrow z \in y]$ - fragmentation axiom
3. $\exists_x\,\forall_{,y}\,[y \leq x]$ - universe axiom (*EMWC*:143-148).

Nowak (2001) has published a list of the axioms and definitions used by Lambert in *EMWC*. Lambert (2006a) is a refinement of the arguments and an application of the formal process metaphysics developed in *EMWC*.

This extensionally defined universe is clearly a radical abstraction, and no effort is made to show how it is capable of dealing with the various phenomena of existence that Aristotle attempted to encompass with his ten metaphysical categories.[24] It is analogous to Whitehead's stripped down extensional universe. Like Whitehead's theory of the Extensive Continuum, Lambert's theory describes a universe stripped of all particularities except for formal relations. While this version of a brick-wall universe does not build directly upon Whitehead's theory of extension, it was inspired by that tradition, and in subsequent work Lambert and colleagues have developed a metaphysical ontology that is derived from Whitehead's approach.[25] Lambert's Boolean universe is thus kindred to earlier attempts to formulate in mathematical terms the logical foundations that underpin our notion of a geometrically coherent universe. It is interesting to note that the Polish mathematician and philosopher, Stanislaw Leśniewski, created a similar extensional set theory in 1916. His 'mereology' assumed objects as the universe of discourse and accepted the notion of part as sole primitive. Lesniewski rejected the Platonic view that logical systems are timeless constellations, and adopted a nominalism so radical that he was forced to see logical systems not as abstract sets of propositions, but as dynamic "assemblages of physical marks" that are "growing in time by the addition of new marks called theses, according to well defined rules" (Simons, 1987:18).[26] Edmund Husserl and Alfred North Whitehead proposed similar mereologies.[27]

In addition to the synchronic dimension, Lambert focuses on the diachronic, or temporal dimension of evolution. This temporal dimension was a primary motivation for the development of Whitehead's mereology.[28] Lambert uses time to individuate

---

[24] Aristotle had proposed that everything that can be said about reality (the subject and predicates of all propositions) fall into one or more of ten categories: substance, quantity, quality, relation, action, passion, place, time, position, and state. Of course, in Aristotle's metaphysical system, substance, not process, is the primary attribute of reality.

[25] The system outlined in *EMWC* has since been modified to bring it into closer alignment with Whitehead's vision. Lambert has stated that this vision was his source of inspiration, but that in *EMWC* he had not worked out the mechanics for executing it. However, in subsequent years his original process system has been modified: "So the metaphysics now comprises a process Boolean algebra with time as an ordered subalgebra with temporal measure and with space as a subalgebra supporting connection, orientation and distance constructs" (Lambert 2009b:2). He has commented that Whitehead's 1911 position, "is exactly the position I am at" (*ibid*).

[26] In the 1930s, Nelson Goodman expounded a similar extensional set theory under the title of nominalism. See Simons (1987) for a history of these developments.

[27] For a comparison of their systems, see Simons (1987). Philosophically, Whitehead adopted a position that resonated with Lesniewski's approach to arithmetical operations. In opposition to Russell's later view that arithmetical operations are tautological, Whitehead argued that each operation should be viewed as a process, issuing in an expression, which becomes part of the material process beyond itself. Illustrating his point, he writes: "A prevalent modern doctrine is that the phrase 'twice-three is six' is a tautology. This means that the phrase 'twice-three is six' is a tautology. This means that 'twice-three' says the same thing as 'six'; so that no new truth is arrived at in the sentence. My contention is that the sentence considers a process and its issue." (*MT*:125) Hence, 'twice-three' and 'thrice-two' are two different processes each issuing in compositions with the same numerical character.

[28] "My root idea", Whitehead said in 1911, "is that an object has essential extension in time as well as in space, and that there are time-parts of an object just as there are space-parts. In fact the time and space extensions are the object." (Whitehead to Russell, Sept. 1911, in Eames,1989:112) By 1920 Whitehead had developed this root idea into a philosophical thesis that the "world we know is a continuous stream of occurrence which we can discriminate into finite events" and these events form

his world of processes, such that "each moment of an object defines a different process, while the totality of its temporal pieces defines a different process again" (*EMWC:* 152). The result is an extensional world where things are the totality of their temporal parts, and objects are annotation of change, not *vice versa*. The need, therefore, is to extend the formal metaphysics of process to incorporate the compositional and chronological structure of time. This is accomplished by embedding the ontological structure of time within the Boolean algebra of processes. In commonsense realism time is viewed as either a point or an interval, as when we speak of a moment *in* time, or an interval *of* time. In Lambert's axiomatization, defining a moment of time as equivalent to an atomic process accommodates these notions (*EMWC:*152-154).

Drawing on the brick wall analogy that was used to illustrate synchronic processes, the *diachronic process* can be visualized as arising from the practice of cementing new columns of bricks onto the end of the wall, "with each moment of time corresponding to a column of bricks", and with the Boolean algebra now operating over columns of bricks rather than individual bricks (*EMWC:*155). In Whiteheadian terms, each new column of bricks is a set of contemporaneous superjects added to the extensive continuum of perished actual entities.[29] Lambert notes that the concept of a brick-wall universe demonstrates that the elementary mathematics proposed "provides a very natural extensional account of reality and exposes unambiguously the sort of inter-relationships I take there to exist between time and the process ontology" (*EMWC:*155). Moments of time (temporal process), from this perspective, are the most primitive temporal intervals from which all other intervals are derived.[30]

Lambert next explains the *chronological structure of time*, capturing the intuitive notion of the 'flow of time' in terms of points and intervals, using an ordering relation.[31] Finally, Lambert introduces the notion of *causation* into his model of temporal processes via axioms that specify the nature of atomic causal influence. Causation as applied in Lambert's model is a loose concept akin to Whitehead's notion of the diverse ways in which a chain of influences can descend down through the members of a society. The result is that "every atom belongs to at least one causal influence chain of temporally adjacent atoms" and is also extended to include non-atomic processes (*EMWC:*163). Thus, at a formal level, Lambert's vision of a process universe is very similar to the Whitehead's extensively connected universe; in both cases the basic 'bricks' of the universe are fragments of process that have the capacity to causally influence subsequent events.[32]

"by their overlappings and containings of each other and separations create a spatio-temporal structure." (*CN* [Cambridge,1920]:172)

[29] There are complex issues related to temporal relativity that Whitehead addressed which need not concern us here; my intention is not to assert a point by point correspondence between Whitehead's metaphysical system and Lambert's formalization of a process vision of commonsense, rather it is to show how Whitehead's conceptualization can be used as a guide to Lambert's project.

[30] Note that all processes involve a diachronic and synchronic dimension, and that time should not be viewed as distinct from space and matter as in the Aristotelian inspired metaphysics of substance. Lambert's theory involves the same problems and a similar solution to that proposed by Whitehead in his theory of extensive abstraction, as elaborated in PNK.

[31] This ordering relation defines time as an unbounded, discrete, linear order of granular moments, embedded in the compositional structure of processes (*EMWC*:156-162)

[32] The process 'bricks' of Whitehead's Universe are 'simple physical feelings'; these physical feelings are the presupposed elementary foundation of Whitehead's cosmology. Whitehead writes that "all our

4.4 Conclusion

In this chapter I have tried to demonstrate that Lambert's brick-wall universe of process fragments can be viewed as analogous to Whitehead's brick-wall universe of actual entities. I have suggested that Whitehead's philosophy provides a rational foundation for linking this process structure to the realm of appearances, and suggested that Lambert's mathematical objectification of the extensive continuum will play an analogous role in Lambert's Post-Classical model of an artificial mind as the extensive continuum in Whitehead's explanation of the human mind's capabilities.

It is important to keep in mind that Lambert's critique of Classical AI was directed toward questioning the metaphysical, epistemological and heuristic adequacy of its notion of representation; representation was identified as the main reason for the failure of Classical AI, and hence developing a different approach to representation is at the core of his proposal for a Post-Classical paradigm. Lambert stresses that the reason he has introduced this formal metaphysical notion of processes "is to understand representation" and provide a source for a system of semantics that avoids the 'manifest image' presuppositions of linguistic realism (*EMWC:*164). Just as Whitehead builds more complex functions upon the basic structure of relational processes represented in the theory of the extensive continuum, so also Lambert builds his superstructure upon his brick-wall Universe. For Lambert, the world is the totality of processes, not the totality of facts or things. Processes are more primitive than either facts or things. They can be described in terms of fragmentation, temporal precedence, and causal influence, and on the basis of this extensional ontology, a non-linguistic notion of intensionality can be developed. This model, Lambert argues, is more appropriate for the task of engineering embedded machines, as the world "does not directly present itself to the robot in terms of pre-packaged linguistic objects or facts", but rather, "through a sequence of causal influences linking processes in the world to processes in the robot" (*EMWC:*165).

Thus, Lambert's process foundation is fulfilling a function analogous to Whitehead's theory of the extensive continuum. It provides a model of a structured domain of temporal and spatial causal connections that is reduced to a bare extensional realm, valid for the world within and the world beyond the robotic agent. Lambert's mathematical model of the universe thus reduces events to patterns of transformation defined purely logically in terms of extension. Such patterns can be intensionally classified in various ways, but they have the same extensional reference.[33] The tree we view is the same extensionally independent of the various intensional ways that I

---

physical relationships are made up of such simple physical feelings, as their **atomic bricks**" (*PR*:238; for further elaboration of this brick metaphor, see *PR*:308-9). For Whitehead, these prehensive feelings are re-enacted in each new subject whereby the cause is transferred into effect; this causal transference embodies the reproductive character of nature. Lambert, like Whitehead, regards the basic temporal-spatial regions of his physical universe as the 'conformal' bricks of his monadic universe; they are his elementary monads stripped to the bare properties which define the logical structure of the Universe (this is the logical structure specified in Whitehead's mereo-topological theory of extension).

[33] Lambert's axiomatization of processes sought to uncover "what there is extensionally independently of how it is intensionally classified" (*EMWC:*166).

classify it. To get to the extensional classification of process it is necessary to begin with an intensional classification and then strip away the qualities employed in the individuation. As Lambert notes:

> Processes are individuated matter stripped of intensional form. All that can then be said of this matter is what its structure is like, and to the author's initial surprise, that resembles a causal, temporally indexed, complete, atomic Boolean algebra. Reality as matter, is understood as a causally evolving slab of fragments in which fragments are assembled into other fragments, fragments are separated into other fragments, and fragments are individuated by the irreducible fragments they can be separated into (*EMWC:*166).

Having formalized existence in terms of his brick-wall universe, Lambert's next step is to formalize a notion of representation. How is a robot to represent significant enduring objects in the flux of process? I will consider Lambert's response to this question in the next chapter.

Chapter 5

A Post-Classical theory of Representation

Chapter 5 – Section Headings

5.0 Introduction – Lambert's theory of representation

> All inferences from experience therefore, are effects of custom, not of reasoning. . . . Custom, then, is the great guide of human life. It is that principle alone which renders our experience useful to us, and makes us expect, for the future, a similar train of events with those which have appeared in the past. Without the influence of custom, we should be entirely ignorant of every matter of fact beyond what is immediately present to the memory and senses. We should never know how to adjust means to ends, or the employ our natural powers in the production of any effect. There would be an end at once of all action, as well as of the chief part of speculation.
>
> \- David Hume.[1]

When David Hume spoke of the rule of custom, his words were motivated by epistemological concerns; when Alfred North Whitehead argued that both the laws of Nature and our capacity to perceive Nature were an expression of habit or custom his concerns were metaphysical; when Dale Lambert took up this theme his concerns were those of a software engineer. The challenge was to translate metaphors and analogies into precise mechanical structures. In the previous chapter I outlined how Lambert transformed a theory of processes analogous to Whitehead's brick-wall universe into a formal theory of existence that was expressed in set theoretical terms. In this chapter my aim is to show how Lambert proposes to use his formal theory of processes to develop a Post-Classical theory of representation. His aim is to develop a machine implementable notion of representation that does not rely upon the Manifest Image presuppositions of linguistic realism and related notions underpinning representation in Classical AI (*EMWC:*167-8).

In Whiteheadian terms, Lambert's aims are firstly to explain the presence of enduring objects (Whitehead's societies) within the flux of process; he does this by regarding 'societies' as relationally connected fragments of process that self-replicate their

[1] Cited Lambert (1996:203), from Hume (1748) *An Enquiry Concerning Human Understanding*, Sect. V. "Sceptical Solution of these Doubts", Part 1, §36.

characteristic form. Second, given that one of these social objects is a robotic agent, Lambert seeks to specify how it translates the notion of custom into machine implementable routines that are in some way isomorphic with the routines that define the societies found in Nature. That is, given that social objects are made up of a changing stream of actual entities, how is an agent able to detect and represent these patterns of change as a unified concept (eternal object)? For example, the desk lamp is a society with its characteristic enduring form; if it is switched on it can illuminate my desk, and it can also be switched off. How can a machine detect and represent the functional significance of that object without appealing to apriori linguistic concepts and associated semantic systems? Lambert shows how this can be accomplished using two classes of physical discriminating devices, and two corresponding classes of symbols, that function in an analogous sense to eternal objects. He demonstrates that the role played by eternal objects in the perspectival objectification of a datum can performed by intensional logical structures. In this manner, Lambert responds to the challenge of providing machines with a functional equivalent of higher level ‘conceptual prehension’ (prehension of propositions) without resorting to the Manifest Image presuppositions of linguistic realism.

The key to Lambert’s Post-Classical theory of representation is the relationship between processes, permanence and representation. The linguistic realism of Classical AI regards permanence and corresponding uniformities in the world as given, and builds these objects into its metaphysical system; process metaphysics places its emphasis on difference and change and seeks to explain the emergence of permanence. In accordance with the process philosophy tradition, Lambert views patterns within the flux of processes as evidence of causal influence passing from one moment to the next; thus, rather than view permanence as a primitive attribute of reality, he argues that permanence arises in the world, “moment to moment, by the causal influences present within it” and suggests that “our ability to negotiate the dynamic world rests with our attention to these repeating patterns of influence” (*EMWC:*168). Hence, the “permanence that we seek amid the change is to be found in the *patterns of influence* that arise as causal influence spreads throughout a collection of processes” (*EMWC:*168).

5.1 Lambert’s self-replicating patterns of influence as enduring objects

Whitehead identified the permanent objects of his universe with self-replicating patterns of societal activity. The flux of events was Whitehead’s primary reality, but some patterns of activity generate uniformities within this flux, and these characteristic forms of uniformity are the basis of enduring objects or ‘societies’. In a similar manner, repeated patterns of influence define the uniformities of Lambert’s universe, and are central to the further development of his system. Not all causal influences need produce uniformities; however, some repeated patterns of change are evident in the guise of enduring objects, and Lambert attributes these repeated patterns to the spread of causal influences between processes. Lambert concludes that for Post-Classical AI these self-repeating patterns of influence are the *central intellectual invention*, rivalling “Aristotelian substance, Cartesian intuition and

Wittgensteinian fact" in providing the conceptual foundation and impetus for a new way of understanding our situation in the world (*EMWC*:168).[2]

Thus, the uniformities in the Post-Classical Universe are not those specified by a language. Identifying facts does not depend upon inferring the existence of some object from the presence of various properties ascribed to linguistically defined things. In opposition to linguistic realism of Classical AI, Post-Classical AI seeks uniformities that are defined solely in terms of their causal temporal structure. Matching patterns of influence are identified by their isomorphic causal, temporal structure.

As an example of such a uniformity Lambert proposes the pattern manifested in the operation of a light appliance, "in which the repeated pattern of influence that is the switch repeatedly influence the repeated pattern of influence that is the light appliance" (*EMWC:*169). In this example, the causal influence manifests itself in the correlation between changes in the position of a switch and changes in the state of the light. Lambert formulates this notion as an <u>influence tenet</u>*;* that "some patterns of influence consist of some patterns of influence repeatedly influencing other patterns of influence" (*EMWC:*169).[3] Lambert generalizes from this notion of influence to develop a notion of 'control' and 'attention' that applies to machines as well as to humans, such that a process is said to "control a uniformity" if it is repeatedly "influencing instances of that uniformity" (*EMWC:*170). For example, from this process perspective a switch can be viewed as 'controlling' an appliance, while an appliance is 'attending' to a switch. From a Whiteheadian view, these notions reflect the two ends of a chain of prehensive influence; 'controlling' is manifest where a transitional process habitually leads to a given outcome, and 'attending' is manifest where a process of concrescence habitually responds to a given datum in a similar fashion. Likewise, a societal structure exercises a control function over another structure when the habitual behavior of the control system constrains the habitual attention of the other system. In either case this influence manifests itself in uniform patterns of change; it is these uniformities that are the determinate objects of a process world.

Lambert ties the existence of representation directly to the existence of uniformities as repeated patterns of influence. Every occurrence of a repeated pattern of influence is in some sense representative of every other occurrence of that repeated pattern of influence, "as each [occurrence] duplicates the requisite causal-temporal structure noted of the others" (*EMWC:*170). That is, it manifests the same structure of relational potential ('eternal object' in Whitehead's terms). Representation in this objective sense means that a process repeats or duplicates the pattern of influence of another process (subjectivity and the question of intentionality will be considered in the next section when representation is related to belief). Thus, different moments in

[2] In later chapters I will argue that Lambert's self-repeating patterns can be understood as Whiteheadian social customs, and that these Whiteheadian customs are the conceptual foundation of Post-Classical, *Social AI*.

[3] Since writing *EMCW*, in an effort to avoid metaphysical baggage associated with the notion of causality, Lambert has dropped all references to causal influence and simply refers to patterns of change that 'change' other patterns of change. Hence, he suggests that his <u>influence tenet</u> can be rephrased as "some patterns of change consist of some patterns of change repeatedly changing other patterns of change" (Lambert 2009b:2).

the life of a tree could be regarded as representing the same tree, or alternatively, because it also duplicates the causal-temporal structure of other trees, could also be regarded as representing other trees in the forest. This allows one process to be representative of another without a presumption of direct influence between them. This notion of representation points toward the endless subtleties of meaning that the notion of 'eternal object' conceals in Whitehead's work; from a philosophical standpoint Lambert's work can be viewed as a suggestive reflection on how to interpret some of the functional roles associated with Whitehead's eternal objects.[4]

5.11 Model theoretical representation of patterns of influence as partial worlds

For Whitehead and Lambert, representation involves paying attention to only some of the fragments of process that make up the world. For both authors, representation understood in its broadest sense may entail the full replication of a previous pattern of influence disclosed in the datum of the brick-wall universe such that, for example, one moment in the temporal duration of a tree is viewed as a representation of the next moment. However, in the more restricted sense of an agent perspective (in the sense of a Whiteheadian subject or a machine society engineered to function as a pseudo high-level subject) representation is always a partial realization of the pattern of influence that it represents, in that it abstracts a pattern from its wider context. Moreover, the society (or process 'fragment') being objectified in a representation may be extremely complex; some societal fragments may contain other fragments, and one such fragment may comply with many patterns of influence. Thus Lambert argues that the same fragment of process may represent different patterns of influence at different levels of abstraction, as for example, when the same black marks on a page can represent a style of font, a form of grammatical structure, or a particular story.[5] Thus, representation entails an abstract individuation process, with each representation constituted to reflect the world at a particular level of abstraction. For Lambert, each level of abstraction is a 'partial world' that describes how the world actually is while omitting some fragments of process in order to attend to the fragments relevant to the description of a particular pattern of influence.

Lambert uses a model theoretic structure as a mathematical device for describing patterns of influence in these partial worlds (*EMWC:*170-3). In accordance with the dictum that representation is mediated by uniformities in repeated patterns of influence, Lambert's model theoretic account of representation specifies that representation occurs when there is a <u>repeated pattern of influence</u>, such that the pattern of influence in the world process is isomorphic to a pattern repeated in a representation process. Each partial world representation is akin to a Whiteheadian proposition (theory) that specifies a simplified version of the characteristic form or eternal object of a Whiteheadian society.[6] The proposition is mediating the

[4] For a computational interpretation of Whitehead's eternal objects see Henry's pioneering contribution (Henry 1993).

[5] That is, the same repeated patterns of influence, such as the patterns of letters on a computer screen, can be individuated in different ways depending upon the functional interests of the agent. Lambert notes that multiple patterns of influence exist within the one process at various levels of abstraction, "with the level of abstraction dictated by the individuation chosen for the process fragments" (*EMWC:*172).

[6] For Whitehead, self replicating patterns of influence that retain a characteristic form are societies, and the enduring objects of our world are all understood as societies. Societies of societies can interact with each other in various ways to constitute more complex societies, with self replicating patterns of

perspectival objectification of a society in the concrescing occasion's datum by simplifying the real world in a manner that retains an isomorphic relation to key identifying characteristics of the enduring social pattern. In an informal sense, Lambert's model-theoretic approach to specifying this characteristic societal form involves defining a structure (i.e. an underlying set together with an interpretation of its functions, constants and relations), thereby generating a set of valid sentences that describe that partial world.[7]

---

influence forming more complex eternal objects. Higher order societies, such as mammals, negotiate these complex environments by drawing upon propositions that specify habitual methods for simplifying this complexity. Propositions (theories) function as 'lures' or programs that are interpreted and satisfied by a datum, yielding a valid perspectival objectification of the situation (e.g., we see a tree).

[7] A model theoretical representation specifies truth in a particular interpretation (see my brief introduction to model theory in Ch.4). Lambert accomplishes this for his partial worlds by assigning values to the variables in his set theoretical structure, thereby generating a set of true formulas that describe a pattern of causal influences. In his formal notation, each pattern of influence is presented as a pair of the form $<s, \Gamma>$, where s is a variable assignment into a partial world x_χ and Γ is a set of formula of type $T = \{\leq, \angle, \mapsto\}$.

In interpreting these symbols recall that the formal relations between the elements of any set of Lambert's processes are those of fragmentation, temporal precedence and causal influence (these relations are informally introduced in my description of Lambert's 'brick wall' universe; see previous chapter, also *EMWC*:143, 158 & 163). Relations of fragmentation, temporal precedence and causal influence are the basis of the model theoretic structure $\mathfrak{R}$ used to model the real world, where:

$\mathfrak{R} = (U, \mathfrak{I})$ of type $\mathbf{T} = \{\leq, \angle, \mapsto\}$, and

$U = \{x | x \leq_{\mathfrak{R}} \Omega\}$, then

$\mathfrak{I}(\leq) = \{<u,v> \mid u \leq \Omega \ \& \ v \leq \Omega \ \& \ u \leq_{\mathfrak{R}} v\}$,

$\mathfrak{I}(\angle) = \{<u,v> \mid u \leq \Omega \ \& \ v \leq \Omega \ \& \ u \angle_{\mathfrak{R}} v\}$ and

$\mathfrak{I}(\mapsto) = \{<u,v> \mid u \leq \Omega \ \& \ v \leq \Omega \ \& \ u \mapsto_{\mathfrak{R}} v\}$.

Informally, these symbols can be interpreted as follows:

U = underlying set; $\mathfrak{I}$ = interpretation of the function, constants and relations; $\leq$ = fragmentation; $\angle$ = temporal precedence; $\mapsto$ = causal influence; Ω = universe as process (everything or the world) and u and v are fragments of the real world (*EMWC*: 170-1).

A formal theory, in model theoretic sense, is required to define incomplete patterns of influence, and therefore Lambert proposes a theory that states (informally) that: "A pattern of influence occurs in the real world if the real world substructure for that partial world is a model for the set of formulae with respect to the variable assignment" (*EMWC*:174).

The substructure x_χ of the real world structure $\mathfrak{R}$ is obtained by not representing some of the processes constitutive of the real world. Lambert notes that such partial worlds can be described using an abstract discrimination function χ: $\{x|x \leq\Omega\} \rightarrow \{x|x \leq\Omega\}$, where there is an association of processes to processes, with each fragment x in the universe Ω either included or excluded in the partial world of x under χ. This partial world is represented by the set x_χ, defined $_{df}$ formally:

$x_\chi =_{df} \{ u | u \leq x \ \& \ \chi(u) \equiv u\}$.

Thus, with Lambert's partial worlds formulation of selective individuation, each self-replicating pattern of influence in the real world characterized by an abstract pattern $\mathfrak{U}$ is expressible as a partial world x_χ that defines the substructure $\mathfrak{R} \mid x_\chi = (x_\chi , \mathfrak{I}_{x,\chi})$ of type

$\mathbf{T} = \{\leq, \angle, \mapsto\}$, where

$\mathfrak{I}_{x,\chi} (\leq) = \{<u,v> \mid u \in x_\chi \ \& \ v \in x_\chi \ \& \ u \leq_{\mathfrak{R}} v\}$,

$\mathfrak{I}_{x,\chi} (\angle) = \{<u,v> \mid u \in x_\chi \ \& \ v \in x_\chi \ \& \ u \angle_{\mathfrak{R}} v\}$ and

$\mathfrak{I}_{x,\chi} (\mapsto) = \{<u,v> \mid u \in x_\chi \ \& \ v \in x_\chi \ \& \ u \mapsto_{\mathfrak{R}} v\}$.

The novelty in Lambert's approach lies in the role played by the world (coupled to appropriate discriminators) in empirically individuating a qualitative interpretation of an abstract mathematical structure (specifying that processes are fragments of that universe with various possible relations of spatial connection, temporal precedence and causal influence), such that the real world processes are isomorphic to the processes represented in the *partial world*.[8] As when Whitehead's prehended datum provides the interpretation that satisfies a perspectival concept of the world, so also the real world is providing the interpretation for a structure that *models* or implicitly defines a set of true sentences that describe Lambert's world. In Lambert's words:

> Once the processes of interest have been selected, *the world* determines a pattern of influence, for the way the world is determines, independently of ourselves, which of the selected processes are fragments of others, which of the selected processes temporally precede others, and which of the selected process influence others. [. . .] A pattern of influence occurs in the real world if the real world substructure for that partial world is a model for the set of formulae with respect to the variable assignment (*EMWC:*171; 174).

Representation as duplication is one type of representation, but the isomorphism[9] between processes need not be causal-temporal in form. What is necessary is that the representing process has an abstractly individuated form that is isomorphic to some individuation of the patterns of influence in the represented process, as in the case of a movie tape representing the dynamic structure of the events represented (*EMWC:*175). The role of isomorphism is essentially the same as that which

---

As noted, each pattern of influence is presented as a pair of the form <s, Γ>, where s is a variable assignment into a partial world x_χ and Γ is a set of formula of type $T = \{\leq, \angle, \mapsto\}$. Each pattern of influence <s, Γ> occurs in the partial world only if its real world substructure $\mathfrak{R} \restriction x_\chi$ allows s to satisfy each well formed formulae of Γ. As representation is mediated by uniformities in repeated patterns of influence, Lambert's model theoretic account of representation specifies that representation occurs when there is a repeated pattern of influence, such that the pattern of influence in the world process x_χ occurs again in a representation process y_χ. Informally, this means that if the set of formula Γ are satisfied, then the set of real world processes x_χ is mirrored in an isomorphic set of representational processes y_χ. For more, see references in Ch.4, fn.4; for more on Lambert's use of this apparatus see *EMWC*:170ff.; also Lambert (2009a:9-11). See section 5.2 of this chapter for an informal presentation of examples of the representational structures used to mediate this model theoretic approach.

[8] Barwise and Perry introduced partial worlds (situations) in their situation semantics in order to model incomplete information, as described in their classic text, *Situations and Attitudes* (Barwise and Perry 1983). It is worth noting that they explicitly link their semantic theory to the 'commonsense' tradition associated with Thomas Reid. Reid had argued that our commonsense understanding of reality is rooted in our natural constitution. Rejecting the external world scepticism of Descartes and Hume, he argued that our bodily organs of perception are attuned to the natural world. Reid's commonsense realism influenced the American pragmatist tradition, and was revived in the psychology of J. J. Gibson and in the philosophy of Hilary Putnam. According to Barwise and Perry, Gibson and Putnam's work is central to the new doctrine of Ecological Realism, which holds the view meaning is "located in the interaction of living things and their environment"; and hence "that there is much more meaning and information in the world and less in the head than the traditional view of meaning assumed" (Barwise and Perry 1983:x). They regard their own work as a contribution toward clarifying the theory of meaning implicit in this new realism. Lambert's work could be viewed in the same vein. See Lambert (1999c) and (2009a) for further references to the work of Barwise and Perry.

[9] Informally, an isomorphism between two processes could be defined as a function that sets up a one to one correspondence between the patterns of influence (causal-temporal structures) that define each process. If two mathematical structures are isomorphic then they are essentially the same; the only difference is the names given to the underlying sets. If two such structures are isomorphic then they share all model-theoretic properties.

Fitzgerald details in her analysis of the model theoretic form of Whitehead's method of extensive abstraction (see chapter 4.2), with patterns of influence in the world represented by specific structures (theories), and these model theoretic structures in turn providing an interpretive context for subsequent patterns of influence.

Lambert makes the point that in a weak sense, representation is also about symbolism. However the symbolism takes place at a rudimentary level prior to facts and objects, where what is symbolized is a pattern of influence manifested in a sequence of events. Lambert notes that in contrast to Classical representation:

> A representation represents a pattern of influence by copying its structure in some form, and probably more often than not, its structure is one unfamiliar to our exotic manifest image symbolism. If viewed as symbols, Post-Classical representations are by comparison, extremely rudimentary (*EMWC:*177-8).

5.12 How does a Post-Classical agent selectively objectify its world

Lambert's model theoretic account of representation, as so far presented, has yet to say anything about the connection between the representation and the represented. In order to designate how a process purposefully represents its environment Lambert introduces his intelligent machine, ρbot (using the Greek letter 'Rho', hence 'rhobot'). In line with the engineering goal and guiding philosophy of his work, ρbot is envisaged as an embedded agent, with both agent and their environment conceived of as process fragments in the wider universe of processes. Lambert identifies several questions concerning the relationship between the representation and the represented that ρbot must be engineered to resolve. Two issues which I shall address in this section are:

- First, if process involves selecting a partial world, how is the partial world chosen and physically realised?
- Second, if representation involves a process practicing selective attention to processes of fragmentation, temporal precedence and influence in the world, how are these selections made and physically realised? (*EMWC*:179)

Lambert's solution to these questions is based on the resemblance between partial worlds and our experience of the world. With respect to the first question, he recognizes that an agent could represent its environment in many different ways, but models his agent on a process interpretation of the human response. He notes that perception entails selectively abstracting and simplifying the mass of processes that make up our situation using the qualities present in various fragments to individuate the objects that we are disposed to recognize. This involves two levels of processing; firstly, something analogous to Whitehead's transitional process, that involves primitive sense data, acting as an efficient cause, that disclose simple eternal objects in the form of qualitative 'sensa'; secondly, something analogous to Whitehead's process of concrescence, whereby a subjective aim and the associated datum of simple sensa, are selectively transmuted so as to pick out the more complex social objects present in a datum. Just as Whitehead's social agents employ habits in the form of pure instinct, to guide their attentiveness to sensa, Lambert's agents employ discriminators to pick out primitive qualities. Likewise, in the same way that Whitehead's higher order social agents draw upon the guiding propositions (theories) embedded in reflex habits to guide the selective unification of the primitive sensa so as to identify societal objects, so Lambert's agent draws upon functionally oriented templates to identify more complex patterns of causal influence.

Thus ρbot's first step in selecting a partial world is based on something analogous to the receptors in the eye, acting as fixed, predefined processes that are influenced by certain other processes (the presence of light). Lambert labels these elements as *discriminators*; they provide the basal level of qualitative individuation that embeds ρbot in the real world. In Whiteheadian terms, discriminators are selective devices that arise through the evolutionary processes that tune social organisms to detect simple qualitative changes in their environment. Without higher-level processing this raw data discloses only the bare form of sensa in the extensive continuum; however, latent within this datum are structural isomorphisms between successive nexi of qualitative data that reflect more complex societal patterns of change in the environment. It is these intensional patterns of qualitative data that are basis of ρbot's partial worlds. They are intensional patterns because ρbot can specify the same fragment of process by combining raw qualitative data in different ways. Thus, Lambert defines ρbot's partial worlds as a selective individuation which arises from qualitative individuation (*EMWC:*179).

In ρbot's world classification occurs whenever **ρ classifies x as $\mathbf{x_q}$**, where ρ is the *classifier process* ρbot, **x** is the process being classified and $\mathbf{x_q}$ is the qualitative classification. Lambert's qualitative classification signifies something analogous to the qualitative intensional classification carried out by Whitehead's subjects, in that both involve classifying devices that have causally distinct, fixed, pre-defined processes that are causally influenced by specific primitive processes in the world (*EMWC:*180-2). That is, as in the case of Whiteheadian agents, the initial raw sense data mediating a subject's response to its environment are qualitative forms of feeling dominated by pure instinct (i.e., non-living physical societies).[10] It is this raw data,

[10] In Whitehead's terms this relationship is the basic form of all prehensive events; namely, a subject feeling (via a qualitative subjective form) an actual entity in its datum. In Henry's (1993) analysis, this relationship forms the centrepiece of his Prolog interpretation of Whitehead's metaphysics. In the highly abbreviated outline of this interpretation summarized in Henry and Valenza (2008), the basal form of all prehensions maps directly onto the three components of Prolog predicate:

Subjective-form(Subject, Datum).

This should be read as meaning that 'the subject *Subject* prehends the datum *Datum* with subjective form *Subjective_form*'. In order to represent the dependency relationship that exists in the concrescent process between the basal simple physical feelings and higher level transmuted feelings, Henry suggests that species of prehension and corresponding phases of concrescence can be captured by a list of facts and rules in a Prolog program. The phases of concrescence then take the form of a rule which suggest how subsequent 'transmuted' feelings are derived from antecedent feelings, beginning with basal physical feelings. This process of concrescence can be represented as:

*Subjective-form_1(Subject, Datum_1)***if**
*Subjective-form_2(Subject, Datum_2)***and**
*Subjective-form_3(Subject, Datum_3)***and**
. . .
Subjective-form_n(Subject, Datum_n)

Henry and Valenza emphasize that in the actual execution of the corresponding Prolog program, there is a dynamic representation of concrescence, which, although it is temporal and skates over many significant issues, "nonetheless cultivates deeper insight into Whitehead's overabundant complexities" by formally representing "something of the interior structure of process metaphysics" (Henry & Valenza 2008:199). In terms of suggesting formal structures that capture aspects of Whitehead's system, Henry's (1993) analysis attempts to capture more of the technical detail of Whitehead's notion of the supplementary phases of concrescence than I deal with in my analysis of Lambert. Yet, from a such as light, heat, or sound, which subsequent higher-order transmutative processes build upon. However, the initial, simple physical purposes connect the subject conformally with its environment; only in subsequent phases of concrescence is this initial objectified datum selectively transformed to disclose more complex societies latent in the initial datum.

For Whitehead, the key factor facilitating the realization of a partial world is the customs or habits that guide the transmutation process. For Lambert this role is facilitated by *templates*. Thus, in addition to discriminators (that detect primitive qualities) Lambert proposes a second class of classifiers – called templates - that detect more complex, functionally individuated relational structures latent in the primitive patterns of influence individuated by discriminators. A light, for example, can be a primitive quality if picked out by a discriminator (a sensor such as a light sensitive diode) but in the context of a pattern of other qualities indicated by a template, it can assume a functional role (for example, the isomorphism between the pattern picked up by discriminators and the pattern of potential relations embedded in the template may indicate that the role of the light picked up by a discriminator is to indicate that an appliance is turned on). While discriminators are simple, fixed, physical filtering devices that persist over time and generate an immediate response to selected stimuli, the identification of a functional role is a more complex process that involves incorporating time in the selection process. Templates functionally differentiate patterns on the basis of the structured transformations of qualities over time. For example, the role of a switch can only be detected in the context of a record of the interdependent qualities of the relationship between switch and its appliance over time.

Thus, the role of ρbot's templates can be likened to the function of higher level habits in Whitehead's organisms. In Whitehead's metaphysical system, propositions (theories) lie at the core of all higher level forms of habit, and are the basis of an organism's capacity to selectively individuate complex patterns of events that are both spatially and temporally distributed. Vision, for example, individuates objects of significance, building upon sensa individuated by simple physical societies and drawing upon habitually selected propositions to transmute and objectify that raw data into the forms defined by complex eternal objects. Templates, in Lambert's system, play a similar functional role. Of course, in ρbot this functional role is mediated by the purely instinctive processes that define non-living mechanical agents; thus in Whiteheadian terms ρbot is a non-living society, but it possesses a complex, highly differentiated structure that enables it to simulate, and to some extent functionally replicate, the higher level processes of Whitehead's living societies.

Hence, in functional terms ρbot's templates are fulfilling a similar role to Whitehead's propositions when they are embedded in the relevant societal datum of a concrescing entity (i.e., reflex habits that embody the memories of the organism). As with Whitehead's social agents, Lambert's Post-Classical agent is a society of causally related processes, selectively attuned to and responsive to qualities disclosed

philosophical standpoint, Lambert's work has some advantages in that while *EMWC* does not draw direct analogies with Whitehead, Lambert's project can (from a practical standpoint) be regarded as a far more complete formalization of Whitehead's process universe as it involves embedding an agent in the real world, and ultimately implementing that vision in an operational system.

by the physical processes that constitute its environment. The agent is not exposed to an unlimited range of possible worlds; rather, they are oriented to contexts defined by the way in which discriminators are organized to support the particular functional templates. In Lambert's words, the template may be a collection of discriminators organized "so as to be able to register, maintain and individuate representation of abstract process structure", with abstract process structure defined by the functional role of processes in a temporal context "relative to a wider system of qualities in which their functional roles are specified"(*EMWC:*183-4). The role of the template (that of individuating processes in a functional context) can be performed by a computer's memory.

In Whitehead's metaphysical scheme, a personal agent, such as a human, can be viewed as a sub-society within the larger societal context of an environment. Thus, at any moment in the subjective life of such a personal society, its initial datum contains a sub-society that is the storehouse for its memories. Through the perishing of the occasion, the 'memories' constituted by the structured relations between the personal nexi and its environment are constantly updated and made available to subsequent higher-order occasions in the personal social nexus that defines the agent. In a similar manner, ρbot is a social process, embedded in a broader social environment of processes. It is a particular social thread in Lambert's brick-wall universe of causally related fragments of process. Fragments of the ρbot societal process are causally related; hence sensor process fragments (that act to discriminate the presence of specific qualities in its environment) and template fragments (that individuate functional patterns amongst the qualities detected by discriminators) interact with societal computer memory fragments to register and maintain the representation of the abstract structure of template processes. Thus, both Whitehead's habits and Lambert's templates can be viewed as substructures embedded in the agent societal structure and these relational substructures provide a purposeful context for the agent interpreting its situation.

The nature of the agent's embedded relation to its environment reflects the purposeful orientation of the agent; that is, behind the agent's simplified representation of its world are functional goals that are related to the agent's role in its environment. Just as Whitehead's social occasions constitute their own partial world through their perspectival unification of a datum, so also Lambert's ρbot uses discriminators and templates to selectively attend to qualities and through this selective attention causally define the primitive contents of partial worlds.[11] The primary difference between the organic and machine 'habits' is that the former arise through natural evolution, while ρbot's guiding habits are humanly engineered to reflect human interests. Hence the machine agent picks up primitive and functional qualities because the classifiers are embedded in the world that they classify, and have been designed and programmed to pick our qualities that are relevant to the machine's functional role in relation to that environment. The qualities classifiers detect are in effect "roles played by processes within patterns of influence" (*EMWC:*185).

[11] Discriminators and templates therefore mediate ρbot's selective prehensive connection with its world. This objectification is based upon a combination of the efficient causation of the world acting upon discriminators, and templates acting as a final cause, selectively unifying the detected qualities in order to satisfy the aim provided by the template.

Thus we attend to qualities because they signify a role that specific processes play in a repeated pattern of influence. We see the primitive quality light, for example, because of the influential role played by light in mediating numerous functionally important habits. Our adaptation to our environment builds upon these qualities, as for example, when light is assigned a role in the operation of traffic signals. In practical terms, ρbot's representation is the result of selective attention, mediated by discriminated qualities, to processes of fragmentation, temporal precedence and influence. Temporal precedence is registered in terms of internal changes in ρbot. Discriminating patterns of influence depends upon monitoring internal change within ρbot and relating this change to external changes. Detecting fragmentation depends upon noting patterns within patterns of influence.[12] These relations of temporal precedence, fragmentation and influence can all be specified in terms of Lambert's set-theoretical notation in order to define a structure and generate a set of valid sentences (a theory) that describes the partial world. However, the structures so specified refer to qualities, hence the structure picked out be a template relies is an intensional structure. Lambert thereby arrives at his Intensionality Tenet: "Each intensional structure is a selective record of patterns of qualitative change, and thus a selective record of patterns of influences formed from changing patterns of influences in the world" (*EMWC:*188).[13]

5.13 The Canonical form of Post-Classical Representation

To sum up, Lambert's introduction of an alternative basis of representation has focused on repeated patterns of influence in the world and shown how they naturally engender a qualitative intensional view of the world. He argues that sophisticated representations of the world rest upon very weak representational devices and the representations that they form about causal influences between fragments. Representation is not about the relationship between manifest objects and facts and their corresponding tokens; it is instead about "a connection between processes in the world and representations recording the elementary structure of past processes" (*EMWC*:189). In Whiteheadian terms, Lambert's discriminators and templates are together picking out the bare logical subjects and individuating them as objects of potential interest. This representation does not involve the complex form of felt

---

[12] Lambert points out that our intentional classifications of the world are a natural product of a similar attentiveness to qualitative change. Our awareness of temporal precedence and patterns of causal influence underpins our capacity to differentiate our world into interrelated fragments or societies of societies in Whitehead's terms. Lambert's machines, like Whitehead's organisms are embedded in a world of sensa to which they have been functionally coupled, via structures isomorphically tuned to each other via qualitative sensitivities. Lambert notes that:

> Intensional representation is the natural outcome of this kind of multi-layered qualitative pattern tracking. The intensional structure that we associate with a clock is indicative, with the clock face providing a pattern of constancy, and each of the hands providing different patterns of constancy within it. Therefore we come to think of the clock as a process fragment possessing a number of components, and depending upon how we attribute these components, arrive at different intensional configurations for the clock (*EMWC*:188). [see clock diagram, figure 5.3]

[13] On a philosophical level, the qualitative structure embedded in a template does not imply ontological or epistemic identity in the patterns of change that the template individuates; rather, Lambert posits a weaker connection - cause and change frequently coincide such that pragmatically, what we do is often influenced by this connection. As with Whitehead's purposeful organisms, the qualitative change detected is more a claim about the organism's practical relationship with its environment, than about the ultimate actualities of that environment, or about claims to knowledge (*EMWC:*187).

satisfaction that mediates human objectification, however, in accordance with philosophical functionalism, Lambert argues that "it is sufficient to capture the function of human representation in some architecture, whether or not it be an architecture identical to that engaged by humans" (*EMWC*:190). Hence, the representational devices used by ρbot may be functionally analogous to those employed by humans, though very different in terms of the way that they are embedded in the agent architecture.

In the next section I will describe how Lambert builds upon this notion of representation to specify the synchronic aspects of templates so as to capture the qualitative composition of the row of bricks that make up a duration of a partial world, and also the diachronic aspect that captures temporal transitions as successive columns of bricks are added to the partial world.

5.2 The Synchronic and Diachronic Dimensions of Representation

The previous section introduced Lambert's alternative philosophy of representation, in which repeated patterns of influence serve as uniformities. In Whiteheadian terms these repeated patterns of uniformities are the eternal objects qualifying societies or complex combinations of societies that are objectified in a datum of an occasion. In Whitehead's philosophy this objectification is mediated by a habitually defined process of simplification and unification leading to a satisfaction. Thus I pick out the society that I objectify as the mug of coffee on my desk by prehending through an indicative system a persisting conjunction between the nexal set of the logical subject and corresponding eternal object. In Lambert's world these indicative uniformities are represented model theoretically by a variable assignment into a partial world that satisfies a set of well formed formula that specify relations of fragmentation, temporal precedence and causal influence between fragments of process.

As noted in the previous section, Lambert proposed that patterns of influence are detected using discriminators and templates. The former are simple, fixed physical devices that disclose primitive qualities, while the latter recover context related functional qualities. Templates are in effect weak representational devices for "identifying clusters of qualities (their synchronic dimension) and for monitoring changes in them over time (their diachronic dimension)" (*EMWC:*191). But how is a machine to represent the changing pattern of qualities manifested by a society during its typical activities? Different phases in the life of a tree society, for example, may be indicated by qualitative changes in its size, in the presence or absence of leaves, flowers, changing colours, and so on. How is the machine to represent these qualitative patterns of change?

Lambert's solution draws upon the notion that the processes constitutive of his brick-wall universe have a synchronic and diachronic dimension; this structure can be envisaged as a sequence of societal states and transitions, whereby at any moment the universe represents a societal state, and the addition of each column of bricks in his universe signifies a transition to a new state. In a similar sense, templates can be viewed synchronically as the representation of *states* of a partial world, and diachronically as the representation of *transitions* between these partial states. "When conceived in this manner", Lambert argues, "templates are state transition networks representing evolving processes" (*EMWC:*192).[14] As state transition networks, templates are constructed from nodes and edges between nodes. Templates, that is, are directed graphs. States of a partial world are "persistent gatherings of individuated processes, and are the building blocks of the represented world" (*EMWC:*195).

5.21 Syncronic representation

The goal of Lambert's compositional analysis is to provide an intensional structure that includes not only the processes being individuated but also the manner in which they are individuated. That is, in Whiteheadian terms, Lambert's aim is to pick out both the public and private sides of the relevant prehension. The process that I am objectifying in the form of a particular eternal object can, in other words, be felt subjectively in different ways. As in Whitehead's system, Lambert asserts that the "basis for this intensionality is selective attention through qualities" (*EMWC:*199). Since qualities are the basis for individuating processes, they must also be represented in the state transition network (*EMWC:*198). Thus, alongside the extensional definition of the basic elementary processes (bricks) in Lambert's set theoretical formalism, Lambert introduces urelements, represented by ur-nodes, to represent the intensional qualities that are the basis of individuating processes in various ways.[15] Hence the state transition network is able to represent extensional processes while distinguishing the qualities that define the process as an intensional object.

State structures represent states of a *partial* world, limited both in terms of processes identified and by the selection and assembly of qualities used to identify those processes. Synchronic representation delineates persistent states in a partial world, using *qualities*, *composition* and *absence* as the only representative elements. A key characteristic of partial states is their composition. For example, if a clock is regarded as a partial world, made up of fragments of process, the world (clock) can be decomposed into face and hands, and the hands further decomposed into hour hand and minute hand. This compositional structure can be represented by allowing nodes to represent processes and edges between them to signify composition.[16] Absence can be signified in a similar manner.[17]

---

[14] A state transition network is a system made up of nodes, representing states, and edges representing transitions. They can be augmented with further notations to form the graphical basis for the analysis of finite state machines (computational automata). In effect, by representing alternative states and the transitions between them, a transition network can be viewed as a system with various memories, each specifying '***if*** *you have arrived in this state,* ***then*** *you must. . .* ' . Transition networks are useful for envisaging and programming the actions of automata in computer games that are triggered by particular events.

[15] In terms of Lambert's set theory, processes are sets, which may be made up of further sets. Urelements (from the German prefix *ur*, meaning primordial), on the other hand, cannot be sets; they are the basic atoms or individual components of sets. In Whiteheadian terms, Lambert's sets are the nexal sets that make up Whitehead's societies, and the analogous qualitative elements used to pick out these societies are the eternal objects that give a characteristic subjective form to each society.

[16] "States therefore have a structure in these template state transition networks formed from edges *between* nodes" with some edges used to construct representations of "states of a partial world" and other edges used to represent "transitions between these partial states" (*EMWC:*195).

[17] Absence refers to a mode of intensional classification of the qualities of some process that are absent. For example, the clock can be defined by various sets of process states that include either {ticking} or {not ticking}. If ticking is represented in the graph by drawing a node as an empty circle, then the quality 'not ticking' is represented by enclosing that circle within another circle. See fig. 5.5; also Lambert (1996:201-2).

These state structures can also be conceived of as sets and hence for practical purposes the framework of set theory can be used to represent states. Thus, to use the previous example of the tree (as process), drawing upon different compositional templates it can be identified intensionally as either ((stem, leaves) trunk) or (stem, leaves, trunk), yet extensionally the reference is still to the same chunk of processes constitutive of the tree (*EMWC:*199). Thus, the same ur-nodes can pick out a different order of set nodes, giving different intensional meanings to the same extensional object. Hence the same state structure can be used to represent one or more concepts of the process object, with the quality urelements represented by empty 'ur-nodes' and process sets represented by solid black nodes. Composition is represented by the edges joining the nodes as in the case of the tree example below (fig 5.2).[18]

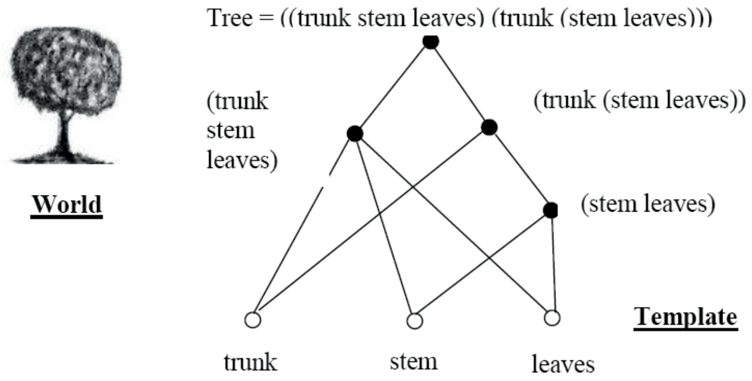


Figure 5.2: State structure representing two alternate concepts of a tree (adapted from *EMWC*:200)

The way in which Lambert's intensional description of a world is related to its extensionally defined composition can be further illustrated with reference to the clock example. If the qualities (urelements) used to individuate the process 'object' sets that constitute the clock are as described textually as follows:

Hour hand = (short, dark);

Clock hands = ((short, dark), (long, dark)); and

Clock = (((short, dark), (long, dark)), (grey, round), ticking),

then the relationship between the qualities and the processes could be represented graphically using a state structure; as before, the quality urelements are represented by ur-nodes and process sets represented by nodes. Partial states are thus finite intensional structures grounded in qualitative identifiers. Composition indicated by the list structure in the textually defined sets is represented by the edges joining the

[18] The tree example illustrates the intensional role of quality nodes in indicating the corresponding set nodes. The urnodes in this case are not the simple raw feelings that a general purpose discriminator would detect, unless it was engineered specifically to detect 'trunk', 'stem' and 'leaf' objects. The relationship between discriminator urnodes and template setnodes is somewhat ambiguous and could perhaps be likened to the relationship between eternal objects of the subjective and objective species (where the EO of the objective species indicates an extensional reference to the relational structure of the nexi prehended in the datum and the EO of the subjective species is an intensional reference to the subjective form felt by the prehending subject, as described in chapter 3.21 of this work and depicted in figure 3.4). In successive transmutations both the objective form of the datum and the subjective form of feeling felt by the subject are transformed, such that corresponding objective and subjective species of EO are transformed, in the fashion indicated by Henry (2003) as outlined in fn. 10 above.

nodes. The same process can also be classified in terms of the absence of a quality; for example, the Clock described in the list above could be represented as *not* ticking.

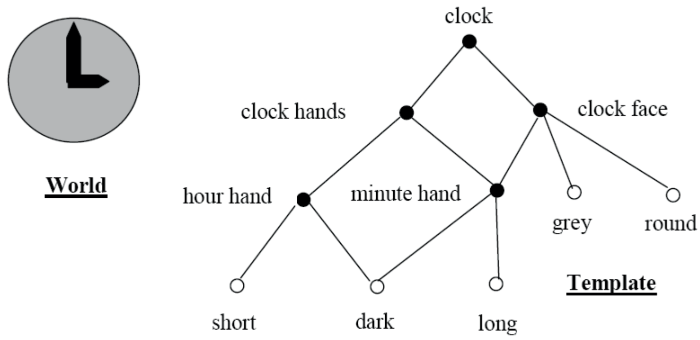


Figure 5.3: Template showing qualitative compositional structure of a world (clock) (adapted from *EMWC*:198)

5.22 Diacronic representation

Partial states rarely persist indefinitely, and so diachronic representation is introduced to deal with the temporal dimension of the template (state transition network) represented by the directed graph. This is accomplished using *transition types* that register the changes in the component nodes as the network changes from one state structure to the next. Three possible changes can take place in the transition from one state to a subsequent state; the previously existing node is retained in the new state structure (*qualitative inheritance*), the previously existing node is discarded in the new state structure (*qualitative termination*), or a previously non-existent node is introduced into the new state structure (*qualitative introduction*). Any qualitative change between successive states can be described in terms of patterns of inheritance, termination or introduction of qualitative influences. To use a musical analogy, diachronic representation could be likened to a musical composition, in which qualitative inheritance is the repetition of the same note through successive beats, qualitative termination involves the termination of a note, while qualitative introduction entail the introduction of a new note. Qualitative transitions between beats (states) thus enable the overall structure of the tune to be represented.

As these patterns are recorded in the robot's memory, they provide the robot with a template for attending to transitions in patterns in the qualitative data derived from the external world. That is, the self replicating patterns of influence recorded in the robot's memory are isomorphic with the self-replicating patterns of influence in the brick-wall universe that differentiate the enduring 'objects' of our world. Thus, in a temporal sense, these enduring objects can be specified in terms of a succession of transitions between state structures, in which the successive states are defined by regular patterns of change, such that a new node is introduced, an old node discarded, or an old node reproduced in the next temporal instance of the state structure. These transition types allow new states to be defined in terms of former states, so that transition types represent sequences of change relating to specific qualitative nodes in the state structure (*EMWC:*203).

Transition deals with what has been referred to as the 'Frame Problem' (McCarthy and Hayes, 1969), namely, given the context sensitivity of knowledge, the problem of delineating the significant factors that *do not change* between states (or 'frames') that may be relevant to our understanding of a state. Lambert's solution to this problem is related to his notion of habit.[19] As in the case of Whitehead's more advanced forms of monadic society, Lambert identifies habit or custom as the crucial factor that facilitates the individuation and objectification of an environment into processes of significance. Lambert's procedure for endowing ρbot with the capacity to discriminate sequences of transition types is an expression of Lambert's vision of a Humean form of intelligence, in which custom rather than discursive reason plays the crucial role in ρbot's grasp of its situation. In Lambert's words:

> Qualitative inheritance, qualitative introduction and qualitative termination are sufficient to describe any qualitative change between successive states. By recording sequential transitions of partial states in memory, ρbot is able to exploit those sequences when they locally repeat in the world. And because they do repeat in the world, and because the nature of those patterns is recorded in the robot's memory, repeated patterns of influence become that which is selectively attended to in the world. This, I contend, exposes the very kernel of our abilities, and resembles the sentiments expressed by David Hume in the epigraph [on the role of custom; cf. the introduction to this chapter]. Recorded patterns of influence account for our commonsense, not the inferentialist principles of rationalism, and in recording these

---

[19] Michael Wheeler (2005) has argued that the crucial test for the viability of any proposed AI program is whether it can solve the frame problem. Thus, if a computer is representing its world in a particular way, and the world is a dynamic, constantly changing world, the computer must decide which changes to the state of the world are significant and, if it does detect a significant change, it must then decide which of its stored beliefs needs to be retrieved and updated, in order to adequately represent and deal with the changed situation? Given that there could be an infinite number of factors, many of which *could* be significant for interpreting the significance of a state, the frame problem seems to suggest the need for theories of the world that deal with these infinite possibilities (patterns of causal influence and their branching potentialities). The challenge of the frame problem is to deal with the potential significance of context without becoming locked into exploring infinite loops of possibilities.

Lambert's solution is based upon the presupposition that the agent, through its stock of routines (mediated by discriminators and templates), is adaptively tuned to its environmental niche, and that (in a Heideggarian sense) the environment calls forth the appropriate response from the agent through the patterns of causal influence that mediate the computer's awareness. From the embedded-situated agent perspective, the agent is not overwhelmed by infinite possibilities. Its perceptual awareness of its environment and of their potential significance is finite. Templates provide a context or focus of attention that limits the scope of the processes that are considered and resource constraints built into the computer's architecture place further limits on the internal activities that the computer engages in. What is not represented is not considered and is therefore not a problem, as the robot does not register it as a concern. The situation is not unlike that which applies to human activity, where the time priorities associated with accomplishing a particular task, coupled to our routine procedures for carrying out that task, steer our attention toward relevant details and decision making timelines.

Thus, Lambert's agents are not overwhelmed by an infinitude of implicit possibilities. What does concern the agent are those things that are explicitly represented in its "lego-block" construction of its situation. A template of a situation represents potential transitions between state instances. When representing significant processes in its environment, qualititative inheritance reports what does not change, while qualitative introduction and qualitiative termination provide explicit reports on what and how a state has changed. Thus, discriminators and templates register significant patterns of influence in the real world, and the only thing the robot does notice are those that are explicitly represented as changes or persistence across contexts, as specified using qualitative inheritance, termination and introduction (*EMWC:*204-5). Further more, because of Lambert's 'semantics of success', the computer's routines are constantly being fine-tuned through use to the agent's environment. Although Lambert's approach does not deal with all aspects of the frame problem as described by Dreyfus, it can be viewed as an alternative form of the embedded strategy advocated by Dreyfus (Dreyfus 2007:262-5).

> patterns, the epistemically unjustified presumption is that the future will to an extent be conformable to the past (*EMWC:*206).

The reference to the 'epistemically unjustified presumption' in the above quotation from *EMWC* is a reference to Hume's position; however, as I have argued previously, Whitehead provides grounds for affirming our commonsense, habitual expectation that the future will to some extent conform with the past. However, this conformity is weak in a causal sense, and as a result higher organisms are adept at adjusting to minor variations in the patterns that they associate with particular phenomena. To achieve similar flexibility the patterns recorded in the robot's memory must exceed sequential transitions of partial states; the same qualitative configuration may repeat on some occasions and be disposed to differ on others. Lambert's templates cater for these differing dispositions by allowing each of the identified dispositional transitions belonging to the same class to be superimposed. Thus any of the multiple transition connections that make up a superimposition class can represent the passage from one state to another. They are in Hume's terms variations of the same 'custom' (*EMWC*:206-7).[20]

The two fundamental assumptions underpinning Lambert's notion of superimposition are those of structural continuity whereby the successor nexi of actual entities in the chain of states constituted by successive social transitions retain the same functional role. The red light in the traffic signal remains the red light even when it is switched off, such that all qualitative indicators of the composition of a state in one moment are continuous with the same elements in successive moments. This is akin, in Whiteheadian terms, to insisting that there is a thread of prehensive continuity between the society and its various societal substructures through successive cycles of concrescence and transition. Thus the societal state indicated in the datum of a concrescence will retain its structural isomorphism with the state of the nexal structures objectified in successive states. Hence, Lambert affirms that a process fragment will play the same role indicated by the same quality in successive states:

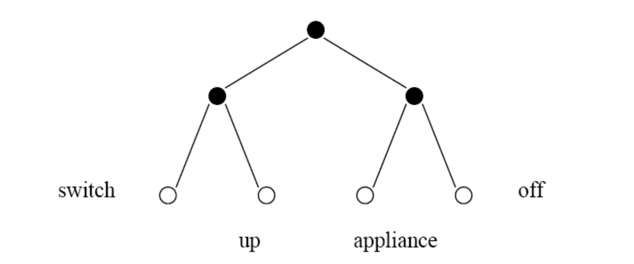
- When superimposing two different instances of transitions from a state, and regarding them as dispositions from the one state, "we assume *the present is conformable to the past* by treating the new instance of the state as the old instance of the state".
- We form this superimposed structure because we expect these dispositional transitions to surface again, "and this *assumes the future will be conformable to the past*" (*EMWC*:207).

Thus, the transitions within a template must, in effect, be segregated into superimposition equivalence classes where the members of the transition class possess a common quality that is prehensively related to the previous subjective quality.

Take, for example, a simple system made up of a switch connected to an appliance. This system can cycle between two states; in one state the switch is up and its appliance off, while in the second state the switch is down and its appliance on. A pictorial representation of one state is given below, with the extensionally defined processes represented by the solid black nodes, and the intensional, qualitative

[20] Lambert explains his notion of 'superimposition' with reference to Hume's conformity presupposition: inferences from experience are based upon custom, whereby a similar train of events in the past is superimposed upon the new train of events, forming the basis of our capacity to anticipate future events (*EMWC*:206-8).

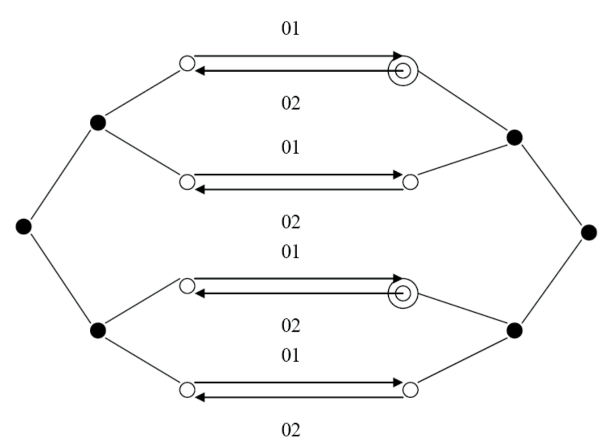
classifications represented by the empty nodes. The upper-central black node in the template represents the state of the partial world 'switch-appliance' in its turned off state, while the lower two solid nodes represent the compositional structure of that partial world (switch-up and appliance-off), with the empty nodes representing the qualities (switch, up, appliance, and off) used to individuate the compositional structure.



Pictorial depiction of a switch / appliance state (Lambert, 1996:209)

Fig. 5.4 Picturing a state.

In the switch-appliance system there is only one transition equivalence class from the first to the second states (inheritance of switch quality x→x, appliance quality u→u, up quality y→y and off quality v→v). From second state to the first involves the same qualitative inheritance transitions (x→x, u→u, y→y, v→v). A pictorial representation of these qualitative inheritance transitions (group of edges labeled 0), together with superimposition transition classes (the groups of edges labeled either 1 or 2), as represented below, shows the structure of the switch template defined in terms of the states and the qualities manifested in transitions as it cycles between the two states (*EMWC*:210).



Pictorial depiction of switch template (Lambert, 1996:210)

Fig 5.5 Picturing a template.

Black nodes indicate the processes (switch; appliance as in fig 5.4) with edges representing their composition. Empty nodes represent the constant qualities (switch; appliance) and variable qualities (up; off) used as a basis for individuating the state of the corresponding processes. The empty nodes with a double circle represent absence of the quality (absence of the quality ‘offness’ and ‘upness’). Qualitative inheritance transitions are portrayed as directed arrows, with those of the same class labeled with the same natural number. The template thus represents the intensional basis used to identify the two possible states of the transition network: ‘switch up and appliance off’ or ‘switch not up and appliance not off’.

Although the terms used to describe this template (switch, appliance, up, off) are theory laden linguistic terms, in the context of the template, they are only tokens signifying a structure. Thus the same switch described textually by:

[((switch up) (appliance off)) ((switch –up) (appliance –off));
Switch→switch, up→up, appliance→appliance, off→off]
[((switch -up) (appliance -off)) ((switch up) (appliance off));
Switch→switch, up→up, appliance→appliance, off→off],

could be represented with no loss of information as:

[((x u) (y v)) ((x -u) (y -v)); x→x, u→u, y→y, v→v]
[((x -u) (y -v)) ((x u) (y v)); x→x, u→u, y→y, v→v].

Alternatively, graphical and textual representations could be combined, as in fig. 7.3.4., with alternate states to the left and right of the diagram, and transition classes listed above and below the diagram.

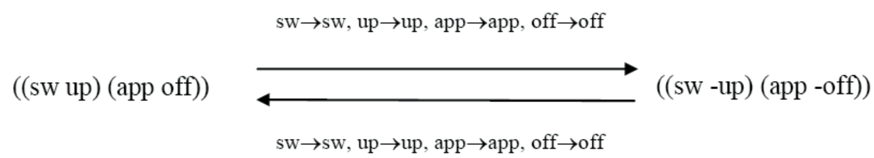


Figure 5.6: Pseudo-pictorial depiction of a switch template (*EMWC*:211)

By building state transition graphs in this manner a template is able to represent a process such as a switch connected to an appliance, transitioning between its states, without introducing theory laden tokens. Instead its structure depicts the composition of a flux in which a succession of states containing multiple objects with both constant and transient qualities represented. In this way the pattern of influence connecting switch to appliance can be represented, irrespective of the qualities used to identify the processes that make up the flux of events. Note that the switch appliance example is extremely simple and does not reflect the potential power of Lambert’s notation. A more complex state structure may include many states, and some of these may involve the introduction of a new qualitative state structure, absent in the current state, into the succeeding state ($\varepsilon \rightarrow y$), or vice versa, discarding a qualitative structure that is present in the current state ($x \rightarrow \varepsilon$).

In order to demonstrate the potential complexity of his representational structures, Lambert introduces a second kind of representative scheme that employs a computer data structure referred to as a ‘stack’. A stack can be envisaged as a collection of lego blocks in which only the most recently added block may be removed.

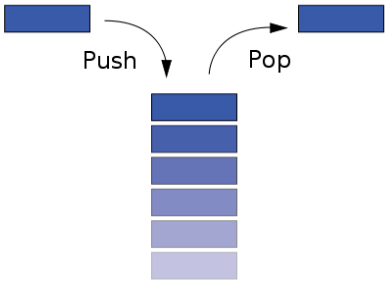


Fig.5.7 Simple Representation of a Stack (Wikipedia)

The addition and removal of items is effected by two basic operations, ‘push’ and ‘pop’, with the last item added to the top of the stack always the first removed, as represented in figure 5.7. Hence it is only possible to access a lower element by first individually removing each of the elements above it. In *EMWC* Lambert provides a pictorial representation of a stack, as depicted in figure 5.8, with the circular arrows represent the two species of recursive push and pop operations, while the straight arrows represent the push and pop transitions from and to an empty stack.

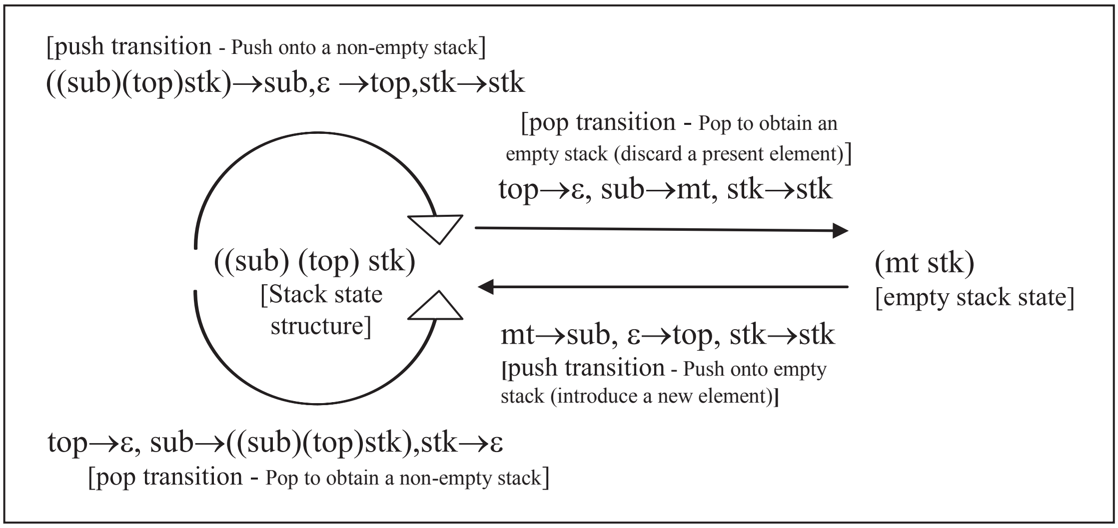


Figure 5.8 Pseudo-pictorial depiction of a stack template (*EMWC*:212)

Thus, as depicted in the diagram, a stack has two possible states:

i. non-empty ((sub) (top) stk)
ii. empty (mt stk),

and four possible transitional operations, each of which can be represented in terms of superimpositions on the state structures (using the class transition notation x→y, y→x, x→ε, ε →y, x→x, y→y). As with the previous graphical mode of representing states and state transitions, stacks are a way of representing structural patterns in qualitative change, such that a new node is introduced, an old node discarded, or an old node reproduced in the next temporal instance of the state structure. Each of the tokens refers to a quality which can be instantiated during execution and the transitions are really specifying patterns for how the instantiations of these qualities can change, so that transition types represent sequences of change relating to specific qualitative nodes in the state structure.[21]

---

[21] Lambert (2009b) has commented on the stack example provided in *EMWC*, and illustrated its application with a Prolog implementation. His implementation complements explanations provided in the next chapter as to how ρbot acquires beliefs. It is also an interesting point of comparison with Henry's (1993) use of Prolog programs to represent the process of concrescence. As will be argued in the next chapter, both Lambert and Henry regard forming a belief about a process as analogous to a pattern matching process initiated by a Prolog query. Lambert's example is based upon a pictorial representation of a stack, beginning with a base foundation plate (numbered 0), signifying the empty state, and a stack of lego blocks (numbered 1, 2, and 3), that are 'pushed' onto the empty stack.
In part, Lambert's commentary reads as follows:

> As characterised in *EMWC*, there are two states a stack can be in: (a) an empty stack and (b) a non empty stack. In *EMWC* I represented these by: (a) (stk, mt) and (b) (stk, (sub), (top)) respectively (actually the order of arguments was slightly different). I then identified that four transitions were possible.
>
> i. Push onto empty stack:
>
> mt→sub, ε→top, stk→stk
>
> ii. Pop to obtain an empty stack:
>
> top→ ε, sub→mt, stk→stk
>
> iii. Push onto a non-empty stack:
>
> (stk, (sub), (top)) →sub, ε →top, stk→stk
>
> iv. Pop to obtain a non-empty stack:
>
> top→ ε, sub→ (stk, (sub), (top)), stk→ ε.
>
> Suppose I wish to implement this structure. If I do this in Prolog, then the two possible states become: (a) (Stk, Mt) and (b) (Stk, (Sub), (Top)), with the uppercase signifying to Prolog that these are variables.
>
> The transitions can then be captured by patterns of the following forms:
>
> push($State_k$, Top_{k+1}, $State_{k+1}$)
> pop($State_k$, Top_k, $State_{k+1}$).
>
> The first refers to a transition in which we have $State_k$, Top_{k+1} and $State_{k+1}$ consisting of Top_{k+1} on top of $State_k$. As there are two possible forms (a) and (b) for $State_k$, the transitions i and iii noted above are represented in Prolog code by the push transitions
>
> ```
> push((Stk, Mt), Top, (Stk, (Mt), (Top))).
> push((Stk, (Sub), (Top)), Top2, (Stk, (Stk, (Sub), (Top)), (Top2))).
> ```
>
> respectively.
>
> The second refers to a transition in which we have $State_k$, Top_k and $State_{k+1}$ consisting of $State_k$ with top Top_k removed. As there are two possible forms (a) and (b) for $State_k$, the transitions ii and iv noted above are represented in Prolog code by the push transitions
>
> ```
> pop((Stk, (Mt), Top), Top, (Stk, Mt)).
> pop((Stk, (Sub), (Top)), Top, Sub).
> ```
>
> respectively.
>
> To illustrate execution, suppose we name a stack of numbers number_stack and start with an empty stack, taken to be the state (Stk, Mt) = (number_stack, 0) i.e. 0 signifies an empty stack. If we then push consecutive numbers onto the stack and then pop them off, then pattern matching across the nominated transitions
>
> ```
> push((number_stack, 0), 1, S1),
> push(S1, 2, S2),
> push(S2, 3, S3),
> pop(S3, T3, S4),
> pop(S4, T2, S5),
> pop(S5, T1, S6).
> ```
>
> produces
>
> ```
> S1 = (number_stack,0,1)
> S2 = (number_stack,(0,1),2)
> S3 = (number_stack,((0,1),2),3)
> S4 = (number_stack,(0,1),2)
> S5 = (number_stack,0,1)
> S6 = (number_stack,0)
> T1 = 1,
> T2 = 2,
> T3 = 3.
> ```

In the next chapter it will become apparent that stacks provide a key mechanism for matching patterns of qualitative inheritance in the world with patterns of potential inheritance specified by a template. Lego block stacks can be assembled on the fly to represent dynamic situations. These computational structures are mathematical structures – akin in Whiteheadian terms to complex eternal objects – that are capable of depicting changing patterns of qualitative composition.

5.3 Conclusion

In summary, Lambert's representation scheme differs from the mainstream Classical AI system "not only in how it represents, but also in what it represents" (*EMWC:*211-2). Rather than assume that the Manifest Image of the world presented in Natural language provides an adequate system of representation that is implicitly presenting an adequate ontology of the world, Lambert's approach to representation builds upon a bare process ontology, defined in terms of fragments of process that causally influence other fragments, and a notion of representation based upon the presupposition that the spread of causal influences between processes can cause repeated patterns of influence. For Lambert, these repeated patterns define the uniformities that we experience; they naturally engender a qualitative intensional view of the world which is the basis of representation for Post-Classical AI.

In this chapter I have argued that Lambert's self-replicating patterns of influence can be likened to Whitehead's societies, and suggested that the partial world representations mediated by Lambert's templates are akin to a Whiteheadian proposition (theory) that specifies a simplified version of the characteristic form or eternal object of a Whiteheadian society. The analogous function between Lambert's discriminators and Whitehead's conformal physical feelings, mean that the habitual patterns of activity captured by Lambert's templates, like the societal habits that define Whitehead's universe, have a reference to the real world.

Although I have used words to describe Lambert's templates the words are unnecessary for the fulfillment of their semantic function; any token would suffice, as the template makes no use of the semantic content of the words. The meaning of the template is embedded in its representation of the qualitative intensional structure that underpin both the synchronic and diachronic aspects of processes; the routine specified by the template provides a mechanism that disposes the robot to attend to repeated patterns of influence that match the template.

The role of Lambert's urnodes and setnodes can be likened to Whitehead's eternal objects of the subjective and objective species, while the combination of nodes that make up a template correspond to the characteristic complex eternal object of a society, in that they specify an abstract pattern that matches the qualitative pattern of influence manifested by a nexal set of actual entities prehended in a datum. As in the case of Whitehead's propositions, Lambert's templates are functionally oriented structures, with a selective role analogous to customs or habits. Templates, via appropriate discriminators, embed the robot in its environment, and provide a qualitative context for interpreting that environment. Lambert, in a defining remark, states:

> I contend that structured 're-presentations' of fluxes in qualitative composition of just this sort provide the basis for our attention to processes in the world and serve as the

> *drivers of commonsense*. Post-Classical representation is founded upon this idea. The Manifest Image symbolism of Classical AI is discarded in favor of post-classical representations structurally reporting patterns of influence in the world (*EMWC:*210 [my italics]).

Hence, representation is not about the relationship between manifest objects and facts and their corresponding tokens; it is instead about "a connection between processes in the world and representations recording the elementary structure of past processes" (*EMWC*:189). Routines that selectively draw our attention to patterns of influence in the world, and provide a context of anticipations for interpreting those patterns, are the drivers of commonsense. The next chapter will outline how the structured representation of qualitative fluxes fulfills the function of Whitehead's prehensions of subjective form, providing a link between qualities and processes, and allowing ρbot to form beliefs about the world.

Chapter 6

Towards a Whiteheadian interpretation of Lambert's Post-Classical philosophy of Belief

Chapter 6 – Section Headings

6.0 Introduction: Process, Representation and Belief

The robot believes what the robot classifies.
(Lambert, in *EMWC*:219)

People find it useful to explain other people's behavior in terms of the beliefs motivating their actions; most AI engineers likewise assume that it should be possible to understand the behavior of an intelligent machine in terms of the machine's beliefs. But what are 'beliefs' and can they be implemented in a machine?

Since the time of Plato philosophers have defined knowledge in terms of true beliefs and sought to clarify how we come to possess true beliefs; with Descartes and Locke this debate was framed in terms of epistemology, however with Kant the question of how ideas mediate and constrain our quest for the truth came to the fore, and by the time of Wittgenstein and Heidegger, the epistemic quest was largely framed within debates concerning the nature and role of language. Instead of viewing linguistic utterances as the neutral instrument of thought and inquiry, language and its representational role itself became a focus of analysis. Thus Analytic philosophers asked how linguistically framed ideas connect us with the world and sought to answer this question by analyzing the representative function of language, while Continental philosophers asked whether the hermeneutic circle of interpretation and our bodily being in the world provided any grounds for privileging one system of beliefs over another.[1] These debates formed the context of the Classical AI community's attempt to engineer a machine that could simulate the human capacity for analyzing symbolically coded information and formulating justifiable beliefs.

Lambert suggests that most proponents of Classical AI framed their notion of belief within the presuppositions of analytic philosophy (as discussed in chapter two). Analytic philosophers have generally assumed that beliefs are logico-linguistic entities that instantiate an intentional connection between the content of the belief and the fact that it signifies. Typically, this connection takes the form of a propositional attitude, whereby an agent has a mental state that expresses some attitude about the content of some proposition ("I believe that snow is white"), and the meaning content of the proposition is true if there is a corresponding factual referent (that snow really is white). Hence beliefs have an intentional quality, in that each belief is *about* that which the proposition refers to. This intentional quality of beliefs was the focus of my discussion in chapter two of this work; the internalist and externalist positions discussed in that section were characterized by a realist or semi-realist orientation toward belief.[2] However, as argued by Lambert, analytic philosophers have tended to rely upon the discredited manifest image notion of linguistic intentionality, and hence Classical AI approaches to belief shared in the failings of analytic philosophy. His Post-Classical paradigm, as outlined in *EMWC* was an attempt to address these failings.

Lambert's Post-Classical AI is based upon process ontological foundations, and a chain of conceptual dependency that leads from his process formulation of 'what there is', to a process formulation of 'what is represented', and a process grounded approach to 'what is known'. Lambert brings together the process notion of a world made up of patterns of influence and the notion that appropriate discriminators and templates can re-present the patterns of influence that define the uniformities in the real world, and builds upon these prior conceptual foundations to provide an answer to the question of 'what is known'. His response is formulated in terms of a theory that he terms 'Compliant Realism'.

---

[1] Both Analytic and Continental philosophers recognized that thoughts and linguistically framed propositions share a property of intentionality; they reach beyond themselves to suggest something of something which might or might not be true, however the question of why some propositions are regarded as valid beliefs remained clouded in controversy. Psychological inquiry was mired in similar debates. The founding fathers of the discipline had generally equated beliefs with the content of mental representations and viewed these representations as the basis of our capacity to reflect upon events in our world; yet as in philosophy, no consensus concerning the nature of representation or of belief has emerged, with divisions between schools of thought reflecting the bifurcations within philosophy. Thus, while the intentionality of linguistic representation, and the grounds of the semantic content that make representations true of false have become a focus of interest, the basic nature of belief remains a contested notion.

[2] Lambert argues that proponents of Classical AI have generally taken up one of three positions regarding belief. Either they are belief realists (beliefs are actual states that exist independently of observer attribution), belief semi-realists (beliefs are in part a product of observer ascription), belief eliminativists (the philosophical notion of 'belief' arose as a postulate of a false theory and will ultimately be abandoned) (*EMWC:*214-217).

Before analyzing what Compliant Realism means as a Post-Classical philosophy of belief, I will review the Whiteheadian process philosophical foundations that frame my interpretation of Lambert's position. I will then consider Lambert's suggestion that belief can be understood in Post-Classical terms as an act of classification, and explore what this means in semantic terms. This will clarify the 'semi-realist' position that Lambert stakes out for his Post-Classical paradigm (Compliant Realism). I will then outline how Lambert builds upon his formal notions of process and representation in order to specify a formal extensional semantics. Lastly, I will summarize the implications of Lambert's position from a Whiteheadian standpoint.

6.1 From Whitehead's notion of belief to Lambert's Post-Classical Belief

In theorizing the nature of the intentional connection between language and the world, Lambert has argued that Classical AI's primary sources of philosophical inspiration were the 'externalist' referential and 'internalist' psychological approaches developed within the Anglo-American analytic tradition. As outlined in chapter two, the competing Realist intuitions that have informed these orientations towards semantic content were that:

- tokens are about things in the world, and hence content is reference into the world, or
- tokens are about interpretations in the head, and hence content is psychology.

Lambert pointed out that the apparent polar relationship between these positions was undermined by weaknesses that led each back toward its opposite; the externalist back into the mind, and the internalists into the world. From Lambert's standpoint both externalist and internalist positions were flawed but partially valid explanations of belief. What was needed, Lambert argued, was a Post-Classical explanation of belief, framed in terms of a metaphysical system that recognized the complementary relationship between the internalist and externalist positions. The necessary metaphysical foundation, Lambert suggested, could be found by turning to the process philosophical tradition initiated by Heraclitus.

I have interpreted Lambert's Post-Classical project as an attempt to apply the formal tools of AI within a process oriented philosophical context. His effort is an AI equivalent to Whitehead's (1906) attempt to formulate a mathematical model of a Leibnizian process world as an alternative to the Newtonian world. In approaching belief from a process standpoint Lambert had at his disposal the Tarski inspired model-theoretic tools of linguistically oriented extensional semantic systems; however, this tool kit could not be directly applied to Post-Classical AI as Lambert had rejected the linguistic focus of Tarski's model-theoretic semantics. Since Lambert's primary motivation was to develop an understanding of commonsense that rejects the 'Manifest Image' and rationalist philosophical presuppositions that have underpinned linguistically oriented approaches to model theoretic semantics, a key consideration was that of developing an alternative to Classical AI's linguistically oriented understanding of reality, representation and commonsense. Lambert found that alternative in a Heraclitean understanding of process combined with a Humean understanding of habit. Patterns of influence, rather than propositions, became the focus of his analysis.

As argued in chapter three, Heraclitean and Humean insights converge in Whitehead's process philosophy. One advantage of the Whiteheadian framework outlined in chapter three, is that it has provided a context for appreciating the close links between model theory and Whitehead's metaphysical system. These links were considered in chapter four with reference to Whitehead's theory of extensive abstraction, where I argued that model theory serves as a bridge between the world of appearances and Whitehead's formulation of a mathematical model of that world. In chapter five I made a connection between Lambert's notion of partial worlds and Whitehead's notion of propositions, arguing that Whitehead had reformulated the notion of propositions in terms that were consistent with his mathematical model of a 'brick-wall' universe. In this section I will briefly review the role that propositions play in Whitehead's process Universe and consider how this relates to Lambert's attempt to formulate a Post-Classical approach towards belief. Again, Whitehead's mathematically inspired metaphysical system will provide the horizon for appreciating the deeper structural linkage underpinning Lambert's model theoretic connection between patterns of influence and patterns of belief.

6.11 Whitehead's metaphysical approach to belief

The process approach to the issue of belief, as highlighted in my analysis of Whitehead's metaphysical system in chapter three, moves the emphasis away from the role of propositions in their declarative role as an expression of pure acts of intellectual judgment and considers the procedural role of propositions in the habit-defined, pre-conscious processing that underpin such judgments. Whitehead stresses that the primary role of propositions is to focus the organism's attention in a way that is consistent with its habit defined routines. Deeply rooted webs of social habits define the commonsense background for all beliefs, indicating the relevant societal context of particular judgments of truth or falsehood. At the base of these indicated contexts is the basal society of simple physical events whose habits define the spatial and temporal structure of the extensive continuum. This is the ultimate relational ground of agent representation and objectification that Whitehead and Lambert formalized in terms of an ontology of processes.

In explaining representation Whitehead focused on the emergence of societies that share characteristic forms, as disclosed in repetitive qualitative patterns of extensive activity such as manifested by strains. Strains illustrate the most elementary and pervasive form of re-presentation in Whitehead's Universe. Each strain feeling conformally reproduces qualities analogous to the simple sensa that we experience in qualitative feelings of colour or pattern. Strains are the basic societies of Whitehead's Universe, and their characteristic habits are the basis of the deterministic order that defines Nature. As with all of Whitehead's occasions, the patterns of influence that they mediate are in part defined by each occasion's perspectival objectification of its situation, however, in the case of strains the habitual constraints that underpin this selective act of objectification are shaped primarily by the simple aggregation of deterministic social influence, and resultant uniform dispositions.

In contrast, by building upon potentials latent in these basal processes, higher level societies are characterized by the emergence of substructures that physically embed 'propositions' in the society. For example, structures such as sense organs and their associated nervous systems create specialized environments which recursively objectify the parent society's surrounding situation in ways that are relevant to its survival. Thus, for Whitehead, agents have a real and direct strain-based prehensive connection with their environment, but this connection is in part the artifact of societal sub-structures that mediate the transmutation of the chaos of simple feelings into objectifications that reflect the organism's evolutionary adaptation to its world. These specialized physical sub-structures are the basis of the habitual capacity of living societies to creatively unify and objectify their environments; the organism's body is a storehouse of memories in the form of propositions (theories) that involve associating a potential pattern of activity with an indicated situation.[3]

Propositions play a role in Whitehead's social organisms akin to Lambert's templates. They transform an infinite array of hypothetical potentiality latent in an environment into finite, habit defined contexts that disclose the presence or absence of ecological affordances vital to the organism's mode of survival. In placing the accent on the role of habits and propositions in guiding the organism's practical objectification of its environment, Whitehead stresses the pre-conscious functional interests and resultant functional roles that structure an organism's disposition to connect in a particular way with its environment. The bat perspectivally objectifies its environment with habit guided responses to sound patterns discriminated by hearing, an eagle uses sight to discern visual patterns, a wolf experiences a world of patterns individuated by smell; in each case these processes of connecting with an environment are mediated by habits which specify the templates or propositions that provide the context and necessary conditions for satisfying a habitual disposition. Whitehead's analysis suggests that belief is but a specialized embodiment of the role played by habit-based propositions in focusing agent attention and determining agent action. Beliefs are a specialized mode of entertaining propositions.[4] In a general sense, beliefs can be viewed as an instance of the wider (conscious and non-conscious) role played by propositions in the hierarchy of nested societies of habits that in combination direct our attention, determine our aims, and guide us procedurally in realizing those aims. Lambert envisages a similar habit mediated connection between ρbot and its environment.

6.12 Lambert – processing toward belief

As outlined in the previous two chapters, Lambert's interpretation of habit is framed in terms of a process metaphysical notion of existence and a non-linguistic notion of representation. While Whitehead's higher-level organisms bring about the satisfaction of a goal motivated unification of their datum by intermingling feelings, and eliminating contradiction and disharmony, Lambert's agents attain a unified perspectival vision by a linear process of causally mediated binary calculations; this

---

[3] Whitehead's vision of embodied memories as propositions resonates with the renowned ethologist Konrad Lorenz's suggestion that: "Our categories and forms of perception, fixed prior to individual experience, are adapted to the external world for exactly the same reasons as the hoof of the horse is already adapted to the ground of the steppe before the horse is born and the fin of the fish is adapted to the water before the fish hatches" (Lorenz 1982, pp. 124–125).

[4] Beliefs, in Whitehead's analysis, are enacting habit constrained propositions that associate a pattern of potentiality with an indicated situation. In the immediate context of enactment, beliefs involve two subjects, the judging agent and the logical subject of the proposition; propositions also entail an eternal object that manifests itself as an abstract qualitative predication of the logical subject, they also involve indicative systems that link the eternal object with its logical subject, and thereby assign meanings which determine the truth or falsehood of the proposition (cf. chapter 3, pp.101-7).

has important implications in terms of the relative autonomy of the respective systems. From a philosophical standpoint the differences in the meaning of a 'belief' that is grounded in felt values, as opposed to mathematically computed values are of vital significance. However, Lambert's focus is on technical outcomes rather than existential philosophical considerations. Although his Post-Classical paradigm is framed in philosophical terms, it is motivated by the concerns of a practicing AI engineer; thus while he adopts a position on intentionality similar to that of Daniel Dennett (1991) his questions are not those of a philosopher.[5] Hence, he does not ask *whether* a machine can hold beliefs in the sense proposed by some advocates of strong AI but instead seeks to build upon a better understanding of the real basis of human commonsense in order to reform the Classical model of the principles that underwrite commonsense capabilities. The challenge is analogous to the aeronautical engineer that seeks to understand how a bird flies in order to engineer an airplane; not to replicate the organic structure of the bird but to build upon the principles that underwrite the bird's capabilities, such as how airflow over a surface generates lift. An aircraft's wing is, in this sense, structurally analogous to the bird's wing, in that both exploit the same underlying principles in order to achieve the same functional outcome. In a similar sense, Lambert's goal is to engineer a machine that embodies the principles underpinning human commonsense; utilizing these principles to build structures that are in terms of instrumental outcomes the functional equivalent of beliefs.

In the last chapter I suggested that for Lambert's robotic agent ρbot, discriminators and templates play the role of Whitehead's habits, and argued that Whitehead's notions of societies (enduring objects), eternal objects and propositions can be used to explain the functional role of templates in Lambert's process universe. It was suggested that Lambert's self-replicating patterns of influence are akin to Whitehead's societies, and that in both cases representation can be mediated by the selective replication of isomorphic patterns of qualities 'prehended' in the physical world. In Lambert's Post-Classical proposal, this selective representation is mediated by templates that, like Whitehead's propositions (theories) provide a model-theoretic context for interpreting patterns of influence in the external world. Viewing the subject's relation with the world from a process standpoint, Lambert argued that ρbot's selective attention to its world corresponds with the way we do in fact "selectively attend to repeated patterns of influence in the world" (*EMWC:*213).

Thus, in contrast to the rich propositional content embodied in the Manifest Image ontology of natural language, Lambert's system is grounded in a process metaphysical ontology that only represents qualitative composition (synchronic perspective showing the state of a pattern of processes at a given point in time) and temporal

[5] Lambert's approach to belief would appear to be indebted to Daniel Dennett's 'intentional stance'. As with Dennett, he is not interested in the reality of a belief in the realist sense (eg. of an actual state in a person (or robot) that derives its 'belief' content from the connection between some representational token ("it is raining") and an external actuality (rain) or internal referent (feeling of rain)). Rather, Dennett brackets the question of what a 'belief' really consists of and instead adopts pragmatic criteria for defining the notion of belief. It is convenient, he suggests, when predicting the behaviour of people to attribute to them rational motives that are based upon their beliefs, desires and intentionality. To contend that an agent holds beliefs means that it is a system whose behaviour can be predicted by an observer using this intentional strategy. Thus, we can meaningfully talk about a robot's beliefs if ascribing some set of beliefs, desires and intentions to the robot enable us to reliably predict how the robot will behave (Dennett 1991:76-7).

changes to that composition (diachronic representation of transitions between states). Lambert's philosophy of belief rests upon the notion that some self-replicating patterns of influence in his process world are structurally isomorphic to changes that take place in other patterns of influence. In this way processes can be either causally or functionally representative of other processes; such causal and functional selection of patterns of qualities via discriminators and templates is the basis of ρbot's embedded or situated connection with its environment.

As suggested in the epigraph to this chapter, belief for Lambert's robot is an act of classification; to believe is to classify. To classify is to predicate the world, not in terms of language, but rather in terms of intensional types derived from templates that provide an extensional reference to patterns of activity found in the world. That is, the routines or habits defined by templates define how the Post-Classical world is individuated and objectified for ρbot. Thus templates are for Lambert what propositions are for Whitehead, and a valid belief, in both cases has a reference to an indicated object in the external world. Lambert's agents are embedded in the world, and the world itself affirms the validity of a belief through the qualitative patterns of causal influence that it imposes upon the agent.

Before delving further into Lambert's analysis of belief, it should be noted that Lambert's critical remarks concerning propositions are targeting the traditional linguistic notion of a proposition. In Whitehead's philosophy propositions are defined in terms of their metaphysical function as a theory delineating potential pathways for a subject's attainment of a satisfaction.[6] My understanding of Whitehead's propositions is indebted to Henry's (1993) computationally inspired procedural approach to the analysis of Whitehead's metaphysics.[7]

---

[6] Whitehead's metaphysical redefinition of propositions moves away from the traditional declarative focus upon the linguistically signified content of a proposition toward a procedural focus on the way in which a metaphysical proposition (an eternal object predicating a particular logical subject for a percipient subject) acts as a 'lure' (see Ch.Sect.3.3). Traditionally, propositions have been associated with declarative knowledge, "knowing that . . . [the situation described by a proposition, such as 'Socrates is dead']", in contrast to procedural knowledge, such as knowing how to make a cake. For example in Frege's philosophy, the semantics of sense and reference has been understood in declarative terms as the sense of a whole sentence (understood as a 'proposition') while its relation to a referent determines its truth value. Thus, the proposition expressed in declarative terms as:

If chased by an angry bull and you see a tree then climb it.

If expressed in procedural terms the proposition could be broken down into a problem and steps toward a solution. For example:

To escape from an angry bull, find a tree, and then climb it.

In logic programming, both a declarative and procedural representation languages can be used to interpret programs.

For example if, in the language of Prolog, a rule specifies:

```
p :- q, r.
```

The declarative reading is:

p is true if q is true and r is true.

The procedural reading can be stated as

to solve problem p, first solve problem q and then solve problem r,

or alternatively

to satisfy p, first satisfy q and then r.

Henry and Geertsen (1986) and Henry (1993) use both declarative (*logical*) and procedural (*control*) reading of Prolog programs to advance their analysis of Whitehead.

[7] Henry's (1993) use of Prolog programs to illustrate his procedural interpretations of eternal objects and concrescence, has provided an important point of reference for my analysis of Lambert's use of Prolog programs to implement his notion of templates. In the context of this analysis of belief, Henry's

6.2 Post-Classical Belief: Believing as an act of Classification

In terms defined in the previous chapter, the question of belief raises the question of how, given discriminators that detect relevant external qualities and an appropriate

---

work provides some suggestions as to how eternal objects can be related to the elements of Lambert's templates, in their model-theoretic role as mathematical structures.
In brief, Henry argues that Whitehead modelled his understanding of general ideal objectivities after mathematical ones, and called these general objectivities 'eternal objects'. Henry suggests that Whitehead's understanding of eternal objects was based upon a declarative view of meaning, rather than on the "clear procedural understanding" that informed his notion of the "concrescence of actual entities" (Henry 1993:122). Hence Henry's contention that it is more appropriate to interpret eternal objects and propositions from a modern computing perspective rather than view them in terms of a traditional logical or mathematical reading.
In order to illustrate his procedural interpretation of eternal objects Henry builds upon an interpretation of the elements and functioning of a Prolog program (see fn.10 ch.5, also section 1 of chapter 4). He affirms that the essential core to Whitehead's notion of an eternal object is that it is the feeling of a possibility of an actuality that makes no reference to any particular actuality, and thus can be exemplified in actual entities and propositions without being conditioned by that exemplification (akin to Lambert's types). His main aim is to extend Whitehead's concept of abstractive hierarchies to focus on the order among eternal objects of the objective and subjective species. He then relates his procedural notion of a program to an eternal object in two ways; firstly, as formal structures that indicate modes of ingression of eternal objects in concrescence, and secondly, as structures that represent the ideal part of propositions (Henry 1993:93).
As I have argued in chapter three, Whitehead's propositions fulfil a procedural function through an eternal object that predicates an indicated subject to the perceiving or judging subject. Lambert uses discriminators and templates to fulfil this function, both of which can be understood in terms of eternal objects. While the details of Lambert's and Henry's projects are very different there are many interesting connections. For example, I have profited from some of Henry's suggestions regarding the role of different species of eternal objects in the concrescence of an actual entity. For example, Henry groups eternal objects into three classes:

(i) "First, there are constants, such as morose, satisfied, red, etc. that we can use as symbols for felt emotions in our programs. As identified by Whitehead, these are the eternal objects of the subjective species (PR,291). As felt eternal objects, they occupy a datum position of a prehension. They may also, however, occupy a position as a subjective form."

(ii) "Second, there are constants that, for some reason of phenomenological insight, we may think as appropriate as symbols for a datum, but not as symbols for a subjective form." [Eg. The geometrical region abstracted from a spatial region].

(iii) "Third there are atoms and compounds that have generalized subject variables that can refer to any actual entity. . . . These generalized programs can be identified as Whiteheadian eternal objects of the objective species (PR:291)" (Henry 1993:93).

Henry's categories can be related to Lambert's notions of (i) primitive quality nodes, (ii) process nodes, and (iii) templates (these connections will be amplified in my analysis of Lambert's semantics of synchronic and diachronic processes).
Note that Henry downplays the 'eternal' prefix in Whitehead's notion of eternal objects and views programs as the ordered sets of possibilities conditioned by historical processes. I concur; it makes more sense to view these abstract forms of potentiality as emergent historically conditioned properties of events rather than as eternal universals. In a summary statement of his view of eternal objects, Henry suggests that:

> A collection of *atomic_* and *compound_eternal objects* that form a *program* are a grouping of eternal objects that form an abstractive hierarchy within which one can define the concepts of component, derivative, internal, and external relations of eternal objects. The abstractive hierarchies of eternal objects of both the subjective and objective species are of a different sought, normally expressible in set theoretic terms. The order expressed by abstractive hierarchies of eternal objects of eternal objects is conditioned by the long past history of the concrescence of actual entities and their internal and external relationships. This is why the order among some eternal objects does not appear to be eternally fixed (Henry 1993:108).

template structure, Lambert's commonsense machine ρbot arrives at a valid judgment concerning the nature of an external perturbation. How, in functional terms that can be translated into a formal semantic system, does Lambert explain ρbot's capacity to exhibit intentional behavior? For example, if the night silence is suddenly broken by the barking of a guard dog, how does ρbot connect the qualitative fragments that it detects using its discriminators and templates in order to arrive at a judgment about what is taking place?[8] Lambert's answer focuses on the act of classification.

Lambert contends that in functional terms, for ρbot "believing is an act of classification"; thus the functional act of classification provides the link between beliefs and representation and hence the epigraph at the beginning of this chapter: "*The robot believes what the robot classifies*" (*EMWC:*219). Expressed informally, this means that a robotic classifier process ρ classifies a fragment x of process in its world as the qualitative classification x_q.[9] Using the form suggested by Henry and Valenza (2008), a Whiteheadian expression of the same process could be interpreted as meaning that that 'the *Subject* ρ prehends the *Datum* x with a *Feeling of Belief* x_q'.[10] Like Henry and Valenza, Lambert's proposal involves a formal simulation, rather than a replication of the feeling of belief.[11] The connection between Whitehead's and Lambert's notions of belief can be better grasped by digging down into Lambert's analysis of the classification process.

6.21 Agent and Theorist perspectives on the classification process: a Post-Classical solution to the bifurcated semantics of Classical AI

---

[8] In this case a relevant template might be a variant of the simple 'switch-appliance' template structure described in the previous chapter. The template could represent the transition between the quiet dog to the barking dog, and vice versa (using tokens, a list format and including transition classes) as follows:

From state 1 to state 2: [((x u) (y v)) ((x -u) (y -v)); x→x, u→u, y→y, v→v]
From state 2 to state 1: [((x -u) (y -v)) ((x u) (y v)); x→x, u→u, y→y, v→v],

Given that *dog* = x*, barking* = u, *intruder* = y, *present* = v, the above situations could be translated textually as:

[((dog barking) (intruder present)) ((dog not barking) (intruder not present));
Dog→dog, barking→barking, intruder→intruder, present→present]
[((dog not barking) (intruder not present)) ((dog barking) (intruder present));
Dog→dog, barking→barking, intruder→intruder, present→present].

Lambert's analysis of belief asks:
How does ρbot use this particular template to provide a formal context for interpreting the qualitative change in the perturbations that it detects? How is this belief expressed in a formal semantic system? In what sense does ρbot *believe* that the dog started barking because of an intruder?

[9] Lambert's formal definition is: ρ believes $x_q =_{df}$ ρ classifies x as x_q (*EMWC*:219)

[10] Keep in mind that the terms *feeling* and *subject* apply to all monadic levels of existence in Whitehead's philosophy. In a high-level act of classification the basal element would undergo various levels of social transmutation. Following Henry's suggestion the robot's act of believing could be represented in the form of a Prolog rule, where the initial process fragment x of a societal datum (represented as *Datum_* x*n*) undergoes successive transmutations until it attains the final satisfaction x_q represented as *Feeling-of-belief_*x_{q1}:

*Feeling-of-belief_*x_{q1} *(Subject_*ρ*, Datum_* x*1)*if
Subjective-form_* x_{q2}(Subjec t_*ρ*, Datum_* x*2)*and
Subjective-form_*x_{q3}(Subject_*ρ*, Datum_* x*3)*and
. . .
Qualitative-physical-feeling_* x_{qn}(Subject_*ρ*, Datum_* x*n*.
(see fn.10, Ch.5).

[11] See the distinction made in chapter 3, section 5.

For Lambert, believing is classifying; the act of ρ classifying x as x_q means that ρ believes x_q. In order to explain how this act of classification acquires semantic content Lambert adapts the referential and psychological arguments developed by analytic philosophers. Hence he proposes a third person *theorist* perspective, in which beliefs are viewed as "internal Working Memory instantiation of Long-Term Memory representations", and a first person *agent* perspective, in which beliefs are "dynamic, set-like, qualitative, intensional classifications of processes in the world" (*EMWC*:237). Lambert suggests that in some respects these two orientations toward belief can be likened to the positivist outlook of the neurologist in contrast to the phenomenological outlook of the continental philosopher; however, he shows that from a process standpoint the theorist and agent perspectives are complementary and provide a semantic outlook that, while similar to the internalist psychological and externalist referential positions examined in chapter two of this work, manage to overcome the problems associated with these positions.

Framed in the terms of the referential and psychological approaches to intentionality, Lambert's theorist and agent perspective would appear to mirror the 'fatal' metaphysical gap in the Cartesian 'theory of representative perception', as discussed in chapter three with respect to the bifurcated Cartesian world. However, in the terms defined in my analysis of Whitehead's metaphysical system these two standpoints can be reconciled; the third person account of the person in the room and the first person account of the room in my mind simply reflect the public and private sides of every transaction. Every act has its extensional reference to a factual datum and its intensional reference to the qualitative properties disclosed in the subjective form of the concrescing occasion.

According to this analogy, the agent standpoint is not an attempt to feel what is private to the agent, but rather an attempt to give a content to the doxastic form of an agent's belief; to identify the complex eternal object mediating the subject's final grasp of their situation – their 'satisfaction' - whereby the object of their attention is disclosed as a determinate intensional whole (the perspectival *what* that mediates a belief, for example, the qualitative feeling of a tree as a patchwork of dappled green, prehended in the mode of presentational immediacy).[12] That is, in Whiteheadian terms the agent perspective focuses on the eternal objects mediating the agent's perspectival objectification of their datum. The theorist viewpoint is akin to Whitehead's analysis of how this objectification is attained.[13] This analogy can be clarified by examining the semantic applications of Lambert's notions of theorist and agent classification.

[12] Although Lambert is using the language of the philosophy of mind, akin to the 'dual aspect' theorist's advocacy of non-reducible first and third person perspectives, there is a fundamental difference. Lambert's reference to the first person stance involves simulating the structural form or object mediating the subjective orientation (in Husserlian terms, it involves a simulation of the noematic, rather than the noetic dimension of intentionality). Lambert's intensional structures imply no correlated experiential dimension; only a functionally analogous formal equivalent (symbolic representation) of the intensional qualitative forms mediating human experiences. On first and third person accounts in the philosophy of mind, see Velmans (2008, 2009 ch.13).

[13] Whitehead suggested that the 'satisfaction' of an actual entity can be divided up in two ways; a genetic division of the process of concrescence and co-ordinate division of the concrete temporal and spatial region objectified in the datum (*PR*:283-4). In that one focuses on the subjective and the other on the objective ends of the prehensive process it could be assumed that there might be some simple correlation between the perspectives advocated by Lambert. However, there is no ready fit between Whitehead's methodological division and Lambert's first and third person orientations. Neither Lambert's nor Whitehead's orientations are strongly 'phenomenological' in the Husserlian sense;

6.22 The Post-Classical semantics of Belief from a Theorist Perspective

The theorist perspective views beliefs as structures that arise within the working memory of ρbot as a result of *how* ρbot classifies its acquaintance with the extensional object of its belief. Lambert suggests that this can be illustrated in terms of three externally related processes; x is a fragment of process situated in an environment external to ρbot, q is a primitive or functional quality associated with the process x, p_q a classifier process situated in ρbot's long term memory and x_q the belief arrived at in ρbot's working memory. In Whiteheadian terms, x is a social nexus of actual entities, q is an eternal object in the form of a sensa or a combination of sensa that define the characteristic form of the social nexus, and x_q is a product of ρbot's classifier device p_q (discriminator or template) attending to the causal efficacy of the process x. Working memory, long term memory and the object classified are all enduring societies interacting in accordance with their characteristic social habits (as represented in fig. 6.1).

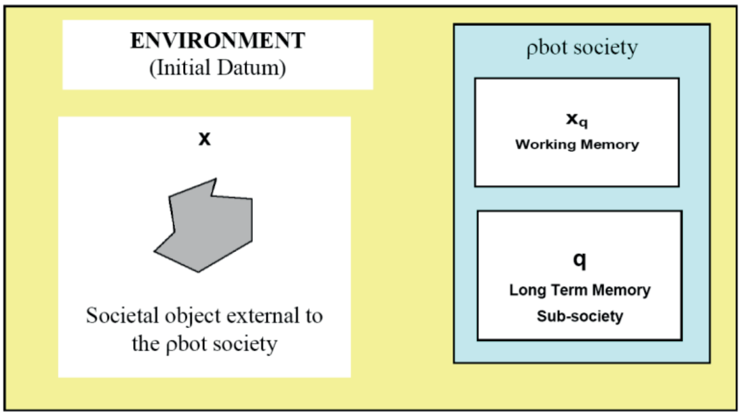


Figure 6.1 Processes in theorist classification (*EMWC*:223, modified)

Thus ρbot selectively attends to qualities q that it detects in causal perturbations arising from process x via the operation of discriminator and template processes p_q. Through these qualities ρbot assembles an internal classification x_q of x in its working memory. Lambert explains the *how* of this process in terms of the role of the long term and working memory processes involved in classifying primitive and functional qualities. These can be illustrated in terms of classifying a primitive quality such as redness and a functional quality such as a switch.

From the theorist perspective, the classification of a primitive quality q such as redness occurs when ρbot's qualitative discriminator device p_q (a fixed red receptor) tuned to that quality red detects qualitative causal influences arising from the datum x (situated in the external environment of ρbot). From the theorist perspective, the operation of discriminator process p_{red} produces the classification of x as x_{red} and this classification constitutes a belief in the redness of x; the theorist perspective views the

however, as explained in chapter three there is an important phenomenological dimension to Whitehead's work. In some respects the first person orientation enters into both of Whitehead's genetic and coordinate modes of analysis. Likewise, both involve a speculative third person dimension. Thus, it is possible to draw analogies between aspects of both of Whitehead's methodological orientations, and Lambert's theorist and agent perspectives, however the correlations are only partial.

belief as a process fragment in ρbot's working memory sub-society produced by the imposition of x into an agent 'mental process' that involves q. The classification of functional qualities follows a similar pattern: from the theorist perspective the focus is on the internal properties of ρbot.

As described in chapter five, functional qualities are defined in the context of other qualities, with superimposition, based on memories stored in ρbot's long term memory forming that basis of the classificatory process. Functional classification takes place via templates that record patterns of past experience and applying copies of those previous experiences to new experiences. Lambert points out that this is essentially a variant of case-based reasoning.[14] The application of the template to the new process takes place in ρbot's working memory. Unlike a primitive classifier device, a template can represent functional processes that involve complex changes over time, such as a musical tune. The template therefore represents states and potential transitions between states. For example, the switch template previously described has two possible state instance structures; hence any classification of a process as a switch involves detecting that the structural composition of the primitive qualitative processes associated with x conforms to a state instance q represented by the quality nodes associated with one of the template's possible states (*EMWC*:220-1). State-instances thus correspond to node-instances, as each node in a template defines a role in relation to other nodes. A process classified as having a functional quality is understood to perform the role of the node instances whose qualities have been detected and matched with the template. For example, if ρbot has a switch template in its long-term memory, then "ρ classifies x as x_{switch} whenever process x causally invokes a node instance having a quality functionally qualified by the node-instances associated with the nodes defining the role performed by a switch in the switch template." The process x_{switch} with its qualitative reference to process x then becomes a belief in ρbot's working memory (*EMWC:*222).

Unlike beliefs in Classical AI, Lambert's Post-Classical beliefs have no reference to natural language linguistic primitives; their reference is solely to simple representational structures that register the patterns of qualitative transitions associated with patterns of influence that ρbot has been programmed to detect. The formal language employed by ρbot is based upon set theoretical representations of patterns of influence within the brick-wall universe of processes that Lambert likens to lego™ block models. According to this analogy, state-instances are like lego-blocks, with each state distinguished solely by its fixed shape and the layout of the connection protrusions and cavities (set node instances and quality node instances) that constrain how blocks can be locked together. These blocks derive their collective functional significance from the tie between ρbot's architecture and the historic capacity of specific block structures (templates) to match qualitative patterns detected in ρbot's environment. However, the types of structures instantiated in assemblages of set and quality noted are universals (primitive quality such as redness or a functional quality such as a switch) and just as Whitehead's universals (eternal objects) can potentially be joined together to form complex structures and put to use in indicating

[14] Case-based reasoning involves solving new problems by adapting solutions that were used to solve old problems. This differs from rule-based reasoning which solves problems by chaining rules of inference together.

and objectifying diverse patterns of influence, so Lambert's lego blocks can be used to model a variety of states.[15]


[16]

Thus, from the theorist perspective, both primitive beliefs like x_{red} and functional beliefs like x_{switch} are processes located in ρbot's working memory that can be explained in terms of *how* ρbot classified the processes that it was acquainted with (either through a discriminator device or by joining lego-blocks together to form an appropriate state-instance match).

6.23 Theorist belief and Functional Role semantics

Thus, according to Lambert, from the theorist standpoint a belief (such as x_q) "is a structure *within* the Working Memory of ρbot which arises from causal imposition of x onto ρbot . . . [and this belief structure] is ultimately a matter of how ρbot's templates *use* the influence of x" (*EMWC*:225). The orientation is similar to Whitehead's theory of concrescence, in so far as it is also concerned with tracking the internal processes arising from the causal imposition of an external process on the inner-workings of an agent. For Whitehead this begins with physical prehension and then proceeds to conceptual prehension and various supplementary processes toward perspectival unification, while for Lambert the sequence begins with the causal influence of a process x on ρbot's discriminators and progresses through a selective unification process mediated by ρbot's store of templates; in both cases the focus is on the practical 'how' of the unification process. Thus, Lambert's theorist perspective places an emphasis upon both a causal reference and a holistic criterion determined by the way ρbot uses its templates. Lambert argues that this provides a basis for regarding the theorist perspective as an "unusual kind of two factor Functional Role Semantics" analogous to the psychological approach to belief employed by analytic philosophers (see Ch.Sect.2.4).[17]

---

[15] Just as Whitehead's eternal objects are patterns of relational potentiality totally abstracted from any concrete instantiation, in a similar sense Lambert's templates transcend their historical origins and acquire a potential that may be satisfied in very different situations. In Lambert's words:

> The value of the state, like the lego block, lies solely with the capabilities of its structure, and of itself, it is representative of nothing other than its compositional structure. But when put to use, the block, in collaboration with other blocks can model many varied functions, for a lego block might on one day function as part of the roof of a house, and on the next day as part of the tail of a helicopter. Such is the intention with state-instances as they dynamically contribute classifications of processes in the world, and templates enrich these classifications further by prescribing recipes for swapping lego blocks (state instances) to reflect the dynamic nature of the processes being classified. This is structurally a very different basis for belief (*EMWC*:222).

[16] Image adapted from Wikipedia: "Lego".

[17] Two factor Functional Role semantics is a variation of Conceptual Role Semantics (CRS). According to Ned Block (1998), CRS explains the meaning of a representation in terms of "the role of that representation in the cognitive life of the agent, e.g. in perception, thought and decision-making". According to Block, CRS extends the 'use' theory of meaning, "according to which the meaning of a word is its use in communication and more generally, in social interaction", but "supplements external use by including the role of a symbol *inside* a computer or a brain" (Block 1998:652). The two factor

Two factor functional role semantics can encounter problems due to an incompatibility between the meaning derived from reference and the meaning derived from how a symbol is used inside a brain or computer. This can be illustrated using Putnam's Twin Earth mind experiment, as discussed previously (ch.2.43). For example, if on Earth's (almost identical) planetary twin, Twin Earth, the fluid substance X_YZ that fulfills the function of H_2O is not in some technical sense 'wet', yet has an *identical use value* for the citizens of both planets, if water is defined purely in terms of use value, then it is no longer true that all water is wet. However, since in referential terms the only true water is that which is found on planet Earth, then it remains true that all water is wet. Thus the two factor theory arrives at a contradictory conclusion: all water is not wet and all water is wet.

According to Lambert, the problem arises because the standard twin factor theory allows water and wet to be content types independent of their use value. His theorist perspective avoids this problem by not being a full two-factor theory of meaning; contrary to the example given, 'water' and 'wet' can have no independent referential content, as reference and use are tied together. Just as Whitehead's prehensional 'bricks' tie together the extensional datum and the qualitative type associated with that process, so also Lambert's set and quality nodes lock meaning and reference together. For ρbot meaning content "is the attribution of a type to an extensional process", with *use* determining the type and *causal reference* determining the extensional process (*EMWC*:226). That is, when a template state instance is matched against an externally imposed pattern of influence, the type (or eternal object in Whiteheadian terms) exemplified in both cases is tied via causal reference (Whitehead's physical feelings of causal efficacy) to the indicated external process. Types (or eternal objects) acquire their meaning or value (become true or false) in terms of the process that is being classified (the indicated logical subject).

I will say no more about Lambert's further analysis of semantic issues associated with the theorist perspective, since the essential point has been made; namely that from the theorist perspective the belief x_q arises within ρbot as a result of the imposition of the causally indicated extensional process x into the mental qualitative classificatory framework q. The result is a semantic stance not unlike the psychological position, but which, because of its process ground, avoids problems such as the Twin Earth paradox.

6.24 Post-Classical Beliefs from an Agent perspective

Lambert suggests that the theorist and agent perspectives are defined by their respective preoccupation with the *how* and *what* of classification. Lambert's *theorist* perspective is analogous to the psychological (internalist) account of belief; the

---

variant of CRS was a response to Putnam's "twin earth" experiment (see chapter 3) and the implication that meaning depends in part "upon something other than the role of a representation in a person's head, namely upon what happens to be in their external environment"; proponents of the two facto version of CRS by arguing that meaning consists of "an internal, 'narrow' aspect of meaning - which is handled by functional roles that are within the body - and an external referential/truth-theoretic aspect of meaning, which might handled by some of the other metaphysical theories of meaning (e.g. a causal one)" (*ibid*:653-5).

classification of x as x_q in the agent's working memory is about *how* ρbot's template uses the influences from x. This contrasts with the agent perspective, which Lambert suggests is analogous to the referential (externalist) account of belief. From the agent perspective, the classification of x as x_q is about *what* has been classified as process x. While the theorist standpoint associated the quality q with its internal classification mechanisms, the agent standpoint associates q with an abstract *type*, instances of which are resident in its memory, and instances of which are also detected in the external environment. This means that from the agent perspective, ρbot is attending to a domain that is populated by abstract entities of the sort that are stored in its memory. According to Lambert, these typed processes are *intensional* entities, since "the one underlying extensional process can be typed in numerous ways" (*EMWC*:223). The contrast between the theorist and agent perspectives are indicated in figure 6.2; however, before discussing that diagram several points require clarification.

Lambert's usage of the term *types* carries with it several meanings. It retains both a reference to the everyday meaning of types, as well as to the more specialized philosophical and mathematical notion of types, and the technical reference derived from computer science.[18] From a philosophical standpoint types are usually regarded as Universals. As Linda Wetzel (2006) has argued, like Universals (and in contrast to Particulars) they are usually viewed as "*having instances, being repeatable, being abstract, being acausal, lacking a spatio-temporal location* and *being predicable of things*" (Wetzel 2006, her italics). I have been informally relating types to Whitehead's eternal objects. As with Whitehead's eternal objects, Lambert's types have no reality outside of their realization in concrete actual entities.[19] When Lambert suggests that ρbot believes in a "world populated by abstract entities", he is not suggesting that the world is composed of abstract Platonic Ideas. Rather, types (such as redness or switch) are "themselves extensional processes located in ρbot's memory" and belief in x_q is "an association between a process x in the world and processes, including q, within ρbot's memory" (*EMWC*:224). Just as Whitehead's higher level agent's objectify their world in terms of propositions (such that the percipient subject objectifies - predicates - an indicated logical subject in terms of an eternal object), Lambert's agent predicates its world of extensional processes as particular typed processes. Thus the typed process (the belief) x_q is an intensional entity made up of "two associated elements"; "(i) extensional processes; and (ii) process types" (*EMWC*:224). Yet like Whitehead's higher level societies, ρbot can represent

[18] The history of *type theory* goes back to Russell's attempt to solve paradoxes associated with set theory. In computing, assigning types means assigning a meaning to a sequence of bits; objects and classes may also be typed. A *typed system* specifies how typed programs behave. Lambert (*EMWC*:251) quotes a passage from Cardelli and Wegner (1985) that summarizes his view of types:

> As soon as we start working in an untyped universe, we begin to organize it in different ways for different purposes. Types arise informally in any domain to categorize objects according to their usage and behaviour. The classification of objects in terms of the purposes for which they are used eventually results in a more or less well-defined type system. Types arise naturally, even starting from untyped universes (Cardelli & Wegner 1985:473).

[19] Whitehead's distinction between physical and conceptual prehensions could be interpreted as licensing some form of substance dualism within each monad, with physical and conceptual prehensions indicating two types of actuality. However, Whitehead is specific in stating that eternal objects have no actuality. They only attain actuality through realization in a physical process. Likewise, Lambert's *types* are patterns of potentiality that are implicit in the functional role of his discriminators. Though physically instantiated in discriminators and templates, like eternal objects they are *functionally* a potential actuality, as they can be joined together to represent various entities that may or may not exist.

processes that may of may not exist, and therefore the truth of x_q is conditional on the existence of processes of the requisite type.

Thus, in semantic terms, the association between a process and a type is *truth conditional*. That is, if ρbot is to believe x_q then ρbot believes that x is an occurrence of type q (in formal terms, x : q).[20] What this means is that a pattern of influence is exhibited in the world that ρbot classifies as q; hence, repeated patterns of influence "effectively define process types, and processes having those patterns are occurrences of those process types" (*EMWC*:224). Processes can be represented in terms of the formal calculus introduced to describe the brick-wall universe, and types defined formally in terms of the model theoretic language introduced to describe representations. Thus the definition x : q specifies an existential semantics for processes, whereby the process must exhibit a specific pattern of influence if it is to satisfy the conditions for being a process of that type q. When ρbot's classifiers represent a particular pattern of influence, they can be viewed as an abstract representation of that type. Hence, with the relevant classifiers ρbot can represent a pattern of influence that signifies a tree on the moon, even though ρbot might not be able to causally validate that proposition. Provided ρbot has appropriate access and the appropriate classifiers it is able to validate whether a pattern of influence really is a tree on the moon.

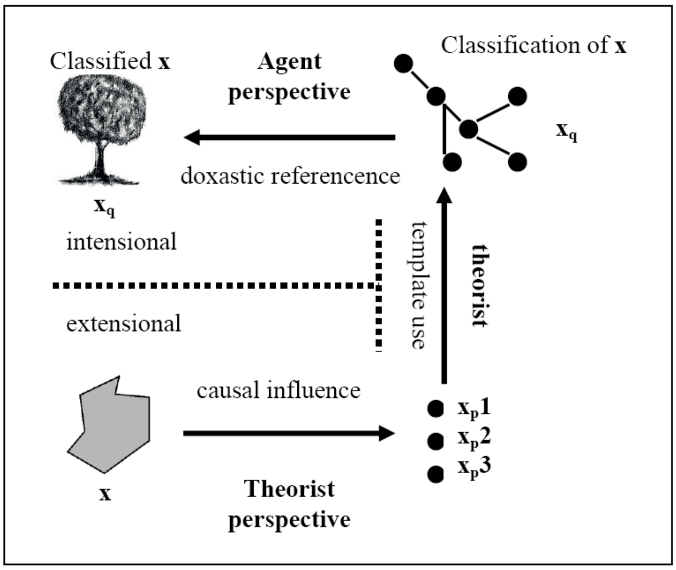


Fig. 6.2 Agent and theorist perspectives in the classification process (*EMWC*:223).

In model theoretic terms, to say that x is a process of type q means that ρbot has been able to detect and represent with its classifiers a pattern of influence q that it has repeatedly encountered in the world. This set-theoretic, intensional structure then serves as a model for interpreting subsequent patterns of influence, which it can classify as real (the interpretation is a model of the pattern detected, formally validating that the belief is real); likewise through this interpretation it is able to formally determine that its belief is true (it is therefore meaningful to speak of abstract

[20] The symbol "**:**" is the computer science notation for type. Where it is necessary to indicate the type of classifying mechanisms **s** (discriminators for primitive qualities and templates for functional qualities) the notation used is **x :$_s$ q**.

facts). Lambert's agent perspective thus focuses on the *what* of classification, such that x_q is an abstract intensional structure that may or may not be real, which ρbot has formulated as a belief on the basis of a reasoning process. The belief x_q thus has a doxastic[21] reference to the world, as represented in figure 6.2.

From the theorist perspective ρbot selectively attends to a causal influence x through the qualities it detects in x. The dynamic structure of the qualities that its classifiers detect (x_p1, x_p2, x_p3) match a template q in ρbot's long term memory, provoking ρbot to view x as a state instance of the classification type x_q. From the agent perspective, x_q is the abstraction q imposed on to process x; the agent believes that its world is populated by trees of the form x_q. This agent perspective involves intensional classification, with ρbot imposing its belief framework for tree onto x.[22] When ρbot classifies **x** as a tree it first detects qualities that match a tree template in its long-term memory, and this match becomes a state-instance of the tree in its working memory. Since the one tree process can be typed in numerous ways, the classification of the tree process x qualitatively as a state instance x_q is an intensional entity, while the causal influence of x on x_p1, x_p2, x_p3 constitute an extensional reference.[23]

Adopting the agent view of belief means that the notion of intension has a different meaning than is the case in classical logic. Rather than situation types being defined in terms of intensional sets defined from more primitive intension sets that are obtained from a plethora of linguistic primitives, the process types are intensional sets obtained from extensional entities. Thus, although both the situation theorist utilizing classical AI and the process theorist drawing upon post-classical AI utilize an abstract intensional outlook, "only the process theorist requires that it be grounded in selected patterns of influence of the extensional world" (*EMWC:*228-9). By combining an intensional sense and an extensional reference in this way Lambert not only sidesteps Putnam's traditional 'twin earth' style of problem (meaning as psychological use) but he also avoids Frege's 'morning star'/'evening star' problem (meaning as extensional reference).[24] How Lambert's theory achieves this semantic clarification can be better appreciated by contrasting the referential orientation of his agent stance with that of Situation Semantics.

6.25 Agent belief and Situation Semantics

I have suggested that the theorist and agent perspectives refer to two contrasting orientations to classifying fragments of process. The contrasting standpoints are analogous to the psychological and referential perspectives of the same perceptual act,

[21] In this case I assume that *doxastic reference* simply means a reference to the object of a belief.

[22] An intensional reference is made up of an extensional process **x** and a process type **q** of $\mathbf{x_q}$ (where **q** is a quality representation of either a primitive quality type, like the colour red, or a functional quality type, such as a light switch). Process types are themselves extensional processes resident in pbot's memory, such that belief in $\mathbf{x_q}$ is the association between a process **x** in the external world and a process **q** in the ρbot's memory. When ρbot believes $\mathbf{x_q}$ it believes that **x** is an occurrence of type **q**, such that repeated patterns of influence "effectively define process types, and processes having those patterns are occurrences of those process types" (*EMWC:*224).

[23] For example, the causal influence of a tree could be classified in various ways as outlined in chapter section 5.21.

[24] For a recent elaboration of these arguments, see Lambert (2009:11).

referring in the first instance to a chain of causal influences that terminate with a classification of content identified with a process in the agent's working memory, and in the second case to the content (type) identified with an external process by an act of doxastic reference. Just as the theorist perspective has a close relationship to the two factor Functional Role semantics, the agent perspective is allied to Situation Semantics.[25] Like Barwise and Perry's (1983) Situation Semantics, Lambert's agent perspective presents an "abstract, intensional, partial classification of the actual world in which the classifications can be either real of fictional" (*EMWC*:228). However, his Agent process semantics differs in terms of its process primitives, the intentionality associated with process, and the subjectivity that is associated with the agent standpoint. I will briefly address these three distinctions.

Firstly, Lambert rejects the Situation Theorists linguistically constituted Manifest Image[26] view of the world in favor of a world defined in terms Patterns of Influence. Lambert follows Barwise and Perry (1983) in using model theoretic structures to determine truth claims, however unlike Situation Semantics, which attempts to extract an ontology from Natural Language usage, Lambert builds upon a process ontology of spatio-temporal patterns of causal influence. Thus, rather than attempt to define situation types in terms of intensional sets obtained from linguistic primitives Lambert's defines process types in terms of intensional sets obtained from extensional entities. As Lambert notes, although his approach also promotes an abstract intensional outlook on the world, "only the process theorist requires that it be grounded in selected patterns of influence of the extensional world" (*EMWC*:229).

In terms of intentionality, the distinction between Lambert's semantics and Barwise and Perry's Situation Semantics again arises from the respective linguistic as opposed to the process grounds of each theory. In analyzing Classical AI's approach to intentionality, I discussed various strategies employed by analytic philosophers to explain the 'aboutness' of intentional statements. A number of those strategies built upon Tarski's semantic concept of truth, and used model theoretic techniques to show how mathematical structures could be used to assign truth values to formal sentences. However, the problem still remained as to how the world or partial world captured by the mathematical structure was related to the real world. Lambert's Post-Classical theory provides a process grounded solution to this problem. Rather than focus on abstract situations, Lambert situates his robot in the real world, and has it interpret events within the context of ongoing activity.

---

[25] The aim of Barwise and Perry' (1983) Situation Semantics was to understand commonsense use of Natural Language by interpreting utterances in terms of real situations. Situation theory attempts to specify these situational contexts in terms of a mathematical ontology abstracted "from analyses of natural language use" (Devlin 2006:1). The core idea in situation theory is that utterances convey information ('infons') that are not true of false in isolation, but as with Whitehead's theory of propositions, are true or false in terms of an indicated real situation. Thus, for a situation, s, and an infon σ,

$$s \vDash \sigma$$

meaning "s supports σ", or in other words that the infon σ is made factual by the situation s. Devlin notes that the actuality $s \vDash \sigma$ is referred to as a proposition. Types are a basic way of carving up situations. Situation theorists employ a similar notation as Lambert; given an object, x, and a type, T, they write $x : T$ to indicate that the object x is *of* type T. For a brief introduction, see Devlin (2006).

[26] 'Manifest Image' refers to the notion that the ontology implicit in the commonsense usage of Natural Language maps onto the world itself. This notion can be regarded as the enduring core of the philosophical ontology initiated by Aristotle. See ch.2, fn.16.

Barwise and Perry follow a similar strategy, however in their Situation Semantics it is a property of a sentence that determines a context, while for Lambert the context is defined by functional qualities associated with a particular template. If a process is classified as a particular type, it means that it has been situated in the context of various functional roles associated with the template. The situation with Lambert's template based classifications is not unlike Whitehead's habit based prehensive unifications in the mode of presentational immediacy, in that in both cases "the causal referent of the type nodes is the process [in the datum]", and that same process is the "doxastic referent of the type node" as the intensionally classified process object (*EMWC*:230). As with Whitehead's habit based transmutations, Lambert's template based classifications will carry different connotations depending upon the way lego-blocks have been used to type the situation.

If, for example, a ρbot template classifies the tree as a type, and thereby associates a functional quality with the tree, the causal referent of the type node is a 'tree process', and the doxastic referent of the type node is that same process as the intensionally classified tree. If the functional qualities associated with the tree are defined in terms of a system of other inter-dependently defined qualities, associated with tree states, such as dry branches, burning wood, tree-cutting, timber-milling, and so on, then this may lead to ρbot making existential and dispositional presumptions about its situation. For example, depending upon the other templates in its memory the robot might classify the tree as a source of furniture timber, a place of refuge from a charging bull, a source of fire wood, an obstacle to avoid when driving, and so on, and these classifications might influence ρbot's future behavior.

As in the case of human intentional acts, such dispositions are fallible; the referent processes of ρbot's classificatory act may not exist (in actuality the wood may not be suitable for furniture timber), it may also refer to an unknown future state (the possibility that the tree will die and the wood become burnable). Thus, Lambert's classification process rests upon associating a functional quality with the external event. The resultant classification is a holistic abstract association of a type quality with a process, in the context of a "system of other inter-dependently defined qualities through the template", and this, Lambert argues, "leads to existential and dispositional presumptions on the part of pbot, presumptions which influence his choice of behaviour" (*EMWC*:230).

Finally, Lambert argues that his process version of Situation Semantics is based upon a first-person agent perspective rather than the third person perspective of Natural Language Situation Semantics. He relates his stance in this regard to Continental Philosophy and also to Thomas Nagel's (1974) challenge to think about what it means to experience the world as a bat. Nagel suggested that subjective phenomenological facts are in a sense objective, in that although there is a subjective element particular to each individual, there is a *shared type of experience* that can be objectified and understood by others. Lambert claims that the agent perspective is not about experiencing *being* a machine engaged in a classification process, but seeking an appreciation of the *type of experience* a machine has. For example, from a theorist perspective ρbot's act of classifying a switch is viewed as an extensional description of a pattern of influences, but from an agent perspective ρbot's act should be analysed in terms of the dynamic structure of the intensional sets which preserve the robot's point of view toward the classified pattern of influence. Lambert contends that ρbot "experiences a typed *intensional* world which derives from the manner in which it uses past experiences", and that describing the dynamic structure of these sets constitutes a *phenomenology* for ρbot that "preserves ρbot's point of view toward the pattern of influence" (*EMWC*:231).

Lambert's suggestion is a contentious claim in that it seems to fly in the face of Nagel's requirement that the possibility of an objective account of experience is dependent upon a minimal subjective similarity between the types of experience in question (a capacity to stand in the other's shoes).[27] I will not enter into this question as it would seem to be peripheral to Lambert's project and to my Whiteheadian interpretation of that project (I have already explained in chapter three as to why Turing-based approaches to computation cannot replicate human intelligence as argued by advocates of Strong AI). I do not regard Lambert's talk of preserving "ρbot's point of view" as a first step toward a machine ethnography; rather, I see it as short-hand way for speaking about the functioning of the machine in a way that that focuses attention on the real internal structure of intensionally typed 'beliefs' that shape its behavior. It does not entail the anthropocentric projection upon a machine of the qualitative experience of an intensionally typed world, such as a human's experience.

Thus, in summary, Lambert arrives at a semi-realist account of belief from two different standpoints. From my process perspective the agent and theorist perspectives can be related to aspects of Whitehead's theory of extension and his theory of concrescence – with an important qualification; namely, the inclusion of intensional elements into the theory of extension as per the strategy outlined by Lambert.[28] According to this analogy, the *agent* or externalist (referential) standpoint is akin to the subject's final grasp of their situation – their 'satisfaction' so to speak – whereby they prehend the objective form of their situation as an intensional whole (for example, the qualitative feeling of a tree is prehended in the mode of presentational

[27] Nagel (1974) outlines his position in his famous article "What is it Like to Be a Bat?" Elsewhere Lambert has suggested that his aim is to create a machine with *some* functional capacities analogous to those that underpin human commonsense (just as, for example, early aviators attempted to create machines that embodied *some* functional capacities analogous to those that facilitate the flight of birds). Although the nature and role of subjective experience is a valid question, in the context of Lambert's project it has no more bearing than, for example, the question of "what is it like to be a metal aircraft?" As I have argued elsewhere, there are fundamental distinctions between the processes that underpin human and machine 'computation', namely, the organization of the human body maximizes the intensity and diversity of our acquaintance with those energetic processes that Whitehead classified as 'prehended feelings', while computers are designed to deliberately minimize both the intensity and diversity of such energetic processes, such that the only significant 'feeling prehensions' in a computer chip are binary contrasts operating in an environment carefully designed to ensure uninterrupted linear chains of physical causality. As in the case of the aircraft-bird analogy, this does not mean that machines cannot be engineered to functionally replicate some of the processes that give rise to human capabilities (see Henry and Valenza, 2008; Bickhard 2009).

[28] As outlined in Chapter three, Whitehead regarded formal logic as a domain stripped of intensional reference, and hence at a superficial level my association of the agent (referential) perspective with the theory of extension and the theorist (psychological) perspective would appear to run counter to Whitehead's focus on the (extensional) public and (intensional) private dimensions of process. As a model of Whiteheadian concrescence the computational model has a limited validity, in that it is using a linear temporal process to represent a non-linear non-temporal process. However, as a model, it is abstracting from the presumed reality in order to represent in a simplified and idealized fashion aspects of that reality.

immediacy). This, in Whiteheadian terms, is the perspectival *what* that is objectifying the public domain from the subject-agent's standpoint. The *theorist* or internalist (psychological) viewpoint is akin to aspects of Whitehead's perspective on the process of concrescence, whereby the mechanisms that mediate the prehensive unification of the initial datum are outlined (*how* the vectors objectifying extensional processes via raw sensa, as patches of green and brown for example, are transmuted through a unifying procedure into the qualitative feeling of a tree). This analogy can be illustrated in terms of the diagram presented in chapter three of this thesis (fig. 3.4); a comparison between figures 3.4 and 6.2 provides an indication of the structural analogies that I have argued exist between Whitehead's and Lambert's theories of belief.

6.3 Lambert's Compliant Realism

In terms of the broader philosophical landscape, Lambert suggests that his approach to ρbot's beliefs constitutes a semi-realism which he labels 'compliant realism'. Compliant realism is Lambert's Post-Classical philosophy of belief. It has two main tenets; it asserts on the one hand that our commonsense world is structured by the forms we impose upon it and, on the other hand, it asserts that this commonsense world is at least partially disclosing what the world is really like.[29] That is, in Whiteheadian terms, the first tenet asserts that our habits are imposing forms upon the world, while the second tenet asserts the qualitative forms disclosed through those habits are prehensions of patterns of feeling and functional relations that are a real part of the world. Lambert's understanding of classification is central to affirming both of these tenets, for:

> Classification explains how a doxastic, qualitative, intensional view of the world can be understood through a methodological and ontological foundation of processes. With this we have a conception of a process *reality*, and a conception of how a fragment of that reality might methodologically organize process *representations* so as to *believe* that fragments in that process reality have qualitative intensional structure (*EMWC:*233).

In relating his philosophical position to the dominant ontological, epistemological and semantic positions identified by Dummett (1994), Lambert contrasts the differing chains of dependence between <u>what there is</u>, <u>what is represented</u> and <u>what is known</u> that characterize each position. He notes that his schema resembles Dummett's ontological class of philosophical system, in which what is known depends upon what is represented which in turn depends upon what there is. However, he contrasts his process version of this position with the dominant object oriented ontology, noting that:

---

[29] Lambert summarizes the two main tenets of his compliant realism as:

> <u>Doxatic tenet</u>: the commonsense world is a world *structured* by the forms we impose in order to understand it, and as a consequence, our understanding of the world is limited by the forms we bring to bear [including our sensory architecture], and
> <u>Methodological tenet</u>: the commonsense world nonetheless, at least partially exposes what the world is really like, because the forms adapt to comply with patterns of influence which are (in a semi-realist sense) in the world to be selected and which the world imposes on the classifying process [the ontological foundation of what is represented is evident in the recurrent patterns of influence that are classified, whether or not I possess experiential knowledge of these patterns – cf. the semantics of success below, section 3.4] (*EMWC:*235).

> The ontological basis for object orientation is generally, and certainly in Aristotle's case, intended as a rigid and complete affirmation of the way the world actually is. The process theorist is more conciliatory. Our objective understanding is of a reality revealed by qualities, but where [depending on the appraising apparatus of ρbot] these qualities may exist at levels of individuation beyond our experience and/or in realms adjacent to our experience, and that each of these qualities exposes certain process fragments comprising reality (*EMWC:*233-4).

Lambert states that the *principia intellectus* of Compliant Realism is to "seek an *equilibrium* for understanding, not a *foundation*" (*EMWC:*235). Hence, instead of the linear foundational chains of dependence that Lambert associates with Dummett's Ages of Ontology, Epistemology and Semantics, his Compliant Realism envisages a circular dependence between existence, representation and belief (as suggested by fig. 6.3). Thus, instead of sure foundations in either what there is, what is believed or what is represented, Lambert seeks only a functional equilibrium between these major components of understanding. Beliefs are confirmed when they successfully guide actions. He expresses this pragmatic standpoint as follows:

> Understanding is the product of a non-well founded success theme in which what is represented is charged to maintain a successful equilibrium between what there is and what is believed. On the basis of successful compliance, we recover patterns of the world at various partial levels of abstraction, which are than absorbed for re-use in the doxastic framework. This is what we do, implicitly organized upon the premise that patterns of influence in reality will repeat themselves, that nature will not suddenly change its course (*EMWC:*235).

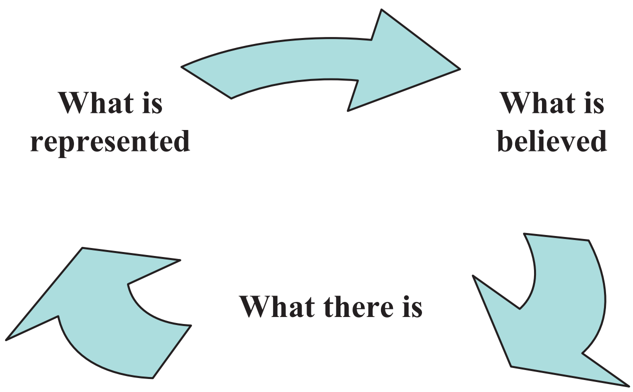


Figure 6.3 Non-well foundedness as successful maintenance of an equilibrium (*EMWC*:235)

Thus, belief in Lambert's Compliant Realism involves a balance between the causal influences of existence, and the selective constraint of representation. Unlike Kant's unidirectional foundational apriori concepts, Lambert's discriminators and templates can be changed in response to patterns of influence from the world. In either case Compliant Realism views knowledge from a practical standpoint, as a function of interaction between existence, representation and belief in response to purposes motivating an agent. Defining and fine-tuning this purpose is crucial, as otherwise equilibrium between patterns of causal influence and pattern classification could produce an endless quantity of beliefs characterized by sheer triviality. Thus a crucial issue in classification is that of ρbot's goals. This question will be considered further in the context of Lambert's Post-Classical approach to mind and information fusion.

6.4 A Post-Classical Extensional Semantics of Belief

In the previous section I discussed Lambert's semi-realist philosophy of belief and argued that it entailed two complementary perspectives. From the theorist's standpoint beliefs were viewed as instantiations in internal Working Memory of Long-Term memory templates, while from the agent's standpoint these same beliefs were viewed as 'dynamic, set-like, qualitative, intensional classifications of the world'. Instead of a semantics of natural language, in which the sense and reference of words mediate the intensional and extensional semantics, the dynamic structure of Lambert's templates mediate the processes through which ρbot arrives at a belief (i.e. makes a qualitative classification and specifies the projected external processes that must correspond to the qualitative states if the classification is to be true). Hence an explanation of belief requires both an extensional and intensional semantics since, in Lambert's words:

> Extensional semantics reports what ρbot is classifying, while the intensional semantics derives from what ρbot is classifying it as. The intensional semantics features beliefs from ρbot's point of view. The extensional semantics defines what must be the case for ρbot's beliefs to be true (*EMWC*:237).

In *EMWC* Lambert develops an extensional semantics for Post-Classscial AI. both the synchronic and diachronic dimensions of an extensional semantics. An extensional semantics must be able to make the connection between the pattern of influences in the world x and a primitive or functional quality q (related to a classifier process performed by ρbot); this act of ρ classifying x as x_q means that ρ believes x_q. Lambert maps out a technical solution to this problem, firstly with respect to synchronic and then with respect to diachronic processes. Together "they identify a belief as extensionally being about a pattern of influence expressed through the [formal Post-Classical] theory of processes" (*EMWC*:237).

I will not provide a detailed analysis of these sections; however, it is important not to anthropomorphize notions such as 'belief', and hence it is important to be able appreciate what ρbot's capacity to arrive at 'meaningful' classifications signifies. Therefore I will give some indication of Lambert's formal system of semantics. Essentially, ρbot's capacity to connect x and q is based upon a semantics of the graph structures embodied in templates. A template may represent a number of states, along with qualitative contexts for identifying particular states and transitions between states. At any instance, a pattern of influence in the world may correspond to one of these states and subsequent instances may comply with the types of potential transitions specified by the template structure. How do we connect the graph structure of a template with a pattern of action in the world? As this work is oriented toward those who, like me have a background in the humanities rather than computer science, I will firstly provide a brief overview of what is meant by a formal semantics of graphs.

6.41 A brief introduction to the formal semantics of graphs[30]

[30] This introduction is based primarily on the first two chapters of Hedman (2006).

In the study of natural language, semantics generally concerns how we assign meaning to words or symbols while syntax refers to the rules or general principles that guide how words or symbols are formed into sentences or formulas. Semantics and syntax have a similar meaning in formal logic, except that the resultant sentences must be capable of being expressed and interpreted within a mechanical calculus. That is, there must be a mechanical procedure that specifies the grammar of the language (enabling the mechanical construction of well formed sentences), and similarly, their must be a mechanical procedure for interpreting or giving meaning to those sentences. In order to achieve this goal, Lambert's formal semantics builds upon the syntax previously used to specify his formal theory of processes. Whether the structure of a template is expressed in first order logic, or represented graphically, the same semantic principles can be applied.

All formal semantic systems share a common form; hence the general principles of Lambert's semantics can be illustrated with reference to the semantics of propositional logic and first order logic. Additional qualifications will show how these general principles are adapted to the specific graphical syntax that Lambert uses to specify the underlying universe of processes represented by his system.

The form that is common to all formal semantic systems involves a general procedure for mechanically interpreting the well formed sentences of a logic. Formal semantics is a means for evaluating the truth or falsehood of a well formed sentence. Semantics rests upon symbols and rules specified by a syntax, together with an interpretation of the symbols and a systematic procedure for generating the truth value of a sentence. The nature of this close relationship between syntax and semantics can be illustrated using propositional logic.

As discussed in the first chapter of this work, every type of logic specifies a language made up of primitive 'words' that must be interpreted. For example, in propositional logic, the symbols $\neg$, $\wedge$, $\vee$, $\rightarrow$ and $\leftrightarrow$ are interpreted as 'not', 'and', 'or', 'implies' and 'if and only if'. These words can be combined together with propositions (atomic formulas represented by capital letters, such as *A* and *B*), to form formulas. Note that the content of the propositions is not relevant to formal semantics; content is stripped away leaving only the token signifying an unspecified atomic content. The grammar of the language (syntax) consists of rules and conventions regarding how these atomic formulas can be combined into larger formulas. A large formula may contain many sub-formulas.

The semantics of propositional logic are defined in terms of a rule specifying which sub-formulas in a sentence are considered first together with truth tables for interpreting these formulas. Formal semantics is not concerned with the content of the propositions; it is not an epistemological quest to determine the truth of a specific proposition such as whether Socrates really is dead. Rather, it concerns the general implication of truth or falsehood for a sentence given the truth or falsehood of its constituent propositions. For example, on purely logical grounds, if the formula $(A \wedge B)$ is true then the propositions (atomic formulas) represented by *A* and *B* must both be true. Formal implications such as this are the basis of truth tables. Thus the rules of propositional logic can be expressed in truth tables which assign truth values to atomic formulas (propositions) and show the consequence for formulas.

An *assignment* assigns truth values to each atomic formula in a formula. As a simple example, consider the truth table for the formula $\neg A$. Such a truth table shows the consequence for the formula $\neg A$ of assigning the values false (0) and true (1) to the atomic formula *A*. The resultant truth table for $\neg A$ can be represented thus:

A	$\neg A$
0	1
1	0

Read discursively, the table states that when *A* is false then $\neg A$ is true, and when *A* is true, then $\neg A$ is false. Such truth tables, when used in conjunction with other rules about which sub-formulas should be considered first, provide a key to interpreting more complex formulas. Such rules permit the mechanical interpretation of complex conjunctions of formulas and are the basis of any formal semantic system. However, the expressive power of such systems of formal representation is greatly extended in first-order logic.

A chief aim of Lambert's semantics of the synchronic and diachronic dimensions of belief is to show that his templates can be expressed in first-order logic. Semantics in first-order logic is similar in principal to that of propositional logic; however the expanded vocabulary of the language has some important consequences. In addition to including quantifiers, the basic atomic formulas in first order logic are made up of terms which can include variables, constants and functions, combined in accordance with specific rules. A sentence in first order logic is one that has no free variables (non-quantified variables). Unlike propositional logic, truth is only defined for sentences, as it makes no sense to ask whether a formula containing a free variable is true since the variable can represent any value (whether *x* is fat depends upon who *x* is). Thus, whether a first-order logic formula is true depends upon first eliminating any free variables (quantifying any free variable so as to create a sentence) and then assigning meaning to the sentence by specifying a context. Quantification means binding a variable to existential or universal quantifiers ($\exists x$ "there exists *x* such that", $\forall x$ "for all *x*"). Context is specified by a *mathematical structure M*, which is formally defined in terms of an *underlying set U*, together with an interpretation of a *vocabulary V* (made up of a set of function, relation and constant symbols) that assigns meanings from the underlying set (or Universe) to the vocabulary.[31]

Mathematical structures play an analogous role in first-order logic to assignments in propositional logic. Different universes and interpretations of vocabularies define different structures, and these structures provide the context for interpreting a sentence. Thus, each well formed sentence specified in the language of a particular structure is necessarily true or false. Hence, any mathematical structure *M* which, when used to assign meanings to the vocabulary of a sentence φ, results in a true

[31] Any mathematical structure will consist of a non-empty set and an interpretation of a vocabulary. Shawn Hedman provides the following example. The structure $M = (\mathbb{Z}|f,R,c)$ is completely described by the following interpretation; the underlying set *U* of a structure *M* is the set of integers $\mathbb{Z} = \{\ldots -3,-2,-1,0,1,2,3,\ldots\}$, and the vocabulary *V* is made up of a set of function, relation and constant symbols $\{f,R,c\}$, where f is interpreted as a unary function $f(x)=x^2$, *R* is interpreted as a relation, $R(x,y)$ such that $x < y$, and *c* a constant, is assigned an element from *U* such that $c = 3$ (Hedman 2006:59).

sentence, is said to *model* that sentence ($M \models \varphi$, meaning that sentence φ is true of structure M). A sentence can be true in some structures and not in others; if a sentence holds for some structure then it is *satisfiable* and if not then it is *unsatisfiable.*

This approach to semantics can be easily applied to mathematical graph theory. Graphs are a class of mathematical structures that can be translated into first-order logic and are intuitively easy to grasp. A graph is made up of a set of nodes (*vertices*) joined by lines (*edges*), such that every edge starts and ends at a vertex. As a structure, the underlying set of a graph is the set of vertices, and the vocabulary is made up of a single binary relation (R, the edge relation). Take, for example, the graph of G, made up of two nodes a and b and a connecting edge ab, represented graphically in figure 6.4.

Figure 6.4 Simple Graph *G*

As a structure this graph is fully defined by the underlying set made up of the nodes, together with the vocabulary defined by the single edge. The resultant graph structure G, interprets R as an edge relation joining (a, b), such that $G \models R(a, b)$, meaning that it is true that the graph has an edge that joins a and b. This graph models the first-order logic sentences φ and ψ,

$\forall x\ \forall y(R(x, y) \leftrightarrow R(y, x))$ . . . [φ]

$\forall x \exists y R(x, y)$ . . . [ψ]

meaning in the case of the first sentence (φ), that for G it is true that R is symmetric ($G \models \varphi$), and in the case of the second sentence (ψ), that for G it is true that each node is adjacent to another node ($G \models \psi$).[32]

Hedman defines *model theory* as "the branch of logic concerned with the interplay between mathematical structures and sentences of a formal language" (Hedman 2006:89). This model theoretic approach to truth is an extension of Tarski's semantic notion of truth. Model theoretical semantics prove that in first-order logic, any structure M, determines a set of first-order sentences φ_M called the *theory of M*, denoted *Th*(M).[33] This introduction should provide sufficient background to appreciate the form of Lambert's semantics; that is, to understand the general strategy that Lambert is following when he attempts to specify "what must be the case for ρbot's beliefs to be true" (*EMWC*:237). This mathematical background should also be

[32] That is, if the structure defined by the graph G models the sentences φ and ψ, then it is true in the world G that "for all x and for all y, x is related to y if and only if y is related to x", and likewise, it is true in G that "for all x there exists some y, such that x is related to y."

[33] The converse is also true, that any set of first-order sentences Γ determines a class of structures Mod(Γ). A theory is a consistent set of sentences, where consistency means that no contradiction can be derived from those sentences. The theory of a structure provides an insight into the structure, as the sentences that make up the theory describe the structure. Conversely, understanding the models of a theory (the class of structures that model a theory) lends insight into the theory. For a more thorough introduction to these issues, including further detail concerning the application of semantics to formal logic, including graphs, first-order logic, and the interplay between structures and sentences in these formal languages (i.e., model theory) see Hedman (2006), especially chapters 1 & 2.

kept in mind when considering Whitehead's notion of propositions (theories) and of the way that contexts determine their satisfaction.

6.42 The Semantics of the Synchronic dimension of Post-Classical Beliefs

As described in the previous chapter, enduring objects in the universe can be represented as a succession of temporal states, each of which embodies a pattern of influence that is passed on to a successor state. In Whiteheadian terms the nexus of actual entities in one temporal moment in the life of a society passes on its characteristic form to a successor nexus. At any moment, a member of that society prehends and objectifies the state of that society in its datum. This prehensive process mediates the constant transition of the society from state to state, as represented by Lambert by the constant addition of brick-wall columns to the extensive continuum of process events. The task of the semantics of synchronic processes is to specify how ρbot forms a belief about a state (i.e., classifies an unchanging pattern) in the brick wall universe. The key to this classification process is combining patterns of qualities causally arising in the world with patterns of potentiality suggested by templates stored in ρbot's memory.

Together Lambert's extensional semantics of states and transitions between states connects his theory of processes Th_{Ω} with his theory of representation Th_{ρ}. This connection involves establishing a match between the qualitative patterns detected in the world and the patterns of potential qualitative and set state transitions stored as templates in ρbot's long term memory. This means being able to represent the composition of potentially very complex structures, defined diachronically in terms of potential transitions between states, and synchronically in terms of the compositional structure of a particular state, defined in term of set nodes and corresponding quality nodes.

In Whiteheadian terms, representing the synchronic dimension of belief can be likened to the task of detailing the compositional structure of the datum at a particular moment in time; this structure is disclosed by qualitative forms subjectively prehended (quality nodes) and their indicated reference in the datum (set nodes). Whitehead's concrescing organism supplements this initial datum, drawing upon habits stored in the organism's bodily memory to selectively transmute and unify this qualitative data in order to arrive at a determinate perspectival objectification of its world. Lambert illustrates the basic principles of this process using the clock face example introduced in chapter five.

As detailed in my discussion of Lambert's philosophy of representation, his approach to representing the synchronic dimension of state structures entails notation for representing *composition*, *qualities* and *absence*. Using the syntax described in the chapter on representation he specifies a unary predicate logic. To see how Lambert builds upon his syntax for representing the composition of states, to create a formal semantics, consider the clock previously described in graphical notation (see fig. 5.3). Let the set nodes correspond to fragments of the clock process in the real world, represented in the graph below (fig. 6.5) as a state structure composed of uniquely numbered nodes, representing the clock, its hands and its face.

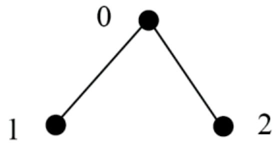


Figure 6.5 Graph representing Composition of a Clock

Each numbered node in the state structure can serve as an index for the corresponding variable, such that x_0 represents the clock, x_1 its hands, and x_2 its face. Since the process fragments are existentially quantified variables, the graph can be interpreted as a structure that supports various true sentences. For example, one belief supported by this structure is that there exist three process fragments that are related to each other. This could be expressed formally using the sentence (using the notation introduced to describe Lambert's brick-wall universe),[34]

$$\exists x_1 \exists x_2 \exists x_0 \, (\alpha \, (x_0, x_1, x_2))$$

where α is an unspecified process composition relationship signifying how the nodes x_0, x_1, and x_2 are related to each other.

The structure also supports other valid beliefs. For example, since the hands x_1 and face x_2 are fragments of clock x_0, the process composition relationship α could be interpreted as fragmentation, resulting in the true sentence,

$$\exists x_1 \exists x_2 \exists x_0 \, ((x_1 \leq x_0) \; \& \; (x_2 \leq x_0))$$

meaning that the hands, clock face, and clock processes exist, and the hands are a process fragment of the clock, and the face is a process fragment of the clock. A further option would be to identify the process composition relationship as one of fusion, resulting in the sentence

$$\exists x_1 \exists x_2 \exists x_0 \, (x_0 \equiv (x_2 + x_1))$$

meaning that the clock as a process fragment, is identical to the fusion of the hand and face processes (*EMWC*: 238).

Similar strategies are used to introduce symbols representing quality nodes, suitably indexed. So for example, the existence of a grey clock face (a process fragment of the clock previously described) could be represented by a simple graph composed of two nodes, indexed such that a specific quality node Q_3 (the quality 'greyness') is associated with a process fragment p (the clock face) represented by the set node x_2 (depicted in fig. 6.5 (a)).

(a) (b)

Figure 6.5 Graph of Quality Node predicating a Set Node.

---

[34] As noted in chapter four (fn.22) for Lambert, composition of processes is signified by the symbol $\in$ and the symbol $\equiv$ is used to signify process identity. The symbol $\leq$ is used to signify that processes are fragments of the evolving universe. The symbol Ω is used to signify a universal process, called the *universe*. This symbol is also used to denote *everything* and the *world*. The symbol for complement is $\complement$, and the complement of everything ($\complement\,\Omega$) is nothing, denoted by $\perp$. The variables x and y are fragments of the universe Ω. Fusion of x and y is denoted $x + y$ (the union of x & y), while fission of x and y is denoted by either the inner fission product x with y ($x \bullet y$) or the outer fission product ($x - y$), where inner fission product denotes the greatest process enveloped by both of the given processes x and y (akin to the intersection of x & y), while the outer fission product consists of the remainder when the inner fission product is removed (akin to the complement of the intersection of x & y).

Lambert suggests that this existentially quantified quality node is most naturally interpreted as an *indexed unary predicate*, such that process made up of a single quality node Q_3 joined to a state node x_2 could be interpreted as $\exists x_2(Q_3(x_2))$, meaning that there exists a clock face that is grey. Using a similar procedure, if a quality is not possessed by an associated set node the absent quality can be represented by negating the associated quality predicate (fig. 6.5 (b)), and interpreted as $\exists x_2(\neg Q_3(x_2)$, meaning that there exists a clock face that is not grey. Similar procedures can be used to represent the absence of set nodes, absentee nodes in compositional structures representing fusion, fission and so on. This leads into a range of subtleties which I will not enter into. However, using this procedure, complex arrays of nodes can be used to indicate the existence and composition of very complex state structures. Classification (belief) involves matching dynamic patterns of qualitiative change found in the world with isomorphic patterns of state changes recorded in templates. How this occurs will be addressed in the section on diachronic semantics.

The unary character of this logic arises from "the fact that qualities are individually attached to processes, rather like Aristotelian properties (forms)" (*EMWC*:249). Lambert's templates overcomes the limitations associated with a unary predicate logics inability to express relations by allowing the quality nodes attached to process nodes to be themselves defined functionally by a template. Templates can define potentially complex relations without introducing specific relation symbols. Qualifications of qualities must be interpreted in terms of either primitive qualities or functional qualities defined by other templates. Thus a quality node can be associated with either primitive qualities detected by discriminators, such as 'greyness', or a functional quality picked out by a template, such as 'clock face'. A template may be composed of sub-structures representing different functional roles (such as the case in the appliance – switch example). These substructures can be abstracted from various templates and joined together "lego style", to create exceedingly complex state structures. Lambert also demonstrates that any of his state representations can be mechanically translated into a unary predicate first-order language (and vice versa).[35]

In sum, by using Lambert's semantics of synchronic processes it is possible to formulate existentially quantified statements about the world that make truth claims. In contrast to the semantics of Classical AI, Post-Classical beliefs about the world refer to patterns of influence that ρbot attends to through the state representation mediated by mathematical artifacts, rather than in terms of the manifest image propositions that mediate Classical beliefs about the world (*EMWC*:250). It is the resultant mathematical form of the structure, not our linguistically mediated interpretations of what these structures represent (clocks, trees, etc.) that determines Post-Classical belief. However, the semantics of synchronic processes does not detail how these truth claims can be narrowed down to a claim about a specific situation in a specific moment in time. That is, when ρbot attends to patterns of influence in the world that are picked out by discriminators designed to detect primitive or more complex functional qualities, it formulates a belief about a state at that instant.

How ρbot affects this match between functional templates and occurrences of patterns of influence is the subject of Lambert's discussion of the diachronic dimension of

[35] Appendix D of *EMWC* describes programs that automate this translation process.

belief. As will be shown in the next section, the use of quality nodes, suitably indexed to represent the diachronic dimension of existence, is the key to ρbot's capacity to assemble "lego block" belief structures that classify dynamically changing processes.

6.43 The Semantics of the Diachronic dimension of Post-Classical Beliefs

Having analyzed the synchronic dimension of ρbot's beliefs, Lambert then turns to the diachronic dimension of belief via an extensional semantics of state-instances. The key to the semantics of diachronic process is the transitory association of a state with its quality content. These transitory associations are referred to as 'state-instances'. A state-instance provides contextual content to a state, as it replicates the node structure of its state through node-instances, but can vary context through the qualification of its quality nodes (*EMWC*:251). A *belief state-instance* can thus be thought of "*as a state with quality content*" (*EMWC*:252; my italics). That is, in Whiteheadian terms, the state instance is defined by the extensionally defined pattern of potentiality objectified in the datum of an actual entity (a discrete system of positively prehended regions in the initial datum of the occasion that expresses that monadic occasion's perspective of its world) together with the qualitative subjective form through which the extensional reference to the datum is disclosed. Mediating the subject's prehension of this unified object are eternal objects of the objective and subjective species, corresponding to the extensional reference and its qualitative form. The prehended object could have both a synchronic and diachronic dimension; in musical terms it could represent the synchronic chord structure of the tune and the diachronic progression of those chords as they give expression to the melody. A belief state instance refers to one of those chords, and specifies its temporal and qualitative identity. The semantics of transitional classes embeds state instances in a system of temporal succession such that the musical chord becomes part of an anticipated progression of notes associated with the melody. Or in terms more appropriate to a discussion of semantics, transition classes entail beliefs that specify transitions to other beliefs, where the entire succession of beliefs describe a pattern of change in the world.

Thus in his semantics of diachronic processes, Lambert uses state-instances to represent "processes having a certain state structure for some period of time before reconfiguring to another state-instance" (*EMWC*:252). This time structure can be represented by modifying the indexed notation used to describe the nodes of a graph, such that existentially quantified variables are replaced with the inner fission product t with x_k (using the notation $t \bullet x_k$), where t is the period of time defining the state-instance S of process *p* at time t, and x_k the indexed variables that make up the state graph. For example, the qualitative association of greyness with a clock face, previously interpreted as $\exists x_2(Q_3(x_2))$, if regarded as the state instance S for process p at time t, now becomes,

$$S_p(t) =_{df} \text{period }(t) \ \& \ \exists x_2(Q_{p3}(t \bullet x_2)),$$

meaning that at a time period t there exists a clock face with the quality of greyness. Note that the quality predicate Q_{p3} has a dual sub-scripting (Q_{pi}, where p is the state and i the node) as in more complex templates, representing two or more states, it is necessary to be able to specify which state is predicated by which quality for that particular state instance.

The semantics of state instances becomes exceedingly complex, as it involves both describing the node structure of a state, and tracking changes in the content of that state through time. By developing a notation that associates content with a state for a period of time, time is formally introduced into state-instances, however, a template is a collection of state structures linked by a number of transition classes. Hence the extensional semantics of template use must include both the extensional semantics of state-instances and the extensional semantics of transition classes. Transition classes are fundamental to Lambert's system as they specify how the 'lego-block' structures of templates are clipped together. As with Whitehead's prehensive vectors, these transition classes are the real bricks of Lambert's universe. Transition classes mediate the process of classification that determines what ρbot believes.

According to Lambert a "*transition class* describes how content is transferred from one state-instance to the next" (*EMWC*:254). Transition classes specify that two state instances are locked together by their qualitative content in a specific way. For example, given the switch-appliance system described in chapter five (in which the template specifies identical transition classes between two states), if the state instance specifies the switch-appliance system in its 'off' state, then the transition class will specify the transfer of qualitative content to the system in its 'on' state. This transition class expresses the disposition of that state instance. Figure 6.6 (below) illustrates the relationship between transition T, states S and state-instances S(t).[36]

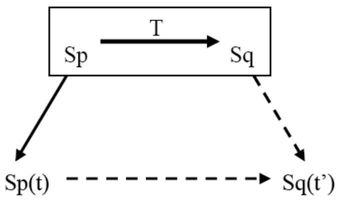


Figure 6.6 The relationship between states and state instances (EMWC:255).

The extensional semantics of the transition class linking the state **p** to the state **q** is denoted by $\mathbf{T_{p,q}(t,t')}$. This transition class "specifies the state-instance $\mathbf{S_q(t')}$ to be associated with the next state $\mathbf{S_q}$ at time **t'** given some state-instance $\mathbf{S_p(t)}$ of the current state $\mathbf{S_p}$ at time **t**." Expressed formally, when "the transition class to state **q** is the only transition class emanating from state **p**, the interpretation is

$$\forall t\ (S_p(t) \Rightarrow \exists t'\ (t\ ;\ t'\ \&\ T_{p,q}(t,t')),$$

where $\mathbf{S_q(t')}$ is a consequence of $\mathbf{T_{p,q}(t,t')}$" (*EMWC*:254-5). This can be read as saying that 'whenever a state-instance of state **p** persists for some period, it will subsequently transition to a state-instance of state **q** for some period'.

If a template represents a complex network with several transition classes emanating from the same state **p** then disjunctive subsequent states arise as possible alternate outcomes. Lambert suggests that the interpretation of the template network then becomes a formal theory made up of a finite number of "dispositional sentences for each state of the template with outgoing transitions" (*EMWC*:255). A disposition is

[36] Using S_p and S_q to signify successive conditions of a state S brought about by the transition T, exemplified in successive state-instances $S_p(t)$ and $S_q(t')$ arising in temporal intervals t and t'.

the tendency, disclosed in past occurrences of that state instance, for a system of associated causal influences to bring about a transition to a subsequent state instance. Each disposition can be formalized in a sentence expressing a transition class (or classes) that follow from a state instance.

The complete theory of template use is therefore specified by a set of sentences that describe all the dispositions associated with the template state network, where each sentence describes a state in the network along with the alternate transition pathways leading from that state to subsequent states. Each sentence in the theory is thus a proposition stating the transition class $\mathbf{T_{p,q}(t,t')}$ that specifies which subsequent state instance $\mathbf{S_q(t')}$ will follow from a prior state instance $\mathbf{S_p(t)}$. For example, the template network sketched in fig. 6.7 represents the sequence of possible states that can follow from an initial state S_1.

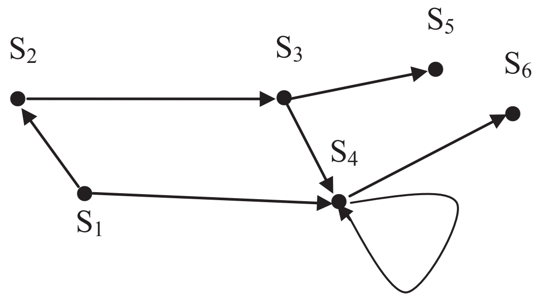


Figure 6.7 Template state transition network (*EMWC*:255).

The template can be interpreted in the model-theoretic sense as the formal theory specified by the set of sentences:

$$\{\forall t\ (S_1(t) \Rightarrow\ \exists t'\ (t\ ;\ t'\ \&\ (T_{1,2}(t,t') \vee T_{1,4}(t,t')))),$$
$$\forall t\ (S_2(t) \Rightarrow\ \exists t'\ (t\ ;\ t'\ \&\ (T_{2,3}(t,t'))),$$
$$\forall t\ (S_3(t) \Rightarrow\ \exists t'\ (t\ ;\ t'\ \&\ (T_{3,4}(t,t') \vee T_{3,5}(t,t')))),$$
$$\forall t\ (S_4(t) \Rightarrow\ \exists t'\ (t\ ;\ t'\ \&\ (T_{4,4}(t,t') \vee T_{4,6}(t,t'))))\ \}$$

This template network corresponds, in my interpretation of Whitehead, to the relational structure of the complex eternal object which defines the proposition (theory) instantiated in a particular habit; the theory includes sentences representing branching paths, 'either – or' possibilities in its finite list of possible transitions from one state to subsequent states. What this theory describes is a set of patterns of influence that together define a societal object in the real world. For a concrescing occasion to form a belief means that it has arrived at a satisfaction by interpreting prehended qualitative forms of feeling in terms of its switch template. This means matching the pattern of transfer of qualitative content between state instances of the real switch with patterns of potential transfer specified by its switch template. The process is analogous to that which Henry (1993) illustrated in general terms using Prolog, however Lambert provides a practical strategy for fulfilling the Henry vision. His extensional semantics of transition classes is the keystone to this strategy.

6.44 Extensional Semantics of Transition Classes

Lambert completes his semantic account of template use with an account of the existential semantics of a general transition class $T_{p,q}(t,t')$. The extensional semantics of transition classes specify the assignments that determine the true sentences constitutive of a theory of the transition class. As illustrated with respect to the switch-appliance template in the previous chapter, transition classes describe how qualitative content is transferred from one state instance to the next. They provide an intensional concept of persistent processes. Lambert's extensional semantics of transition classes rests on the presupposition that these persistent processes (repetitive patterns of process, manifesting themselves qualitatively in successive state instances), are a function of spreading patterns of causal influences associated with the transition process.[37] These spreading patterns are disclosed in qualitative indicators, either direct causal indicators or functional indicators. These qualitative indicators enable repetitious patterns of change to be represented in templates as discrete repetitious patterns of qualitative change; the pattern can be viewed as a finite system of states, with potential transitions between states described in terms of qualities that are *inherited*, *introduced*, or *terminated*.

In Whiteheadian terms, qualitative content is introduced into a state instance when a subject prehends the subjective forms of socially reproduced patterns of feeling disclosed in a datum. The societal pattern disclosed in the datum is analogous to a state instance. A society is a structured system of processes that can be individuated as a system of functional roles; the structured roles that define a society will determine how the qualitative patterns prehended in a datum will be inherited by successive occasions of that society. In Whitehead's metaphysics, this causal influence is explained in terms of transitional processes dominated by causal conformity, and supplementary concrescent processes dominated by dispositions derived from purposive habits. In *EMWC* Lambert also posited a basic relationship of causal influence between fragments of societal process as a foundational assumption.

Thus, just as the identity of Whitehead's societies rest upon the notion of self-replicating societal patterns of influence that constrain and enable each transitional process, Lambert's semantics of transition classes rests on the presumption that external influence accompanies each successful existential state-instance change. This flows from the heuristic presumption "that the repetitive nature of qualitative state transition derives from instances of the transitioning states influencing instances of the recipient states"; alternatively, in his formal notation, Lambert's influence presumption can be represented as $(t \bullet x) \longmapsto (t \bullet y)$, meaning that "some fragment of $(t \bullet x)$ influences some fragment of $(t \bullet y)$" (*EMWC*:256).

6.45 Qualitative Inheritance, Qualitative Introduction and Qualitative Termination

Qualitative inheritance, introduction and termination define the intensional form of transitions between successive states. In Lambert's Post-Classical approach to belief, when ρbot detects repetitive patterns of influence in its environment it is able to record them as sequential transitions of partial states in its memory; these recorded patterns of influence constitute the habits that provide ρbot with its commonsense capacity. This capacity rests upon the principle that the same patterns of influence that

[37] That is, "the repetitive nature of qualitative state transition derives from instances of the transitioning states influencing instances of the recipient states" (*EMWC*:256)

occurred repeatedly in the past will continue to occur in the future, and that ρbot's habits (templates) therefore provide a solid basis for inference about likely occurrences in the world. I will detail the formal notation that Lambert introduces to deal with the semantics of qualitative inheritance, introduction and termination, as it is too technical for my purposes; however it builds upon the previous notation in a manner that can be indicated informally.

In Whitehead's language, Lambert's self-replicating patterns of influence are societal objects; just as societies are disclosed via the subjective prehension of qualitative forms that disclose the enduring characteristic form of the society, so too Lambert's enduring objects manifest themselves in the repetitive patterns disclosed in the qualitative forms mediating our experience of transitions between successive states of our world. Lambert's proposal that qualitative inheritance, qualitative introduction and qualitative termination are sufficient to describe these patterns of qualitative change thus promises to clarify Whitehead's notion of the characteristic forms that mediate a subject's perspectival objectification of a societal process; it does this by providing a model of how the eternal objects function, not as static ideal forms such as the Platonic numbers or geometric forms, but rather as dynamic ideal forms that specify transitional processes. In order to reflect this dynamic role, eternal objects must specify how the qualities that predicate the subjective form of the present occasion are related to the qualities that predicated the subjective form of previous and future occasions. That is, they must specify a historical dimension. They must also be able to individuate qualities functionally in terms of the historical role that they are performing in a given process. Together, this historical reference individuated in terms of a function (role), enables a pattern of societal change to be tracked through a pattern of inherited qualitative transitions. In addition to inherited change the complex eternal object objectifying a society must also accommodate the introduction and termination of qualitative roles. Eternal objects of this type would enable an occasion to classify external patterns of qualitative influence in a manner that reflected their diachronic societal significance. Lambert's semantics of transitional classes provide a mechanism for accomplishing this function.

My Whiteheadian gloss to Lambert's notion of the qualitative predication of patterns of change, in terms of inherited, introduced or terminated roles can be illustrated in terms of a family. As a social object a family is made up of patterns of activity that can be individuated in terms of roles (father, mother, daughter, son, together with nominated related role activities.). Each role is itself a society and can be broken down into further societal substructures in the Whiteheadian sense. However as functionally defined qualitative entities, a state transition graph of the family could consist of the introduction and termination of roles (e.g., signifying births, deaths, rights of passage that introduce or terminate the quality defining a role in the family), together with the inheritance of that quality over successive moments of that role (i.e., inherited quality 'son', 'daughter', etc., identified with a particular role). In a similar sense, Lambert attempts to individuate and represent the historical context and changes that characterize the typical patterns of change associated with a simple switch-appliance society. Roles such as those of switch or appliance are tracked and situated in a context. Representation of these state transitions within ρbot's memory provides a basis for matching specific qualitative patterns of change with events that have a causal reference to processes occurring in the external world. Hence in order to define the transition class $T_{p,q}(t,t')$ it is necessary to show what quality content has been inherited, introduced or terminated.

According to Lambert, the "superimposition of qualitative introduction, qualitative inheritance and qualitative termination transitions define transition classes", and of these "qualitative inheritance provides the avenue for content attribution within a template" (*EMWC*:256). Inheritance is the most important factor in a society and the most difficult to represent. In Whitehead's metaphysical scheme it is fundamental attribute of all transitions, even simple physical feelings, as it is the basis of the continuity of qualitative feeling that is constitutive of his process universe. Thus, when a past actual entity is prehended as part of the datum of a new occasion, the new occasion also prehends the qualitative feeling of the subjective form of feeling that characterized that past actual entity's satisfaction. However, what was felt as a subjective form by the previous occasion is now felt as a datum. For Whitehead, the inheritance of the past is not simply based upon prehending a past actual entity; it also involves prehending the qualitative feeling felt by the past entity. This feeling of what was felt in a mode that indicates the transition between occasions is the basis of the qualitative passage of time and underpins all routine anticipations. Henry (1993) provides a formal representation of the intrinsic part that historical reference plays in simple physical feelings.[38] Lambert's specification of qualitative inheritance involves a similar historical reference in transmitted qualitative forms.

Lambert draws upon the stack example of representation (section 5.21 in the previous chapter) in order to show how the historical dimension of inheritance can be illustrated in terms of a non-empty stack undergoing a push operation. Push and pop operations are the basis of transitions between states represented by the stack where, as described in the previous chapter, each operation can be defined in terms of a transitional class that incorporates into its form a historical reference to the previous state. Thus, a push transition will bring about a new state instance along with a transformation in the content or qualities associated with its node instances. The top of the stack element changes, with the substack at time t_1 inheriting the qualitative content of its node structure from the previous stack at time t_0. While the substack inherits the compositional form of its previous instance it does not inherit exactly the same characteristics as it is no longer the top element of the stack; however it does inherit the historical element of *having been* the top element.[39] In Lambert's words:

---

[38] A physical prehension (or physical feeling) is a subject's feeling of another feeling in a specific actual entity. It is an example of the form of a pure physical feeling in which past societal entity is "objectified by one of its own physical feelings" in a subsequent prehending occasion (PR:245). Henry (1993) provides an example "that symbolizes a conformal simple physical feeling by the dominant occasion of a personal society , named *p1*, of a feeling of *anticipation* held by a past dominant occasion of that society, called *p3*." Henry represents this prehension of a simple physical feeling formally as:

Anticipation(p1, anticipation(p3, Physical_datum)).

Where:

The constant *anticipation* refers to a specific emotion.
The constant *p1* refers to the subject of the physical feeling.
anticipation(p3, Physical_datum) is the form of the feeling felt.
The constant *p3* refers to the subject of the feeling felt.
Physical_datum is the form of the physical datum of the feeling felt (Henry 1993:42).

[39] To represent the historical dimension of inheritance involves superseding the unary notation indicating a quality predicate ($Q_{p,i}(-)$) with a binary predicate ($Q_{p,i}(-,-)$); eg. For parametric x, $Q_{p,i}(x,t)$ "expresses the property (quality) of having the historical quality $Q_{p,i}(t)$ of time t" (*EMWC*:257).

> What should be inherited is the historical quality of having been the top of stack element at t_0. Qualitative inheritance is about new processes inheriting the compositional form of old processes, together with the historical qualities of those processes" (*EMWC*:257; Lambert's underlining).

In addition to introducing notation for representing a historical parameter, Lambert also shows how to parameterize the role of individual nodes. In inheriting the functional qualities of a node structure, each node can be individuated in terms of its functional role in delivering a specific quality to another node (for example, the transitions between state instances in the switch-appliance template represented in fig.5.4 are defined in terms of 'switch quality', 'appliance quality', 'up quality' and 'off quality').[40] I will not detail his notation for dealing with this, but essentially the push and pop operations can be thought of as an 'if-then' pattern matching procedure that ideally arrives at a true or false classification. Qualitative inheritance involves pattern-matching a node instance structure from a current state with a node structure in a new state, to produce a new node instance with a structure homomorphic to the new node structure.[41] Lambert defines the four possible kinds of qualitative inheritance transitions as:

$\mathbf{quality}_s(n_{p,i}) \rightarrow \mathbf{quality}_s(n_{q,j})$; $\mathbf{set\text{-}node}_s(n_{p,i}) \rightarrow \mathbf{quality}_s(n_{q,j})$;
$\mathbf{quality}_s(n_{p,i}) \rightarrow \mathbf{set\text{-}node}_s(n_{q,j})$; $\mathbf{set\text{-}node}_s(n_{p,i}) \rightarrow \mathbf{set\text{-}node}_s(n_{q,j})$.

The first two are the basis of the stack push transition class referred to above, while the second two involve inheriting structure, rather than quality, and are hence more complicated to describe.[42] Because they specify inheriting structure, several qualitative assignments can proceed from the one transition (*EMWC*:260).

Thus qualitative inheritance is about new processes inheriting the compositional form and historical qualities of old processes. It specifies how qualities predicating set nodes in one state instance transition to the next state instance; by selectively tracking qualities through successive temporal states qualitative inheritance enables transition classes to identify persistent process 'objects'.[43] Lambert concludes that qualitative

---

[40] One in which a quality node inherits recursively from itself; and another in which a quality inherits from a process node. The functional quality of performing the role of x_0 at time t_0 is defined as the role played by the node that delivers a quality to a subsequent node. If a process x_0 delivers an inherited quality in that state instance $S_1(t_0)$ then this role of delivering the functional quality can be parameterized as $Q_{1,0}(u, t_0)$ using the parameter u performing the role of x_0 in the state instance $S_1(t_0)$. If the quality is externally qualified (either as an external primitive or functional quality) then it must be interpreted as an inherited external quality at that time instance (*EMWC*:258).

[41] In mathematics, a homomorphism is a structure-preserving map between two algebraic structures.

[42] Lambert illustrates the latter two inheritance transition classes using the previous stack example when a stack state undergoes a pop operation. In both cases there is a search that proceeds down through the stack structure representing the state, on and 'if – then' basis, where fail leads to an 'or-else if-then' until a match results. It there is no match there is no inheritance.

[43] Lambert (2009b) provides a simple example of how the matching procedure used in *EMWC* can be implemented using Prolog in order to store beliefs in ρbot's memory. This example builds upon the Prolog example of representation [footnote 21, chapter 5] provided in the previous chapter. There he demonstrated how the state transition structure for routines can be implemented through pattern matching predicates "with variables acquiring values providing the ability to instantiate qualities". In explaining how this example applies to beliefs Lambert writes:

> The stack example [c.f. footnote 21, chapter 5] shows how variables in pattern matching programming languages can provide the contextual qualitative capability characterised in *EMWC*. But these contexts only survive while that piece of code is running through the <variable, value> assignments to its variables. To retain state instances, they need to be stored in memory. So the stack states:

(a) (stk, mt) and (b) (stk, (sub), (top))

might then instead be stored in memory in the form

(a) stack(stk, mt) and (b) stack(stk, (sub), (top))

Propositions then hold the structure which, as noted in *EMWC*, itself has a propositional interpretation. The stack transitions are then recoded as follows, retracting and asserting state instances to memory (":-" means "if" if you are unfamiliar with Prolog):

i. Push onto empty stack:

```
push(Stk, Top) :-
        retract(stack(Stk, Mt)),
        assert(stack(Stk, (Mt), (Top))).
```

iii. Push onto a non-empty stack:

```
push(Stk, Top2) :-
        retract(stack(Stk, (Sub), (Top))),
        assert(stack(Stk, (Stk, (Sub), (Top)), (Top2))).
```

iv. Pop to obtain a non-empty stack:

```
pop(Stk, Top2) :-
        retract(stack(Stk, (Stk, Sub, Top), (Top2))),
        assert(stack(Stk, Sub, Top)).
```

ii. Pop to obtain an empty stack:

```
pop(Stk, Top) :-
        retract(stack(Stk, (Sub), (Top))),
        assert(stack(Stk, Sub)).
```

The following then illustrates the same interaction as before but with storage to memory.

```
| ?- assert(stack(number_stack, 0)).
yes
| ?- stack(number_stack, Mt).
Mt = 0 ?
yes
| ?- push(number_stack, 1).
yes
| ?- stack(number_stack, Sub, Top).
Sub = 0,
Top = 1 ?
yes
| ?- push(number_stack, 2).
yes
| ?- stack(number_stack, Sub, Top).
Sub = (number_stack,0,1),
Top = 2 ?
yes
| ?- push(number_stack, 3).
yes
| ?- stack(number_stack, Sub, Top).
Sub = (number_stack,(number_stack,0,1),2),
Top = 3 ?
yes
| ?- pop(number_stack, Top).
Top = 3 ?
yes
| ?- stack(number_stack, Sub, Top).
Sub = (number_stack,0,1),
Top = 2 ?
yes
| ?- pop(number_stack, Top).
Top = 2 ?
yes
| ?- stack(number_stack, Sub, Top).
Sub = 0,
Top = 1 ?
yes
| ?- pop(number_stack, Top).
Top = 1 ?
yes
| ?- stack(number_stack, Sub, Top).
no
| ?- stack(number_stack, Mt).
Mt = 0 ?
yes
```

Lambert concludes that "In this way beliefs and other working memory elements come to store propositions that record contextually instantiated state instance structures" (Lambert 2009b:7-9).

inheritance "is essentially a pattern-matching exercise" that "promotes a conception of objects" through tracking qualitative patterns and generating an "intensional conception of these persistent processes" (*EMWC*:261-3).

6.46 Types as the Eternal Objects of Post-Classical AI

Having dealt with qualitative inheritance, Lambert then turns to the remaining transition kinds necessary to define the general transition class $T_{p,q}(t,t')$: qualitative introduction and qualitative termination. Every template usage begins with the initialization of a state instance through the introduction of external content. Lambert notes that this qualitative introduction of external content can take the form of either a primitive or functional quality, where the introduction of an external primitive quality records "the causal influence of a discriminator output", while the introduction of an "external functional quality is a reference to a template type in some other template-instance" (*EMWC*:263). Lambert details a notation for representing the semantics of each quality; the semantics of primitive qualities is dealt with by introducing a predicate for each fixed discriminator,[44] while the semantics of functional qualities is based upon the semantics of template types. I will not say much about these operations (both make a minimal semantic contribution - Lambert describes them as 'bookkeeping' operations for organizing memory within the robot), however Lambert's discussion of qualitative introduction leads into some observations on the *semantics of types* that clarifies and reinforces remarks made in the previous discussion on Lambert's agent perspective.[45]

In that context I noted that types are akin to philosophical universals; in Whiteheadian terms they are the eternal objects of Lambert's paradigm. From Whitehead's metaphysical perspective the meaning of a belief arises in the context of a subject entertaining a proposition; this entails predicating something of an indicated nexus.[46] Similarly, I have suggested that in Post-Classical AI, the meaning of a belief arises in the context of ρbot's attribution of a *type* to an extensional process, where *use* determines the type and *causal reference* determines the extensional process. Lambert's analysis of qualitative introduction reinforces this point. The role of types can be illustrated in terms of Lambert's switch template.

As in *EMWC*, Qualitative inheritance, qualitative introduction and qualitative termination are the three kinds of transition. While each of these operations is defined with respect to nodes in some state structure S_k, in practice each is used as an operation on node-instances in some state-instance Σ_k, structurally homomorphic with S_k (as depicted in figure 6.6). Thus, each transition class expression T linking S_1 to S_2 provides a computation between state instances. The states themselves are not explicitly represented. The transition code does not provide a declarative representation of the state structure; rather, the role of the state structure is encoded procedurally in the compiled activity performed by the transition class T. Each of the compiled transition procedures should return a true (yes) or false (no) to indicate whether the transition was successful. Hence each transition expression defines a transition computation on node-instances, and the "compilation of transition expressions produces executable code placed in long-term memory" (*EMWC*:298).

[44] The primitive quality P associated with the discriminator *i* is represented as the quality P_i predicating process x at time t. Thus each primitive quality $P_i(x, t)$ "registers the influence of x on a designated discriminator output for that primitive quality" (*EMWC*:264). By introducing a discriminator$_i$ predicate for each fixed discriminator, primitive qualities can be defined in formal terms as follows:

$$P_i(x, t) =_{df} \exists p\ (\text{discriminator}_i\ (p)\ \&\ \exists u\ (u \bullet t \leq p\ \&\ x \mapsto u \bullet t)).$$

[45] See section 6.24; also footnotes 12 & 16.

[46] See 'tree growing on the moon' example in chapter section 3.31.

A template type is made up of set nodes; the template type tracks an identical role performed by one set node through the various states of the template. In computer science types are defined from abstract set theoretic types like the natural and real numbers, however in Post-Classical representation, "types can only be set nodes in memory, and can only be generalized by considering set nodes across templates" (EMWC:267). Template types individuate the role played by set nodes within repeated patterns of influence. There can only be one set node in each state that belongs to that template type. For example, in the switch template, an appliance set node is specified for each of its two states. Lambert notes that the "two appliance set nodes track the role of the appliance in relation to the switch, and in doing so, specify what *type* of process an appliance is in relation to a switch"(*EMWC*:263). The semantics of functional qualities is based upon the semantics of template types. Each set node in a given state contributing to a type can be associated with a functional quality (property Q) in a similar manner to qualitative inheritance. "In this sense", Lambert notes, "the set node contributes a quality based upon its *substructure* and this is the quality acquired through external qualification" (*EMWC*:265).

As previously noted, Lambert demonstrated that for any template it is possible to formulate a set of formal sentences that convey the semantics of that template. Templates specify patterns of influence and template types specify the roles disclosed in these repeated patterns; the formal sentences that make up the theory of a template show how each template instance uniquely determines a pattern of influence as indicated by "by qualitatively limiting the choice of variable assignments for which satisfaction obtains" (*EMWC*:267). In specifying the semantics of qualitative introduction for qualities derived from the functional role of nodes in templates Lambert demonstrates that Post-Classical belief delivers a type individuation of the world that delivers to ρbot its capacity to modify templates to cope with novel situations. By individuating roles that make up templates it is possible to identify identical roles in what are otherwise completely different templates. In this way polymorphic typing occurs, where for example, a similar switch role might be identified in a number of different templates.[47]

Thus Lambert's Post-Classical approach to the semantics of belief structures is based, not on linguistic terms, but on representational structures that pick out the functional roles played by collections of set nodes within a template. A template is a representation of patterns of influence that occur in the world; that is, like characteristic eternal objects that define a Whiteheadian society, they specify potential forms of self-replicating patterns of activity. These patterns of activity can be further individuated in terms of functional roles; hence template types are defined in terms of functional types. In a similar sense, the complex eternal object that defines a particular society might consist of a multiplicity of component eternal objects, each of

---

[47] Polymorphic typing refers to the association of multiple types with the one process. Polymorphic typing is often used with finite state machines, such as employed in computer games, as it allows common behavior/roles to be shared by different FSMs, as for example, if the same 'walking' type is utilized by various templates representing different game characters (see examples from games in Raim 2008). Lambert notes that types have been viewed as lattices and analysed algebraically. Complete lattices were viewed as semantic models of primitive data types and their operational outcomes. Subsequently, types were viewed as algebraic ideals and adopted as a means of managing polymorphic typing in programming languages (*EMWC*:227).

which could qualify particular transitional prehensions. Lambert's semantics of functional types provides a model of how a series of prehensions could by typed to indicate the subordinate structural components of the parent society.[48] In this way types become abstract universals that can be applied in contexts beyond those in which they were first individuated. In summary:

- Each template specifies a set of patterns of influence
- Template types select out roles within these patterns of influence
- Template types are composed of set nodes of a template with no more than one set node for each state in the template
- Types can only be set-nodes in memory and can only be generalized by considering set nodes across templates
- The same type can occur in several templates, hence polymorphic typing

Types therefore secure the lego connectivity that enables different routines from ρbot's memory to be coupled together to adapt templates – creating new lego block structures - that provide ρbot with the capacity to classify novel conjunctions of processes. Each such template is, in Whiteheadian terms, a proposition (theory) about a situation, with multiple propositions providing a means for representing complex situations. Thus, as Lambert concludes, types have a ubiquitous place in his system of representation as roles in patterns of influence.[49] This provides an abstract objective criterion for truth. In Lambert's words:

> Fragments of templates composed solely of template set nodes represent abstract functional qualities in the world by specifying the role a process must play in a pattern of process interaction. These fragments of templates have been termed *template types*, and fragment of the world become *occurrences* of those types whenever they abstractly conform to the roles prescribed by those types. Templates therefore secure an objective abstract *type individuation of the world* and the relationship between types and their occurrences secures an abstract objective criterion for *truth*, whereby something is an occurrence of **type τ** of **template s** if and only if it satisfies the role prescribed for τ in s" (EMWC:266).[50]

---

[48] As Lambert notes in the context of discussing the role of sets of set nodes (such as the switch example): "As each of the set nodes is merely an individuated process, any collection of such nodes is also a single process, and so a collective functional role within a template can be identified by a single process, which within the template, is individuated as a collection of nodes." (*EMWC*:265-6)

[49] Lambert notes that although ubiquitous, types are a poor basis for memory organization, and for that reason neither types nor templates are explicitly represented as such in ρbot's memory; their primary function is to illustrate the general principles underpinning Post-Classical AI (*EMWC*:268). How they are translated into programming strategies this will be discussed in the next chapter.

[50] Using the parametric u, to signify a process fragment u of the partial world $\mathfrak{U}$, the latter sentence of this quotation can be expressed formally as

$u :_s \tau$ (meaning that u is an occurrence of type τ in s, where s specifies a template).

This is true only if u satisfies the role prescribed for τ in s.

Lambert notes that he has shown that for any template s it is possible to formulate a set of formal sentences Σ_s that convey the extensional semantics of that template. Furthermore, since this set is of a finite cardinality it is possible to express the set as a single conjunctive sentence σ_s. Lambert then goes on to show that for constant symbol a denoting some process in the Universe Ω, that $a :_s \tau$ can be defined as

$$a :_s \tau =_{df} \sigma_{s\tau}(a)$$

with the set of formulae $\Sigma_{s\tau}(u)$ defined in terms of process predicates derivative of $\{\leq, \underline{\angle}, \mapsto\}$ so that $\Sigma_{s\tau}(u)$ is of type $\mathbf{T} = \{\leq, \underline{\angle}, \mapsto\}$. Lambert comments that: "If all existential quantifications are removed in lieu of the introduction of variable assignment functions, then it becomes a set of formulae $\Gamma_{s\tau}(u)$ of type $\mathbf{T} = \{\leq, \underline{\angle}, \mapsto\}$." Then each template s effectively specifies the set of patterns of influence, and a

6.5 Conclusion

In this chapter I have shown that Lambert's philosophy of belief rests upon the possibility of mechanically replicating and pattern matching aspects of the structure of self-replicating patterns of influence found in the external world. Belief for Lambert's robot is an act of classification; to classify is to predicate the world, not in terms of language, but rather in terms of roles in templates that provide an intensional specification of patterns of activity found in the world. Lambert's robot is designed to pick out patterns of qualities in its external environment, using discriminators and templates; ρbot's beliefs about its world are qualitative intensional representations of these repeated patterns of influence.

I have argued that ρbot's templates are analogous to Whitehead's propositions (theories), in that both function as structured systems of activity that provide a model-theoretic context for interpreting patterns of influence in the external world. Lambert, like Whitehead, assumes that we do in fact possess something analogous to templates that guide our selective attention to repeated patterns of influence in the world; I have referred to these structures as routines or habits. A process world is individuated not by words, but by habits that are individuated in terms of abstract types and extensionally referenced to the world via a chain of causal influences. As with propositions and eternal objects in Whitehead's metaphysical system, templates and types can potentially mediate the representation of the rich diversity of societies that make up our Universe.

I have also described the main tenets of Lambert's Compliant Realism (that our commonsense world is structured by the forms we impose upon it and that our representations are at least partially disclosing what the world is really like) and suggested that, in Whiteheadian terms, this means simply that while our habits are imposing forms upon the world, the qualitative forms disclosed through those habits are prehensions of patterns of feeling and functional relations that are a real part of the world. This extensional reference provides Post-Classical AI with its semantic criteria for truth and reality.

Lastly, I gave some indication of how Lambert translated his philosophy of belief into a practical semantic system that could potentially be implemented in computational code. The focus was on the state instances and transitional classes that give expression to the routines that create lego style structures, made up of process transitions that map the dynamic patterns of influence that constitute Lambert's Post-Classical world. I have also suggested that Lambert's system builds primarily upon a procedural interpretation of events, and drawn upon Henry's (1993) argument to propose that Whitehead's notion of eternal objects, and the process of concrescence, can also be best understood in procedural terms.

---

template instance will specify a pattern of influence by "qualitatively limiting the choice of variable assignments for which satisfaction obtains." (*EMWC*:266-7)

Lambert's Post-Classical philosophy of belief slightly modifies Dummett's vision of an ontological hierarchy in that his Compliant Realism is based upon a circle of dependence rather than a linear hierarchy. This reflects the role of agent intent in defining the world. Lambert's agents are structurally coupled to their environmental niches in a manner analogous to Whitehead's vision of organisms. Ecologically embedded habits provide the deep foundations of agent intent and the commonsense capabilities that are employed to realize agent aims. This evolutionary context is fundamental to the Post-Classical view of language; it underpins the rejection of the Manifest Image view of language and the adoption of a process grounded view of language.

Lambert's analysis of representation and belief has involved introducing Lambert's Post-Classical agent, ρbot. In Whiteheadian terms ρbot is a mechanical, non-living society that is carefully structured in order to functionally simulate some of the strategies that living systems employ to perspectivally objectify their environments. Communications are an obviously vital element in human society and they also play a key role in Lambert's agent society. In Lambert's terms an agent society is a collection of robotic individuals that communicate 'intent', so as to cooperate in the pursuit of shared goals. How intent is introduced into Lambert's multi-agent system will be the subject of the next chapter.

Chapter 7

Towards a Philosophy of Mind for Post-Classical AI

Chapter 7 – Section Headings

7.0 Introduction – Lambert and the Whiteheadian vision of Mind

> I find myself as essentially a unity of emotions, enjoyments, hopes, fears, regrets, valuations of alternatives, decisions – all of them subjective reactions to the environment as active in my nature. My unity . . . is my process of shaping this welter of material into a consistent pattern of feelings . . . which is myself at this moment, and yet, as being myself, it is a continuation of the antecedent world.
>
> Alfred North Whitehead (*MT*:228).

> The character of an actual entity is finally governed by its datum; what ever be the freedom of feeling arising in the concrescence, there can be no transgression of the limitations of capacity inherent in the datum. The datum both limits and supplies. It follows from this doctrine that the character of an organism depends on that of its environment. But the character of an environment is the sum of the characters of the various societies of actual entities which jointly constitute that environment . . ."
>
> Alfred North Whitehead (*PR*:110).

> I contend that the external behaviour we associate with commonsense is really the combination and interaction of environmentally triggered or machine motivated execution of repeatable chunks of internal behaviour. I call these *repeatable chunks of internal behaviour* routines, as the term conveniently marries its everyday sense of 'customary' with its computer science sense of 'procedural'.
>
> Dale Lambert (*EMWC*:288).

Lambert's Post-Classical paradigm is an attempt to envisage an alternative philosophical path along which to pursue the dream of artificial intelligence. The Post-Classical model of an artificial mind, in contrast to the Classical model, does not presuppose that an adequate ontology of the world is captured in language (the 'Manifest Image' fallacy); instead it proposes a process ontology that is based upon a critical analysis of the phenomenal world. Having rejected the Classical presupposition that a rationally coherent and complete ontology is implicated in Natural Language, and rejected the related notion that commonsense is mediated by linguistically based rational calculations, Lambert's Post-Classical paradigm advances an alternate situated view of agents, in which commonsense capabilities are based upon habits or customs that are fitted to the environment of agent action. Unlike previous attempts to model situated agents, Lambert's model builds upon a process vision of the agent universe; a universe in which all enduring entities are self-replicating patterns of influence in the flux of Creation.

According to Nicholas Rescher, the crucial distinction between a substance oriented and process oriented approach to mind, is a shift of focus from *things* to *doings*. From the notion that the unity of the agent lies in the unity of its body, its memory, or its customs, to the notion that agent unity stems from the activities constitutive of the agent's history: "one is the individual that one is by nature of the macroprocess that integrates the microprocesses constituting ones life and career" (Rescher 1996:108). As argued in chapter three of this work, the dominant philosophical figure in reviving this process view of the world has been Alfred North Whitehead.

My attempt to illuminate the trajectory of Lambert's work by situating it within a Whiteheadian context has produced a broad brush-stroke image of situated machines that function by imitating the propositional basis of human habits. I have argued that Lambert's world of processes can be likened to Whitehead's prehensive flux of occasions and similarly, that Lambert's self-replicating patterns of influence are analogous to Whitehead's societies. I have also argued that Lambert's templates play an analogous role in his proposed cognitive machine as do Whitehead's habits in his higher-level societies. For Whitehead the process of shaping the flux of human experience into coherent perspectives is mediated by habits; Lambert's cognitive machine objectifies its world by drawing upon routines. However, Lambert's machine model of commonsense is not attempting to precisely replicate human commonsense. Rather, his work is an attempt to implement an abstract and idealized model of commonsense – a simulation – that begins with an abstract model of the Universe, and proceeds via abstract, formally modelled notions of representation and belief towards a procedural model of a mind. That is, a mind framed in terms of *doings* rather than *things*.

Lambert's work can be viewed as an attempt to build a computational simulation of the role played by Whitehead's propositions in individuating and objectifying the subjective agent's world. This simulation is based upon the presupposition that at a fundamental level the phenomenal universe manifests a mathematical order. I have argued that Whitehead's method of extensive abstraction was a procedure for capturing the implicit order in the phenomenal realm and representing it in terms of a calculus of extension. In presenting a metaphysical model of a generic monadic agent, Whitehead's analysis focused on the role played by subjective feelings; however, in accord with his extensional representation of the Universe, Whitehead emphasised that aspects of the subjective process of concrescence could be represented in mathematical terms. I have argued that for Whitehead this mathematical order is implicit in the indicative structures that tie his metaphysical propositions to the logical subjects that they predicate (see Chapt.3.31 on indicative feelings). Lambert's approach to AI can be viewed as an attempt to extend this abstract formal approach to other elements of Whitehead's metaphysical system: Post-Classical AI is a simulation based upon an abstract, idealized representation of the process universe.

In the preceding chapters I have explained that this Post-Classical simulation is based upon formalizing a theory of processes (Ch. 4), developing representational structures defined in terms of these processes (Ch.5), and using these representational structures to provide a semantic basis for attributing a semi-realist notion of belief to a robot

(Ch.6). Together, these three chapters show that Lambert provides a strategy for translating Whitehead's propositions into intensional structures that mediate classifications of activity extensionally identified with patterned processes in the external world.[1]

The conceptual traffic between Whitehead and Lambert has not been one way. While Whitehead's philosophy provides a unifying context for bringing together Heraclitean process and Humean customs, Lambert's philosophies of representation and belief provide illuminating analogies for dealing with Whitehead's notions of eternal objects and propositions. When considered in the light of Lambert's work, Whitehead's metaphysical propositions can be viewed as an attempt to represent sub-structures within that wider mathematically ordered Universe; substructures (or partial worlds) that provide a basis for situating and evaluating mathematical 'sentences' that represent patterns of extensionally indicated activity taking place in the world. Templates arise from and are fitted to representing recurrent patterns of activity in a particular environment; yet through mediating abstract modes of classification ('types') they can give rise to novel beliefs (classifications effected by novel 'lego' templates, constructed on-the-fly to fit particular situations). Thus, by advancing an environmentally grounded form of abstract representation, Post-Classical AI creates an alternative notion of belief, which forms the basis of an alternative approach for engineering commonsense. This Post-Classical vision constitutes a significant resource for reflecting upon and further developing the Whiteheadian philosophical legacy.

Lambert's *EMWC* proposal for a Post-Classical mind, as outlined in this chapter, is essentially suggesting how the formal apparatus outlined in previous chapters can be put to work in a cognitive machine. The chapter proceeds on the assumption that a metaphysical ontology of processes provides an adequate basis for capturing what takes place in the human world, and that a single generic class of templates can be constructed to represent the recurrent patterns of qualitative change detected amongst the process flux. In the next chapter it will become evident that Lambert modifies this position, and recognizes the need to supplement his formal ontology of processes with material ontologies that build additional systematic constraints (structural categories) into the agent's environment. However, this modification builds upon principles that are explicated in this chapter; hence the view of mind advocated in *EMWC* is a necessary basis for appreciating subsequent developments in Lambert's work. Thus, in this chapter I will outline Lambert's proposal for translating templates into implementable programs (attitude programming) as well as briefly describing Lambert's proposals for organizing the internal architecture of his robot. This chapter thus describes a critical milestone in Lambert's endeavor to translate a novel vision of a process inspired alternative AI into a practical proposal for engineering an artificial mind.

7.1 A Multi-Agent Architecture for a Post-Classical Mind

[1] That is, §3.3 showed that "a robot using these representation structures can be understood to be classifying the world *intensionally* in terms of dynamic qualitative sets, and *extensionally* in terms of repeated patterns of influence" (*EMWC*:269).

The artificial mind of Lambert's ρbot is based on a multi-agent architecture composed of a collection of agents. His design embraces a distributed design philosophy "in which each agent commands its own dedicated physical processor, and communicates with other agents by message passing" (*EMWC*:273). The roles assigned to his various agents are loosely inspired by a functionalist approach to the macroscopic neuroanatomy of the brain.[2] The ρbot multi-agent system is equipped with various sensor and effector agents, mediating inputs and outputs respectively, together with control agents that oversee and motivate the active co-ordination between sensor and effector agents (see fig.7.1). The core of ρbot's capacities lie in the careful functional organization of its referential tools, with agents processing inputs at multiple levels of abstraction in order to generate the flow of data that culminates in an appropriate output. "In practice", Lambert states, "each component [of the architecture] is an autonomous reactive agent with a dedicated functional role to perform, which it accomplishes by communicating with other agents in the architecture" (*EMWC*:273).

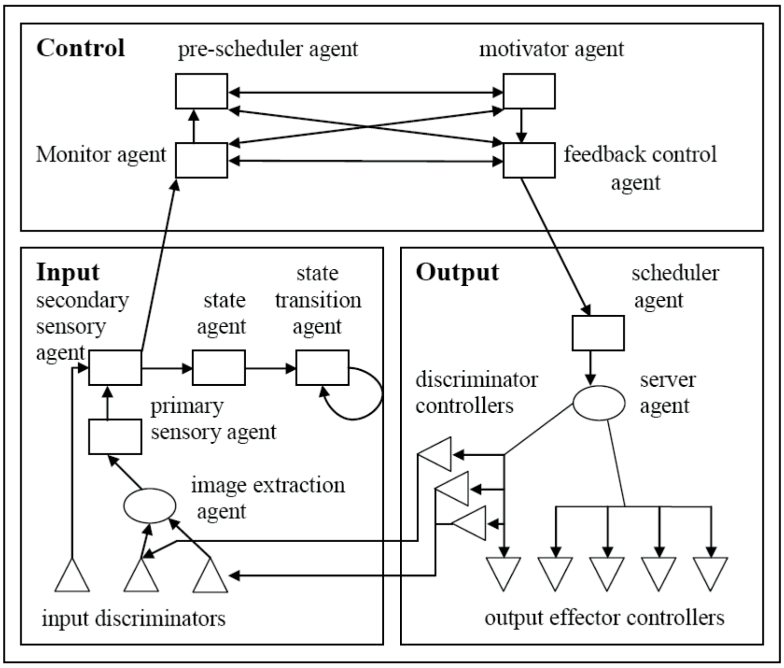


Figure 7.1 A multi-agent architecture (*EMWC*:272)

[2] Lambert (2009b) comments that he does not pursue this neurophysiologically motivated design concept in subsequent work, but has instead "moved toward a cognitive multi-agent architecture with agents on the same or different computational platforms" (2009b:10). However I refer to his neurologically motivated model here as it signifies one of the main challenges facing a Post-Classical AI; namely that of escaping from the mind = brain = computer analogy that has underpinned the Classical Computational-Representational Understanding of Mind (CRUM) as outlined in chapter one (cf. Thagard 2005:10-12). This traditional focus on the functional architecture of the brain to the neglect of the functional architecture of the environment's in which the brain is embedded is one shortcoming of Classical AI that Agre and Horswill (1997) draw attention to in their analysis of the role of the 'life-world'. Privileging the solitary brain metaphor of mind leads to a lop-sided approach to representation that focuses on the internal processes of classification but neglects the external systems of organization in which the agent is embedded - and the role that the embodied mind's interactive participation in these systems play in the development of the brain's neural architecture and its associated processes of classification. My goal in this work is to suggest that Lambert's Post-Classical AI, when framed in terms of Whitehead's philosophy, provides a bridge between cognitive science and social theory, and forms a basis for building both a theoretical perspective and practical procedures for implementing a social interactive understanding of representation and mind.

In order to achieve his goal, Lambert proposes a Post-Classical alternative to the Classical AI notion of knowledge engineering. The guiding theme of his embedded approach to engineering commonsense is that the knowledge engineer "seeks to recover what we do, not what we think we know" (*EMWC*:280). Instead of attempting to program the machine to emulate our flawed, retrospective manifest image rationalizations of why we act in the way we do, Lambert proposes to "embed the engineer inside the machine and confront the engineer with the machine's embedded perspective on the world" (*EMWC*:279). In practical terms this could mean placing ρbot in a virtual world and allowing an engineer to assume the role of the various agents that make up ρbot. By receiving only the input a particular type of agent receives, the engineer could take the functional role of that agent in mediating ρbot's interaction with its environment and use their human intelligence to select out the important patterns of influence from the standpoint of the particular agent. In this way an initial core of routines could be created for each agent, with selected routines forming the basis of a toolbox that could be used to construct more complex routines until a sufficient stock of routines had been established for ρbot to function effectively in its virtual world (*EMWC*:279-80). Thus with embedded engineering, each of the multiple agents in ρbot is programmed from the agent-perspective by an engineer; the agents are then weaned from the engineer so that ρbot functions autonomously in its (virtual) world.

7.11 Attitude Programming

The key to Lambert's Humean (habit-based) approach to intelligent behavior is his proposal to engineer the routines that mediate the agent's adaptation to their environment using templates. In chapter six I sketched Lambert's philosophy of belief, and explained how processes within ρbot could represent processes in the world. To that end, templates were used to classify ρbot's external experiences of the world, in effect recording the beliefs ρbot formed when interacting with processes in the world. The stock of templates in ρbots long term memory shape its capacity to classify processes in its working memory and respond appropriately to its world. Hence when ρ classifies x as x_q, the belief x_q is an instantiation in ρbot's working memory of a long term memory of the qualitative structure type q. In this way the templates of long term memory shape ρbot's classification of (and hence beliefs about) the world. Lambert's vision of an artificial mind that enacts these classifications is based upon a distributed multi-agent architecture in which each component agent performs a dedicated role. Lambert's aim is to distribute long-term memory across the agents that instantiate ρbot's architecture, with each agent's long-term memory servicing its specific role within the ρbot architecture.[3] Lambert further suggested, as a practical compromise, that these long-term memories should be programmed into ρbot using the Embedded Engineering strategy. His proposal for carrying out this instructional programming task is developed in what he terms 'attitude programming'.

[3] Although memories clearly play a role in mediating routines, this approach ignores the dynamic role that carefully structured environments can play in soliciting and specifying routines. Environments are not a neutral space, but a dynamic aspect of the formation and maintenance of appropriate routines, as argued by Agre and Horswill (1997).

The basis of ρbot's intelligent capability is thus a stock of beliefs stored in long-term memory; however beliefs alone do not give rise to commonsense judgments and actions: beliefs work in conjunction with other attitudinal orientations, such as orienting and motivating fears and desires. Hence Lambert's conclusion that "for ρbot to possess even a shadow of commonsense" it must have "*forms of internal behaviour other than belief"* and it must "*record these various forms of internal behaviour in long-term memory"* (*EMWC*:281-2, italics used in the original). Our memories of past experiences are weighted in terms of their emotional significance, and these shape our response to future events. The fear associated with an encounter with a snake, for example, lends caution to future encounters, even though they may occur in very different circumstances. Thus ρbot's Long-Term memory should include, alongside beliefs registering external experiences, a mechanism for registering an internal attitudinal response consistent with its embedded role in the repeated patterns of influence. Thus, templates must be upgraded to record a range of behavioral experiences over and above beliefs about the world. This is the aim of attitude programming.

Lambert discusses various kinds of attitudes, like fear, worry and hope, and argues that belief, desire and expectation seem more primitive. Furthermore the latter attitudes can be implemented as operators on data bases (memory). Propositional attitudes are alleged mental states that can be characterized in a propositional attitude expression. Lambert suggests that *propositional attitude expressions* are typical of our communications with each other and that they can be represented using the form:

<subject> <attitude> that <proposition expression>

where 'subject' denotes the individual whose mental state is being characterized; the 'propositional expression' describes some propositional claim about the world; and 'attitude' is expressing the subject's dispositional attitude toward the claim about the world. For example:

<Fred/Tom/Bill> <believes/fears/expects> that <it is raining/the sky is blue/today is Monday>.

Such expressions enable us to communicate our own mental dispositions to others, and express judgments about the mental activity of others, based upon their external behavior (*EMWC*:282).

Since the process grounded representational strategy developed for his Post-Classical AI involves replacing proposition expressions with transition class expressions, Lambert's approach to formally representing attitudes is to replace propositional attitude expressions with transitional attitude expressions (*EMWC*:284). Transition classes (see chapters 5.2 and 6.4) refer to a new state structure that is made up of elementary transitions from a previous state structure (A → B). A transition class can be viewed as a directed connection between two state structures [state instances], with the aim being to successfully negotiate passage from the first state to the second, successfully undertaking all of the required transitions (as represented in fig. 7.2).

f→(x –y), g→ε, ε→a
((f)g) --→ ((x-y)a)

Figure 7.2 Illustration of a transition class (*EMWC*:285)

Transition classes may succeed or fail in generating the content of the desired successor state.[4] The connection between the old and new state structure is made only if the transitions enable content to be fully passed from the first state form to the second state form; "in this way transition classes engender a semantics of success, with success determined by whether the requirements for transferring content to the successive state are successfully negotiated" (*EMWC*:285). This semantics of success underpins Lambert's notion of commonsense behaviour.

Lambert's approach to explaining and modeling human activity has much in common with Philip Agre's computational approach to modeling human experience. Both Agre and Lambert adopt a Humean perspective on the vital role of custom in coordinating and giving direction to human affairs, and both regard a reactive architecture as an appropriate way to implement something akin to customs in a computer simulation of agent problem solving behavior. Agre drew upon his daily life as a source of inspiration for his reactive architecture. He observed that much of his day-to-day activity was mediated by routine responses interacting dynamically with his environment. For example, he observed that when walking home from the subway he did not use an elaborate world model to plan his journey and then verify and update that plan using temporary situation-bound representations constructed during the course of his journey; instead in an unreflective manner he used context bound rules that were triggered by contact with external stimuli. For example, sight of the subway stairs triggers execution of the rule: "on exiting the station head in the same direction as the subway stairs"; and sight of the three star hotel triggers: "turn left at the first road after the three star hotel"; and so on. He implemented his insight in a computer game, Pengi, in which the penguin seeks to achieve a goal, not by planning and representation, but by "an improvisatory interaction with continually evolving surroundings" (Agre, 1988:189).

Lambert was evidently aware of the parallel tendencies in their work, and drew attention to both similarities and differences between their works. Citing a lengthy passage from Agre and Chapman (1987) he notes that their use of the term 'routine' corresponds to his notion of a 'template instance'.[5] For Lambert, a template instance

[4] In this case the transition between the initial state structure ((f)g) to the state structure ((x -y)a) is successfully achieved by matching the compositional structure of f with (x -y), discarding g and introducing a. The match between f and (x -y) is possible because the unspecified quality content of f (say (α β)) is inherited by the successor state (x -y).

[5] Lambert writes that: "Their [Philip Agre and David Chapman] use of the term 'routine' corresponds to my use of the term 'template instance', and so that substitution should be made when reading the following quotation.

> Routines [template instances] are patterns of interaction between an agent and its world. A routine [template instance] is not a plan or procedure; typically is not represented by the agent. An agent can reliably enter into a particular routine [template instance] without representing it because of regularities in the world. For example, imagine the penguin running from a bee. The penguin will run as far as it can, until it runs into a wall made of blocks. Then it will have to kick its way through the wall. Then it will run some more. Then it will hit another wall. The process could be described by a procedure with two nested loops: running until it hits something, kicking the obstacle, and then repeating.
>
> But this same pattern of activity could equally well arise from a pair of rules: (R1) when you are being chased, run away; (R2) if you run into a wall, kick through it. These rules don't represent the iteration; the loop emerges as a result of the interaction of the rules with the situation. Causality flow into the system from the world, drives the rules which chose what to do, resulting in action which changes the world, and back again into the system, which responds to changes (Agre and Chapman, 1987: 269).

emerges from a pattern of interaction between ρbot executing routines and its environment in which the environment contributes to the routines selected. As noted in the epigraph to this section, Lambert contends "that the external behavior we associate with commonsense is really the combination and interaction of environmentally triggered or machine motivated execution of repeatable chunks of internal behavior", and calls these "*repeatable chunks of internal behavior* routines" (*EMWC*:288). For Lambert, routines are the library of recipes that the agent can draw upon in response to various environmental contingencies. A template is any combination of routines, and a template instance emerges from a pattern of interaction between ρbot executing routines and its environment in which the environment contributes to the routines selected. As a combination of routines, a template will most often be selected because the to-state of one routine complies with the from-state of another routine whose to-state is nearer the goal. That is, the template is created using routines as the 'lego block' connectors that secure the desired transition from one state to the next (*EMWC*: 291).

Routines thus describe chunks of internal behavior necessary for bringing about a transition from some initial state to a desired goal state. An agent's routines are individuated as " *goal annotated state transition networks* in which each network connection is defined by the transition class expression of some transitional attitude, and each goal annotation is defined by the goal expression of that transitional attitude" (*EMWC*: 288). Lambert's routines are thus effectively *directed graphs formed by linking individual transitional attitudes together,* with the satisfaction of each transitional attitude depending upon the successful formation of mediating state instances between the initial from-state to the final to-state that signals the goal of the particular routine. The expression of routines is thus recursively based on joining together transitional attitude expressions.[6]

In summary, Lambert's proposal for programming ρbot is to apply an embedded engineering approach to replace the transition classes between the states of ρbot's templates with routines. These routines define repeatable chunks of internal behaviour and are composed of a network of transition classes, each of which is annotated by a goal to be accomplished before transition to the next state occurs. The satisfaction of each attempted goal turns upon the capacity of ρbot's multiple agents to communicate beliefs, desires, expectations and anticipations. The next section explains how this ontology of routines, beliefs, desires, expectations and anticipations are integrated into the multi-agent architecture of ρbot.

7.12 Agent Architecture

In order to fulfill their roles, the agents that make up the distributed multi-agent architecture of ρbot must be able to communicate with each other. Moreover, they must be able to persuade each other to cooperate in the pursuit of a goal which means in some sense sharing a desire for the same end. Lambert's suggestions as to how his

---

[6] The engineer communicates these routines using 'routine expressions'. Lambert notes that: "Routine expressions describe a chunk of internal behaviour for achieving some goal state, when some initial state prevails. Routine expressions therefore consist of: a from-state, disclosing the intitial state required for routines execution; a to-state, announcing the goal state the routine endeavours to achieve; and a routine-body, which textually describes a directed graph of transitional attitudes" (*EMWC*:288).

agents accomplish this coordination involves developing an architecture that coordinates beliefs (classifications), desires, expectations and anticipations into some kind of memory and communication structure.

Lambert begins by defining each agent as an agent process, in accordance with his theory of processes, and individuates each agent as a fusion process combining a working memory process (**WM**), a Long Term Memory Process (**LTM**) housing routines, and a fixed **Control** process that orchestrates the activity within the agent architecture (as represented in fig.3.6). This can be represented as:

Agent ≡ WM + LTM + Control.

The working memory process can be further individuated as a fusion of a communication process (**Com**), action process (**Act**), **Desire** process and classification process (**Classif**), and represented as:

WM ≡ Com + Act + Desire + Classif.

Operationally the agent architecture combines a post-classical case-based philosophy with a reactive architecture. Each agent performs its role by first communicating a desire to another agent to satisfy a goal and then, having instigated this desire in the other agent, receiving and acting upon the results of that agent's efforts when it attempts to satisfy that desire (*EMWC*:308-11). To facilitate this passing of messages between agents, each agent's processor is endlessly executing a receive process (designed to accept messages from other agents), and a send process (designed to send messages to a particular agent), and a control process (designed to administer the internal activity of the agent). All activity in the agent architecture stems from the performance of actions which originate either from within the agent or externally (see fig. 7.2). Actions are organized into mail boxes. Control within the agent architecture is dictated by a 'semantics of success', and hence time is critical to agent behavior. Thus, desires, expectations and anticipations have an expiry time (*EMWC*:312-9).

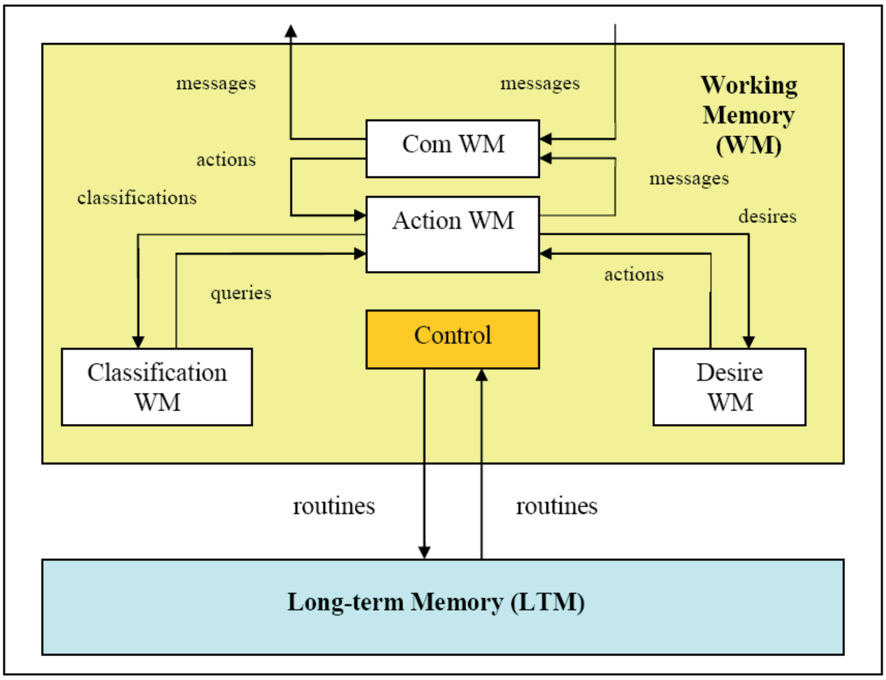


Figure 7.2: Agent Architecture (*EMWC*:311)

Desires play a crucial role in the operation of each agent, as each agent's actions are the result of internal behavior initiated by the agent's desires. These desires are contextual references to goals located in routines in long-term memory. Internal behavior is a consequence of dependent desires contextually traversing transitional attitude paths through routines, exploring alternate paths until its action succeeds in realizing the desired goal.[7] A dependent desire is satisfied if the action succeeds and is unsatisfied if the action fails (*EMWC*:319-21). In explaining the architecture of the working memory (WM) the role of desire (Desire WM) in controlling action (Action WM) is closely related to the role of classification (Classification WM).

The role of Classification Working Memory is best appreciated by recalling that in ρbot classifications function as beliefs. Within ρbot's multi-agent architecture beliefs are distributed across the component agents, with each agent having its own classification working memory. As noted previously, these classifications can also include expectation and anticipations, and so that at any moment each agent's classification working memory can be individuated as a set of beliefs, expectations and anticipations. Thus, each belief, expectation and anticipation is an instance of a classificatory working memory structure administered by a corresponding classification attitude (*EMWC*:323-4). An attitude-semantics, based upon a 'semantics of success', is specified for each of these classifications.

Each agent's classification memory structure is made up of beliefs, expectations and anticipations. These three classification types are used to classify persistent partial states using a state-instance structure (a state-instance is the association of content [an ordered lattice of nodes] with a state for a period of time, as described in chapter 6). Classification working memory can be individuated as a set of beliefs, expectations and anticipations. When applied to the same state-instance, beliefs, expectations and anticipations are distinguished by the different attitude orienting the classification toward that particular state instance. As listed by Lambert, they can be differentiated as follows:

- When created, a *belief* is used by the agent to classify that the state-instance structure of that belief <u>is</u> representative of a persistent state in an occurring partial world.

- When created, an *expectation* is used by the agent to classify that the state-instance structure of that expectation <u>will be</u> representative of a persistent state in occurring partial world.

- When created, an *anticipation* is used by the agent to classify that the state-instance structure of that anticipation <u>might be</u> representative of a persistent state in occurring partial world.

The alternative attitudes toward state-instance structures are possible because of the semantics of success associated with each of the three attitude goals (*EMWC*:325).

[7] Dependent desires can also generate independent desires, in which satisfaction is sought independently of the context and routine execution to which the desire belongs. This in turn leads to desire hierarchies, with the execution of an independent desire involving the linked dependent desires traversing of the transitional attitude paths that make up their routines, until the desire meets with success or failure (*EMWC*:320-1).

In sum, belief, expectation and anticipation are kinds of classification memory structure that are used to classify a persistent partial state using a state-instance structure. Belief thus involves creating a belief state-instance in working memory, while expectations suspend particular internal behaviors until a belief occurs, and anticipations invoke particular internal behaviors when a belief occurs (*EMWC*:326).

7.13 Memory Architecture

Routines are the basis of ρbot's capacity for responding in an intelligent fashion to its environment. ρbot is based upon a multi-agent, reactive architecture in which a stock of routines, stored in each agent's long-term memory, facilitates each agent's active communication and cooperation with the other agents that make up ρbot. Each routine defines a repeatable chunk of internal behaviour which is designed to facilitate a transition from an existing belief state to a desired belief state. In the above sections I have outlined Lambert's suggested method for programming each of ρbot's component agents with a stock of routines (using an embedded engineering technique to attitude-program each agent in ρbot's multi-agent architecture). I then described the function of the main elements of working memory, including control, desire and classification. In this section I complete Lambert's architectural sketch by providing a brief description of the architecture and function of the long-term memory.

Broadly speaking, the long term memory performs three functions, the *addition* of items, the *retrieval* of items, and the *removal* of items. Lambert notes that the removal of items from long term memory is an important issue, but has no specific proposals concerning that operation. In addressing the remaining two operations he first notes that the function of adding items to the memory concerns the question of where the routines of long term memory come from. The answer Lambert provides is that routines are either specified by the engineer or fostered by the machine, based upon its 'experience' with the initial set of routines. The retrieval of items involves three problems; what should be retrieved (selection problem); how to store items retrieved so they can be easily retrieved (storage problem); supplementing the long term memories reactive control mechanism with a deliberative mechanism (control problem) (*EMWC*:333-4).

Lambert's solution to the selection problem (what should be retrieved) involves replacing the semantic pattern matching of Classical AI (based upon the manifest image paradigm with its associated brittleness problems), with 'segregated structural matching'. In a suggestion that has potential sociological implications, Lambert suggests that his architecture segregates its representation structures (routines) by attaching them to different agents, and further differentiates them in terms of memory functions. The segregation of structure between agents and within agent memory contributes to the meaning of a structure; what a structure means is determined by where it resides. In the same way that a person can use the functionally differentiated structure of their city's organizational landscape to enable them to find the service that they are seeking, so also the segregated structure of ρbot's agent architecture can contribute to solving the problem of selecting appropriate routines. Having decided where to look for routines with the appropriate semantic capability, there is than the problem of which routines to select, based upon their *from* and *to* states. Post-classical representation allows this to be reduced to a similarity match between qualitative set structures (*EMWC*:336).[8]

7.2 Concluding remarks on Lambert's proposal for a Post-Classical AI

Lambert's visionary dissertation began with a critique of the epistemological, heuristic and metaphysical adequacy of Classical AI, and culminated with his presentation of an alternative 'Post-Classical' AI paradigm. His proposal addressed the weaknesses identified in the Classical paradigm. Firstly, he outlined a post-classical *metaphysics* in which the world is composed of processes, and is understood in terms of repeated patterns of influence among those processes. Lambert argued that process metaphysics provides a more adequate account of how the world really functions than the substance based categories of Classical AI. Specifically, it provided an *epistemological* basis for explaining how matters of interest come to be selected and represented in our commonsense world. He showed how dynamic qualitative sets could be used to represent repeated patterns of influence. These representations were then used to formulate a semi-realist account of belief, and an extensional semantics was presented to show how these representations classify repeated patterns of influence in the world. Lambert then addressed the question of *heuristic* adequacy by proposing a particular multi-agent architecture grounded in routines that represent patterns of influence in the world. An embedded engineering programme was suggested as replacement for the knowledge engineering programme of classical AI, and attitude programming was introduced as a means of programming the constituent agents with relevant routines. Finally, Lambert outlined an internal structure for the agent architecture.

My account has focused on the intuitions underpinning Lambert's Post-Classical proposal – I have argued that Whitehead's philosophy provides a coherent conceptual framework within which to situate Lambert's Heraclitian vision of a process universe grounded in patterns of influence and his Humean vision of commonsense behavior grounded in customs. My goal has been to describe in a non-technical (non-mathematical) fashion how, by giving a process twist to Barwise and Perry's (1983) realistic, model-theoretic approach to situations and attitudes, and Agre and Chapman's (1987) situated approach to AI, Lambert has developed a strategy for formalizing and implementing his Post-Classical vision of an artificial mind. My account does not do justice to the logical rigor that Lambert brings to his analysis, but does attempt to indicate the general form of his approach to formalization and implementation. What I hope does stand out is the unusual combination of depth of philosophical insight and breadth of technical expertise that informs Lambert's project; Post-Classical AI builds upon the powerful logical legacy of Classical AI and uses that toolset to give practical expression to a grand unifying philosophical vision, the contours of which are visible in *EMWC*. Whether Lambert's dissertation achieves its aim of promoting a new paradigm for artificial intelligence remains to be seen. Historically, novel intellectual visions have often taken many years to find their

---

[8] Lambert notes that a connectionist approach could be used to solve this problem, even though he uses a different approach. He also notes that the storage problem could also be solved using a connectionist approach (ANN). Other approaches are also considered. (*EMWC*:336-7)

audience. However, if this Post-Classical paradigm is to receive the recognition that it deserves and become the *principia intellectus* guiding efforts to engineer artificial intelligence, it needs to be embraced critically, and hence I offer these observations.

Lambert skates rather lightly over some of the wider ontological issues related to his adoption of a process metaphysical system. His exposition begins with his formal representation of the universe in terms of an extensional set theory that is akin to Whitehead's theory of extension as elaborated in Part IV of *Process and Reality* (*PR*), but he fails to outline a metaphysical vision in the wider sense of specifying a system of concepts that encompasses the various kinds of phenomena that constitute our reality. While he grounds his notion of matter, representation belief and mind in a notion of process, the connections are under-theorized, both with respect to contrasting objective and subjective standpoints, and with respect to the 'material ontologies' that define the sphere of human action. Whitehead recognized that a process metaphysical system required an analysis that described both the objective orders which fragments of process constitute and the subjective coming-to-be of each fragment of creation, as reflected in his theory of *extension* and his theory of *concrescence*.[9] Whitehead regarded the objective order of extended relations, and the subjective mode of perception through which the subject apprehends and evaluates that order as complementary modes of analysis. Lambert's philosophies of existence, representation, belief and mind traverse this domain but the underlying process philosophic framework of ideas remains at best implicit.[10]

This philosophical deficit impacts upon the clarity of Lambert's Post-Classical schema. Though very sophisticated in terms of its approach to semantics and various formal issues relating to representation, the metaphysical status of the external 'patterns of influence' that his system attempts to model remain unclear. In my Whiteheadian interpretation of Lambert, I have noted that 'influence' for Whitehead included both the past actualities acting as efficient causation and propositions acting as final causes 'luring' subjects towards a goal. Lambert's Post-Classical paradigm makes a similar distinction between efficient causal forces and final causality in the form of 'beliefs'; but the ontology upon which this philosophy of belief rests flattens

[9] As explained in chapter three, Whitehead adopted Heraclitus' position, and argued that the Cosmos is engaged in a continual process of Becoming. He envisaged the Universe as a creative flux, through which past events, in their perishing, are continually influencing the formation of a new fragment of Being, and that each fragment in turn, as it perishes becomes part of the Universe that influences further fragments, and so on. Nothing survives in Whitehead's cosmos, except the patterns of identity that are preserved in the continual regeneration of the Cosmic flux; this preservation of identity through process takes place in a manner analogous to the way that institutions in a society preserve societal identity in the midst of the constant perishing of the people that fill the roles constitutive of the institution.

[10] For example, in *EMWC* Lambert proposes a model of causal influence to explain representation and perception, however, he does not explicitly address the thorny ontological issues that concern the distinction between human and machine intelligence. These issues are centre stage when Lambert is outlining his process semantics and comparing human and machine forms of grasping the meaning of something. Whitehead explicitly addresses the different qualitative dimensions of various forms of causal influence (forms of prehensive 'feelings' particular to non-living and living matter). Lambert skates over this issue, and fails to provide the coherent approach to ontological levels that Whitehead's conceptual framework provides. Human and machine intentionality are assimilated in the same model without making the kind of distinctions that Whitehead was careful to make between different modes of 'prehension' and the different types of 'society' that arise in conjunction with more complex modes of prehension. The result is a conceptual framework that blurs the distinction between the types of causality that influence machine and human perception.

the corresponding real world domain of influences into a homogeneous domain of bare patterns of causal influence. There is no attempt to provide generic categories for dealing with various phenomenal levels of being. What are the forms of traffic that connect the macro-societal process to the individual micro-processes that shape careers and history? That is, commonsense categories bear witness to the physical, emotional, social and cultural macro-structures that form environments of action; human commonsense is cognizant of these contexts, and it is clear that they have a causal influence on action, but how should these structures be represented and what mediates the traffic of influence from macro-structure to individual agent action? As Talcott Parsons commented with respect to Whitehead, it is not enough to say that "connectedness is the essence of all things"; the analyst must suggest "how and on what terms and with what consequences for each other things are connected" (Ackerman & Parsons 1966:39). Lambert, in developing his Post-Classical notion of representation, leaves much of the detail of his process universe implicit, and the metaphysical ontology that underpins his system provides very little formal indication of the ontological structures that support human action.

Thus, while Lambert's proposal for a Post-Classical artificial mind does involve making explicit suggestions concerning formalizing elements that have been simply assumed in previous chapters, such as the notions of memory, purpose, emotion and the communication of beliefs, the ontological dimension of human commonsense action remains underdeveloped. This possibly reflects the continuing influence of the Classical legacy; that is, like Classical AI, *EMWC* is oriented toward engineering an individual agent, and does not take into explicit consideration the social dimension of human commonsense and its social correlates in machine interaction. This issue is in part addressed in Lambert's subsequent work (as will be discussed in the next chapter) however it remains a lacuna. Given the sociological motivations behind my analysis it is worth while rehearsing the reasons for this deficit.

In the Classical tradition the architecture of the computer mind, was inspired largely by the presumed equation between mind, brain and computation. In my Whiteheadian reading of Lambert I have proposed an analogy between state instances and the societal datums mediating Whitehead's Many, and likened Lambert's transitions to Whitehead's prehensive transitional and concrescent processes. In both cases the selective transition is determined by a qualitative content that is specified by a habitual mode of becoming; a disposition that orients the agent and constrains the path along which it seeks its satisfaction (in both senses of the word). For both Whitehead and Lambert the satisfactory fulfilment of a transitional process means predicating an internal process in the agent with a particular qualitative pattern and matching that against some external process that is the extensional reference of the internal process. This means that the agent must be engineered to couple with the external pattern of influence (via discriminators and templates).

From a Whiteheadian (and a Socionic) perspective, this structural coupling means that the form of the external environment is an intrinsic element in defining the internal structures of an agent. Hence the architecture of the agent 'mind' should proceed in conjunction with consideration of the architecture of the external environment. This would appear obvious to an engineer designing a robot to fit a physical environment, such as a robotic aircraft, submarine, or automobile, but in the case of engineering commonsense there has been an assumption that language has enabled mind to transcend any specific environment. This, I would argue is part of the linguistic Manifest Image fallacy, and is partly responsible for the failure of Classical AI, as argued by Lambert in *EMWC*.

Instead of attempting to model embodied, socially situated agents, Classical AI engineers attempted to extract a precise and rational, universally applicable calculus from human language. However, as Bertrand Russell commented: "Understanding words does not consist in knowing their dictionary definitions, or in being able to specify the objects to which they are appropriate . . . Understanding language is more like understanding cricket: it is a matter of habits, acquired in oneself and rightly presumed in others" (Russell 1921:147). Russell's observation points the way to a socially embedded understanding of language and intelligence; it suggests that more attention needs to be devoted to the dynamic patterns of interaction between agents and concrete external environments; what are the forms of action that agents engage in; what are the 'games' that people play? This, I submit requires more of a sociological focus than is evident in *EMWC.* These issues become more important when attempting to develop a framework of ideas for dealing with hybrid societies capable of differentiating between the capabilities and forms of awareness possessed by machine and human agents. As will become apparent in the next chapter, Lambert has to some extent drawn this same conclusion.[11]

These critical comments aside, *EMWC* is a tremendous achievement. Many of the ideas developed formally in his dissertation have been tried and proven in subsequent work and the formal details can be tracked through the relevant papers. As will become evident in the next chapter, Lambert's post-doctoral career has taken Post-Classical AI out of the realm of speculation into the laboratory and the domain of practical application. In the course of these developments ρbot has evolved from thought experiment into a functioning (experimental) information fusion system, and the metaphor guiding this development has changed from personal robotic assistant interacting with individual humans, to that of a virtual society made up of many societies of computer agents. These machine societies are in turn integrated into human society, constituting a hybrid society distributed across virtual and real worlds. The strong social flavor of some of Lambert's more recent work has brought out more clearly the sociological implications and as yet under-developed sociological potential of his Post-Classical paradigm. As I will attempt to show in the next chapter, the natural fit between Lambert's vision of MAS based AI and sociological theory creates a fertile ground for applying the Socionic methodology of using sociological theory to guide the design of MAS and MAS to implement sociological theory.

I will conclude this review of Lambert's ground-breaking dissertation with a final observation. As presented in *EMWC*, it is difficult to see the forest of Lambert's Post-Classical paradigm through the tree-trunks of technical detail. By presenting Lambert's work against a Whiteheadian philosophical horizon I have attempted to emphasize the intuitions motivating Post Classical AI; to some extent these philosophical intuitions are more important than the formal proposals for their implementation, for if Lambert's critique of Classical AI is valid, and his Post-Classical intuitions remedy the defects in Classical AI, then the way is opened to an era of intelligent machines. And as Lambert notes, regardless of how his Post-

[11] But not to the extent of embracing sociology, which is of course my justification for framing his work in Whiteheadian terms and proposing a Socionic extension!

Classical structures are implemented, "the ramification upon our moral, social, political and cultural lives will be immense if it proves successful and propels us out of the information age and into the intelligent age" (*EMWC*:348).

Chapter 8

Lambert's blueprint for Information Fusion

Chapter 8 – Section Headings

8.0 Introduction

Lambert's critique of Classical AI, and his suggestion for a shift toward a Post-Classical AI paradigm, as outlined in his 1996 dissertation *Engineering Machines with Commonsense (EMWC),* was an ambitious, speculative vision of how to overcome the biggest problem confronting AI: that of engineering machines with a capacity akin to human commonsense. His proposal was based upon re-appropriating the logical and mathematical tools used by Classical AI engineers within a process philosophy inspired framework of understanding. Thus, building upon a foundation of a formal process ontology, Lambert constructed a Post-Classical notion of representation and belief, and used these foundations to develop a multi-agent architecture for an artificial virtual mind. In Lambert's words, his dissertation "markets a vision for how things [pertaining to engineering a commonsense machine] might be understood differently" (*EMWC*:347). As a marketing exercise his thesis clearly found a receptive audience with his employers at the Defence Science and Technology Organization (DSTO). This chapter summarizes the progress that Lambert has made, under the auspices of DSTO, in developing and implementing his Post-Classical vision in a real world information fusion system.

The first challenge in transforming *EMWC* into a marketable military research project was that of repackaging the aims of *EMWC* so that they conformed to the research objectives of DSTO. In practical terms this meant that instead of speaking about engineering a 'commonsense' machine and promoting a new AI paradigm, Lambert began to promote 'automated information fusion', aimed at establishing a superior form of 'situation awareness'. In the jargon of DSTO, the goal of this work was to "pioneer a paradigm shift in command environments through a superior use of capability and greater situation awareness" (Estival et.al. 2003). As we shall see, the rubric of information fusion and situation awareness has enabled Lambert to pursue the Post-Classical agenda sketched out in *EMWC*. The packaging has changed, but the Post-Classical vision outlined in *EMWC* remains at the heart of Lambert's work in information fusion.[1]

In detailing Lambert's progress I will touch briefly upon his vision of the role of information fusion in the operation of the Australian armed forces. The decision to develop an automated information fusion capacity reflects a broader military policy of transforming the armed forces into a networked, highly automated fighting force. This emerging military organization exemplifies the new hybrid form of society that will be created as virtual, machine-agent societies are engineered to operate as an integral part of human organizations. As such, Lambert's project could be approached from the perspective of Socionics. His work at DSTO provided an opportunity to validate and further develop a vision of how to engineer machine intelligence; in addition the proposed integration of machine and human agents into the operation of the military has provided a case study of an emerging hybrid society. Thus, the main aim of this chapter is to examine the progress made in implementing Lambert's vision of a multi-agent system grounded in a process ontology and, as an ancillary goal, to use the envisaged military application of Lambert's information fusion system to illustrate the broader (largely neglected) Socionic context pertaining to the development of hybrid societies.

To examine Lambert's progress in realizing his Post-Classical vision, and provide a Socionic perspective on his project, I will first describe how Lambert re-appropriates the notion of an automated Information Fusion system and transforms it into a vehicle for applying the core ideas of his Post Classical AI. In the second section of this chapter I describe the structure of Lambert's current blue-print for an information fusion system, and show that its content is an elaboration of the principles elaborated in *EMWC.* In the third section I expand upon the social level of Lambert's process ontology which, according to Lambert, is crucial to the operation of the system as a whole. Finally, I note some significant remaining challenges and suggest that the Post-Classical approach outlined in *EMWC,* if joined to the Socionic methodology, may provide a useful basis for further developments.

8.1 Transforming Information Fusion into a vehicle for applying the core ideas of Lambert's Post Classical AI.

The first step in marketing Lambert's Post-Classical paradigm to DSTO was to couch its objective in terms that conformed to the Australian Government mandated objectives of DSTO research. This has been accomplished by repackaging Lambert's AI research within the rubric of information fusion, and presenting information fusion as a key component of DSTO's response to the Government's defence policy

[1] Lambert has recently stated that "*EMWC* is a part of who I am" and acknowledged that the ideas developed in *EMWC* lie "behind my fusion models and my FOCAL development" (Lambert 2009b:12).

objective of further developing and enhancing the defence forces information capability (Oxenham et.al. 2003). In order to position his own work within DSTO's research objectives Lambert appropriated the notion of an automated Information Fusion system and transformed it into a vehicle for applying the core ideas of his Post Classical paradigm. In practical terms, communicating this vision was greatly aided by the implementation of many of Lambert's ideas in the Future Operations Centre Analysis Laboratory (FOCAL). The FOCAL project served as a test bed for a range of technologies needed to implement Lambert's vision of intelligent, computer-generated virtual military advisers assisting human military personnel to carry out a Command and Control role. When the content of Lambert's work is examined (§8.2) it will become evident that Lambert's case for extending the capacity of information fusion systems into the domain of high-level cognitive tasks traditionally performed by military analysts rests largely upon the principles and ideas developed in *EMWC*.

The linkage between research in Artificial Intelligence and research into automated information systems designed to promote situation awareness has its theoretical roots in the work of McCarthy on developing systems with a commonsense capacity to respond to context (McCarthy 1987; McCarthy 1989; Kokar et.al.2009). Lambert's formal approach to representing situations is a process-inspired modification of McCarthy's formal approach to representing context. As recounted in the previous chapters, Lambert's Post-Classical paradigm appropriates many of the logical tools of Classical AI but grounds them in a post-classical metaphysics of processes, and a re-worked philosophy of belief; these innovations were designed to overcome the brittleness problems that McCarthy was also attempting to address. For the process theorist the world is made up of processes. A situation (or context) represents only a limited fragment of the world; hence situations are fragments of process. A semantic theory must be able to specify what is true in a given situation, which in the language of *EMWC* means that template types must be able to provide formal criteria for specifying that an occurrence is an instance of a given type (a pattern matching exercise, as described in chapter seven). Lambert's work in *EMWC* was thus oriented to making sense of situations. It is this Post-Classical concept of AI that lies behind Lambert's later work on information fusion.

Lambert's notion of information fusion is an adaptation of the model of data fusion developed by the Joint Directors of Laboratories - referred to as the JDL model (Lambert 1999; 2003; 2009a). Most defence related research on the topic of data fusion has taken the JDL model as its starting point (Llinas et.al. 2004). The JDL model of data fusion was an attempt to generalize from techniques developed for automatically combining information from sensors (e.g. radar) relating to the movement of objects (e.g. tracking aircraft), so as to outline a procedure that would encompass automated methods of generating and combining data and information from multiple sources. Although the JDL model involves a recursive refinement process, it is basically a linear vision of machine mediated knowledge acquisition that begins with the collection of objective data, uses this data to assemble a holistic view of the situation (facilitating situation awareness), and from this view of the situation generates a range of more or less desirable potential scenarios (facilitating threat assessment).

The JDL model includes the 'higher-level' forms of data fusion (situation and threat assessment), but the primary focus of research has been on assembling information about the movement of physical objects. Lambert's re-worked version of the JDL model not only shifts the focus of development to the higher-levels, it also considers a wider range of objectives than those that were the focus of traditional military planning. Recent developments in the national security context, including the growth of international terrorism, peacekeeping and nation-building roles, in combination with the Australian Government's adoption of the Network Centric Warfare (NCW) policy have increased the demand for more sophisticated forms of information fusion capacities that are capable of dealing with information sources and data bases covering a wide range of phenomena. This demand is reflected in Lambert's interpretation of 'higher-level' fusion.

In Lambert's work, higher-level information fusion is extended to encompass any phenomena that could be of interest to the military in its expanded role. To use Lambert's preferred definition, higher-level information fusion is "the process of utilising one or more data sources over time to assemble a representation of aspects of interest in an environment" (Lambert 2009a:6). Lambert notes that 'aspects of interest in the environment' includes issues pertaining to "biography, economy, society, transport and telecommunications, geography, and politics, in addition to combinations of the aforementioned" (Lambert 2009a:6). Thus, the 'objects' that are the focus of Lambert's fusion process can encompass any facts – including psychological and socio-cultural facts - that are relevant to making sense of a situation. These objects are the basis of Lambert's revised version of machine information fusion, and enter into each of the three main levels of assessment:

- *object assessments*, are representations of *objects* in the world;
- *situation assessments* are representations of *relations between objects*; and
- *impact assessments* are representations of *effects of relations between objects* (Lambert 2009a:8).

Also crucial to Lambert's revised notion of automated information fusion is the need to explicitly incorporate the role of humans into the model of distributed, machine based data fusion; that is, machines and humans are both participants in and objects of the information fusion process. The advent of hybrid societies, incorporating a combination of machine and human decision making processes, means that the notion of data fusion must reflect the human and machine modes of assessment, plus the combined mode of assessment arising from the interaction of people and machines. In Lambert's words:

> . . . there is a need for the concept of data fusion to extend beyond machine based data fusion. In general, the demand is for fusion involving both people and machines, and so combines: an environment of interest; machines that sense that environment of interest; machines that further refine the sensed data and other information; mechanisms for integrating the refined information with people; and the fusion activity undertaken by people, some of whom may also independently sense the environment of interest. The resulting higher-level fusion system considers fusion from the standpoint of machines; fusion from the standpoint of people; and fusion from the standpoint of the integration of the two (Lambert 2009a:7).

Lambert's vision of the integration of machine-based information fusion with human information fusion is informed by Endsley's (1995) exploration of the role of situation awareness in human decision making. Endsley defines Situation Awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"

(Endsley 1995). Lambert notes that situation awareness is “not a computer system or a screen display - it is a state of human awareness” (Lambert, 2001b). Situation assessment is thus the human *perception* of the current situation, *comprehension* of its likely consequences, and *projection* of interventions that would generate alternative scenarios.

Lambert argues that Endsley’s understanding of situation assessment can be integrated with a revised JDL model of high-level machine fusion to form a hybrid model of information fusion. In this hybrid model, human perception, comprehension and projection are regarded as roughly equivalent to machine functions associated with object assessment, situation assessment and impact assessment (Lambert 2001b). However, this does not mean machines achieve ‘situation awareness’. Machines can generate representations and semantic classifications, but this does not imply awareness. *Situation awareness is always a state of human awareness*, but as with other forms of symbolic media, machine generated outputs can contribute to this human state of awareness. From this standpoint, high-level machine information fusion can shape the interaction between machines and humans in a manner that contributes to human situation awareness. This interaction takes place at the three functional levels common to human awareness and Lambert’s revised JDL model of machine fusion.[2]

Lambert’s perspective on machine and human information fusion and situation awareness brings to the fore the importance of the machine’s capacity to present information in a manner that engages the intent of the human actor. This means shifting from a data driven view of information fusion to a goal driven perspective.[3] From Lambert’s process perspective machine-based data fusion (in the form of situation and impact assessments) can contribute to both the comprehension and projection elements of the mind’s situation awareness if it engages our intent. Data, if it is to be relevant, must be presented in a way that engages our interest. In the past, the use of mechanical aids to support shared awareness has been limited to visual displays, such as the traditional ‘pins on map’ representation of a tactical situation. Even in the current era, as Lambert (1999a) has noted, the dominant US approach to situation awareness has remained focused on the “dissemination and display of perception-level data”. His own approach to fostering shared situation awareness

---

[2] In developing his own Post-Classical approach to machine based information fusion Lambert argues that the classic JDL model presumed an Aristotelian world of objects that failed to take adequate account of the relational dimension of events; likewise the JDL model failed to take into account the essential role of higher-level situation and impact assessments in mediating our selection and perception of data at the level of object assessment. In a loop analogous to the ‘non-well founded’ equilibrium (between what there is, what is believed and what is represented) that underpins Lambert’s ‘semantics of success’ (cf. chapter 7), Lambert argues that the three levels of his revised JDL model of machine data fusion are linked together in a relationship of circular dependency, such that the system “observes to explain; explains to predict; and predicts to observe” (Lambert 2009a). This is an expression of the insight that what one observes in part depends upon what one looks for, and what one looks for is shaped by one’s intent, and that intent is in part shaped by what we observe.

[3] Lambert has observed that:

> The bulk of the data fusion community remain heirs to the traditions and techniques of sensor fusion, which is primarily concerned with the use of sensors to form object assessments. This has engendered a data driven view of data fusion, under which a data fusion process is understood as a process that assembles assessments of the environment, based on the data provided to it. An alternative perspective is to embrace a goal driven view of data fusion. Data fusion is then understood as a process that assembles assessments of the environment, based on its goals. In fusion terms, this represents a transition from a level 1 centric view of data fusion to a level 3 centric view of data fusion. (Lambert 2009a:18)

involves focusing on "the story behind the data", through the use of software that enables "the machine to generate stories from its accessible information" (Lambert & Scholz 2005:13-14; 29).

The effort to engineer a story-telling capability reflects Lambert's assumption that a machine generated information system that seeks "to support the comprehension and projection levels of situation awareness needs to be highly effective at establishing, maintaining, and shifting context" (Wark & Lambert 2007:1). Generating gigabytes of crucial information is of little value if it fails to grasp the attention of the appropriate person. In public broadcasting, for example, a key function of news readers is to establish a context for news items that engages the interest of the listener. Similarly, Lambert's hybrid information fusion system envisages virtual news readers that can tailor individual presentations to suit the needs and interests of individual users while still promoting a shared awareness or 'common operation picture' (COP) across the distributed network of human users. In practical terms, this means developing a machine capacity to integrate information and deliver it in the form of persuasive stories (supported by maps, videos, graphics, etc.) that are relevant to the goals of an organization, but tailored to the needs and interests of particular departments or divisions, and of the individual roles in the various sections of the organization.[4]

Shared situation awareness thus depends upon establishing contexts that appeal to the various interests that motivate the members of an organization. To engineer a machine with a functional capacity akin to grasping and telling a story is an extension of themes explored by Lambert in his investigation of engineering commonsense. The Post-Classical perspective developed in *EMWC* provided Lambert with a philosophy, a vision, and the outline of a practical engineering procedure for extending machine reasoning into this fundamental realm of human activity. Lambert's ρbot and its ATTITUDE multi-agent successor are implementations of this design philosophy, in that their architecture reflects principles derived from Lambert's critical analysis of the way in which customary understandings, habits and routines motivate, orient and dynamically steer human responses to situations.

In *EMWC*, Lambert discloses an acute awareness of the philosophical and theoretical contexts informing AI research; however, behind his highly abstract and speculative reflections on general principles, was a practical engineering background that has played a significant role in his effort to transform his Post-Classical paradigm into a

---

[4] Stories, from the standpoint of *EMWC*, are symbolic templates that are modified and reused, as the need arises, by joining together (lego style) the routine fragments that compose the story in a manner that relates the story to the new situation and achieves (satisfies) its goal. More will be said about the nature and role of story telling, both from a sociological and a computational perspective, in the envisaged socionic extension of this project. For now it is sufficient to note that Lambert likens machined based object assessment with establishing the data, and higher-level information fusion with determining the story behind the data. As Lambert (2001b) stated in an early summary of his ATTITUDE-based approach to situation awareness:

> Situation assessments emerge as stories about the world, couched in formal languages capable of expressing relationships between objects. [This paper develops a concept of situation . . .] as the basis by which we organise these stories. A theoretical process ontology is discussed for defining the language of situations and an existing [ATTITUDE] computational implementation of situations is illustrated.

This story telling orientation was central to the FOCAL project. As Wark and Lambert (2007) noted:

> One of the research programs in FOCAL is exploring the use of ECAs, dubbed virtual advisers, to act as automated storytellers, combining narrative with multimedia presentation, or 'multimedia narrative', similar to television news services.

new paradigm for Situation Awareness. *EMWC* was inspired in part by practical work carried out by Lambert while he was working for the Microwave Radar Division of DSTO.[5] During the early nineties DSTO worked in collaboration with Ericsson Defence Systems, developing an AI system for radar control, based upon a reactive planning architecture designed to "reduce some of problems associated with classical AI approaches" (Powis, Lambert, et.al. 1992:428). This work was subsequently developed to form the basis of the MetaCon reactive planner that was developed for control of phased array radars on the Swedish Airborne Early Warning aircraft. The strong influence of ideas presented in Lambert's *EMWC* work is evident in papers that deal with the upgrade of the MetaCon architecture to a multi-agent ATTITUDE architecture (Lambert and Relbe 1998), capable of dealing with epistemic and ontological uncertainty (Fabian & Lambert 1998) that would provide Situation Awareness for military personnel (Lambert 1999a) using an ATTITUDE based multi-agent architecture for intelligent virtual advisers (Lambert 1999b) incorporated into a system designed to provide what Lambert termed 'Ubiquitous Command and Control' or UC^2 (Lambert 1999c).[6]

UC^2 ('you see too' - the pun was intentional) is Lambert's vision of a new form of information system undergirding a decentralized, hybrid military organization. He argues that a shift is taking place in society from the hierarchical social organization typical of the early industrial era, to a more devolved form of decision making and resource management, with a correspondingly flattened organizational structure. The key to his UC^2 proposal is the development of machine agents with the capacity to collaborate in an intelligent fashion with human agents to deliver a heightened shared awareness of situations; heightened both in terms of a better understanding of imminent threats, and of the opportunities latent in a given situation to respond to a threat. The UC^2 ideology promotes automation through software agents that "tightly integrate fusion and resource management" and thereby improve the military's capacity to deliver capabilities where they are most needed and will achieve the greatest effect (Lambert 2009a:19).

Lambert's *EMWC* vision for engineering commonsense was at the core of the UC^2 proposal. The ATTITUDE multi-agent architecture was essentially a development of *EMWC's* ρbot proposal and could thus be viewed as an attempt to mechanically harness the principles that, according to Post-Classical AI, promote human commonsense. The resultant automated representational system delivers a reasoning capacity to Lambert's virtual military advisers, and facilitates the cooperative interaction between the multiplicity of distributed agents that combine to constitute the UC^2 system.

---

[5] Lambert acknowledges the importance of both a four year full-time research scholarship with Flinders University and the support of his current employer (DSTO), in undertaking and submitting the work outlined in *EMWC* (*EMWC*: 'Acknowledgments').

[6] According to Lambert (2009b) there have now been four different agent implementations based around *EMWC*:

- The Scheme code in Appendix D of *EMWC*;
- The Pascal and then Ada MetaCon code for the Swedish AEW;
- The Ada ATTITUDE code that extended MetaCon to multi-agency etc.;
- The Prolog ATTITUDE TOO code written in 2009 (not yet published).

Lambert's argument in favour of a UC^2 system is couched in terms of a process philosophical approach to changes in the historical global situation, as manifested in the changing function and structure of globally oriented organizations. Citing the popular work by Senge (1990) Lambert asserts that organizational identities should be explained in terms of underlying processes that mediate organizational adaptation. Adaptive processes guide the formation of organizational identity, not vice versa. An organization's adaptive action, when understood in this context, is the "utilization of *capability* to achieve *intent*, given *awareness*" (Lambert & Scholz, 2007:160). Lambert notes that while there is a mutual tension between the nodes of this action trinity, intent is the foremost element, in that it establishes future goals that orient action. Capability is linked to intent via *capability options,* which are defined in a manner that resonates with the Gibsonian notion of 'affordances' as "anything that has the capacity to change one's awareness of the world, typically by changing the world" (*ibid*). Mission intent, situation awareness and capability options are thus linked in a relation of circular dependency, with action at the core of the trinity, as represented in figure 8.1.

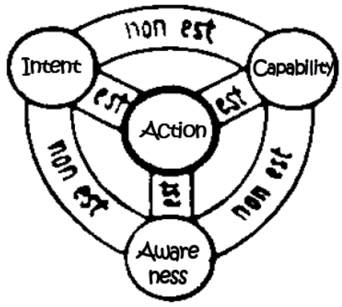


Figure 8.1: Lambert's Action Trinity (Lambert and Scholz, 2007:159)

Lambert's vision of intelligent virtual advisers underpinning a UC^2 system was subsequently put into practical effect by a multi-disciplinary team working under the auspices of the Future Operations Centre Analysis Laboratory (FOCAL). The goal of the FOCAL project was to "pioneer a paradigm shift in command environments through a superior use of capability and greater situation awareness" (Estival et.al. 2003). To that end, Lambert brought together a team of over twenty individuals with expertise ranging across engineering, physics, philosophy, mathematics, computer science, linguistics, human computer interaction and psychology, and employed them to build a virtual command and control centre. The centre piece of the FOCAL laboratory was a semi-immersive virtual reality environment that utilized cinematic technology, animations, and a natural language system to enable virtual advisers to interact directly with humans (see figure 8.2). This laboratory was used to explore the feasibility of Lambert's vision of virtual advisers interfacing with humans in the operation of a highly automated information fusion system.

Fig 8.2: Spherical screen used in the FOCAL laboratory (Wark & Lambert 2007:2).

Thus, FOCAL provided a platform to demonstrate the progress made and challenges still to be met in implementing Lambert's vision of a new form of hybrid command and control system, based upon an integrated society of machine and human agents. Various technologies ancillary to information fusion have been trialled under FOCAL (Andrews et.al. 2006), and Lambert has given further thought to the broader questions relating to the role of information fusion in a networked military organization (Lambert & Scholz 2005). However Lambert's major effort has continued to be directed toward engineering the Post-Classical vision of a 'commonsense' machine that underpins his blueprint for developing an information fusion system (Lambert 2001b; 2001c; 2003a; 2003b; 2003c; 2006a; 2006b; 2007a; 2007b; 2009a; Lambert & Nowak 2006; Lambert & Nowak 2008; Lambert, Wark, et.al. 2009; Nowak 2001; 2003; Nowak & Lambert 2005).

Lambert's endeavour to apply his research vision in a domain where practical outcomes are the measure of success has carried some costs. *EMWC's* substantial philosophical contribution toward framing the broader paradigm issues facing artificial intelligence has been followed up. Within DSTO, obtaining funding for original applied research, let alone funding for philosophical forays into paradigm assumptions, is subject to competitive pressures. Lambert's response to these pressures was to frame his research in terms of the militaries NCW policy. This institutional context has placed constraints on Lambert's approach to information fusion. The demand of the military for practical technology has tended to push information fusion development in all countries, including the USA, toward utilizing existing off-the-shelf software, and adapting that to the requirements of the military, rather than engaging in the risky business of developing new technology.[7] Lambert has been subject to these pressures, and has taken advantage of off-the-shelf software

[7] For example, in 2003, the same year that Australia launched its Network Centric Warfare policy, the US Army embarked on an ambitious, two to three hundred billion dollar Future Combat Systems (FCS) program that is designed to transform its army into a robotic fighting force. At the heart of this robotic force is the FCS network, described by the Army as "a complex communication system with one common cause: Deliver a platform for data and knowledge exchange that enables forces to dominate land combat". The network is being assembled from existing software, in the hope of avoiding costly and time consuming delays that have so often hindered the implementation of entirely new software systems within military platforms. See the US Army FCS website - https://www.fcs.army.mil/systems/index.html

where it is useful;[8] however, behind the emphasis on obtaining results there is clearly a continued commitment to the process-based approach to engineering commonsense outlined in *EMWC*.[9]

Thus, despite constraint, Lambert retains a strong commitment to realizing his Post-Classical vision. Lambert's information fusion system is more than an innovative piece of technology; it is the exemplification of a new scientific paradigm, the expression of a novel set of philosophical presuppositions, harnessed to a vision of a practical artefact. The vision behind the practical engineering ambitions is not always easy to see, however, in the light of a reading *EMWC* it becomes quite transparent. This can be readily illustrated with reference to Lambert's (2009a) paper, "A blue-print for higher-level information fusion systems"; this paper confirms that the developments over the last decade have remained embedded within his Post-Classical conceptualization.[10] In the next section I will use the 'blue-print' paper to frame my own summary of Lambert's progress in implementing the Post-Classical vision developed in *EMWC*.

8.2 Implementing Post-Classical AI in a higher-level information fusion system

Lambert's Post-Classical conceptualization was an effort to philosophically re-frame artificial intelligence so as to avoid the rationalism, solipsism and linguistic realism of Cartesian AI, and "deliver an alternative functional architecture for commonsense machines" (*EMWC:*133). To that end, Lambert developed a four step argument, beginning with a metaphysical theory of existence that was used to ground an alternative approach to representation, belief and mind. As previously recounted, he first re-conceptualized the ontological foundations of AI in terms of process metaphysics. Next he suggested that repeated patterns of influence between processes are the foundation of commonsense representation, and showed how patterns of influence in the world could be represented using state-transition graphs. Then, on the basis of this process-based approach to representation he introduced a semi-realist account of belief; beliefs are characterized using a formal semantics that enables each represented belief to associate a pattern of influence with a process in the world. Finally, Lambert pulled together his Post-Classical approach to AI by providing an account of mind as a multi-agent architecture, in which each individual agent is viewed as an evolving process, drawing upon its own specialized store of routines and

[8] As is evident in recent commercial agreement between DSTO and the independent Australian ICT research consortium, NICTA, aimed at adapting an off-the shelf first order theorem prover, and a Descriptive Logic (DL) reasoner (see Baader et.al.2008). The DSTO information fusion project is part of a wider collaboration between allied nations, such as the UK, USA and Canada (see Lambert, Wark, et.al. 2009) and Lambert's research effort is part of the Australian effort to develop an information fusion system that complements and is compatible with technologies being developed by these strategic partners (see for example, the recommendations in Lambert, Saulwick, et.al., 2008).

[9] Richer (2006) notes some of the conflicts that have arisen over the need to justify research and pressure to deliver capability in NCW related research and development. Constraints of this sort encourage projects that build upon existing technologies and procedures to produce rapid, demonstrable capacities, rather than innovative philosophically framed research such as outlined in Lambert's *EMWC*.

[10] Lambert (2009a) was written in 2007; its contents are amplified in Lambert, 2007b and Lambert & Nowak 2008. Lambert, Wark, et.al. (2009), provides an overview of recent developments, as presented at a conference in Seattle, in July 2009.

beliefs to adaptively interact with its environment (including other agents) so as to transform its own state and the state of the world, in conformity with its goals. As I shall show, this Post-Classical vision has framed Lambert's vision of information fusion and underpins his approach to engineering this technology.

The continuity between *EMWC* and Lambert's (2009a) blue-print for information fusion can be best exhibited by comparing the practical examples used to illustrate his ideas. In *EMWC* Lambert illustrates the Post-Classical notions of existence, representation, belief and agent reasoning by showing how set theory, model theory and state transition networks can be used to represent a process fragment or dynamic series of such fragments, such as the processes constitutive of a tree or a light switch. In *EMWC,* Lambert builds upon these principles to suggest an architecture for ρbot. Lambert's (2009a) blue-print illustrates the same principles but uses examples with a military flavor. Sets, model theory and state transition graphs are used to show how a multi-agent system could identify and assess military situations. In the latter analysis Lambert does not refer to the role that *EMWC* has played in the development of his ideas, but the same key elements are present.[11] Hence, in this brief review of the intellectual genesis and subsequent evolution of Lambert's project, I will touch upon the main elements of Lambert's Post-Classical paradigm, and show how these elements continue to structure Lambert's approach to information fusion and situation awareness.

8.21 A Philosophy of Existence - process ontology – the world as fragments of process.

From the launch of his new paradigm of information fusion and situation assessment in 1999, Lambert has promoted the view that situations are fragments of the world that can be represented by sets of assertions (Lambert 1999a, 1999b & 1999c through to Lambert 2009). These fragments of the world are understood "through an extensional concept of process, where a process is any spatio-temporal fragment of the universe" (Lambert 2006a). The genesis of Lambert's process ontology can be followed through papers that reiterate his critique of the Aristotelian substance oriented ontology and defend Heraclitus's process orientation as a preliminary to the continued elaboration and development of the process ontology first presented in *EMWC.* The link to Lambert's theory of processes in *EMWC* is explicitly referred to in Lambert (2001b) and in Nowak (2001; 2003); in subsequent papers Lambert then cites these earlier papers.[12]

The key paper dealing with Lambert's 'philosophy of existence' is Lambert (2006).[13] In this paper he presents the arguments developed in *EMWC*, including his critique

---

[11] The essential continuity in Lambert's project does not mean that there has not been change and development. As will be discussed in the next section, this is particularly evident in the conceptualization of the process ontology grounding Lambert's system.

[12] Lambert and Nowak also explore the practicality of implementing the FOCAL vision using various off-the-shelf products (Nowak & Lambert, 2005), but the process orientation is clearly Lambert's preferred option.

[13] 'Philosophy of Existence' does not imply existential ruminations of the Continental variety; rather, Lambert's concerns are framed by the concerns of analytic philosophers. Logical questions concerning existential quantification, and other considerations underpinning his formal ontology of processes, are the focus of Lambert's philosophy of existence; this is followed by an outline of his formal theory under the heading of a 'Mathematics of Existence' (Lambert 2006).

and modification of the Zermelo-Fraenkel Set Theory with the Axiom of Choice (ZFC). He also cites Whitehead's *Process and Reality*, before plunging into his formal exposition of his theory of existence. He notes that the ZFC version of set theory promotes an Aristotelian outlook and modifies the axioms of ZFC to present his own extensional set theory (mereology) in the 'nominalist' tradition of Lesniewski, Goodman and Lewis. As discussed in *EMWC*, this nominalism is an important complement to his process philosophy for it separates meaning content from ontological commitment, allowing a formal system to reason about process fragments that may not exist.[14] He then presents a similar formal theory of fragments of process as detailed in *EMWC*, and suggests how it should be extended to encompass space and time. Finally, he suggests how the formal metaphysical ontology, expressed as a Boolean algebra, could be integrated mathematically with the hierarchy of semantic theories that make up his expanded process conceptualization of the world. This hierarchy of semantic theories represents a significant innovation and supplementation to the ideas expressed in *EMWC*, but it is still within the spirit of his Post-Classical paradigm, and in Lambert's words, has evolved "according to philosophical standpoints like nominalism and process philosophy" (Lambert 2009a:9).

The various levels in Lambert's process ontology are first introduced in Nowak (2003), building on earlier suggestions made by Lambert (2003c) at a presentation given in 2001.[15] Lambert appears to have adopted these ontological levels as a way of extending the capability of the templates that were the focus of *EMWC*. As noted in my concluding observations in chapter seven, by treating causality as homogeneous 'patterns of influence' *EMWC* obscured significant distinctions that philosophers such as Whitehead have sought to clarify. Influence works in a variety of ways, and it is obviously important when dealing with strategic considerations to be aware of the various forms of influence that can bring about change (the military has traditionally focused on 'kinetic' factors, but this is but one element in broader strategic picture).

Lambert's post *EMWC* work deals with the notion of influence by rejecting the notion of causality as an explicit primitive concept and instead introducing various notions of causal influence implicitly, via a multi-level ontology. This ontology provides a rich array of concepts for indicating the nature of the patterns of influence constraining and enabling action. A significant influence in choosing this path was the work of Daniel Dennett.

The trigger for the development of Lambert's revised approach to ontology appears to have been questions that arose through consideration of the 'impact assessment' level of his revised JDL model. Impact fusion involves making assessments about how changing situations will impact upon the intentions guiding the action of cooperating and competing agents (friends and enemies). This means fusing several levels of

---

[14] Lambert's existential process metaphysics is a nominalist, mathematical (formal) theory of existence that indicates what is meant by existence, but does not specify what exists. Thus it can refer to things that do not exist. By separating content and ontological commitment it is possible to reason about entities whose existence has not been established. This enables the linguistic content of a statement to be defined in intensional terms, without entailing the existence of a corresponding identity (see my summary in chapter four and also Lambert 2006:2-3).

[15] This 'Mephisto' ontology is further developed by Lambert (2003a), Nowak and Lambert (2005), while Lambert and Nowak (2006) provides the most complete exposition. It remains foundational in Lambert's model of information fusion (Lambert 2009a).

explanation. Dennett (1971) argued that our predictions are grounded in three main forms of explanation: physical causes (laws of physics), functional design (what artifacts were designed to do), and intentional goals (what people intended to make happen).[16] Lambert's reflections upon Dennett's observation that we draw upon physical, functional and intentional explanations in dealing with our world, seems to have provided the impetus to extending the process ontology outlined in *EMWC*.

Lambert notes that the intentional level of explanation is crucial to impact assessment, and that it proceeds by "ascribing routines, beliefs, intentions and other mental states to individuals and then predicting their behavior" (Lambert 2003c:14). He observes that the ATTITUDE agent architecture makes these routines, beliefs and intentions explicit, and suggests in ATTITUDE mediated information fusion the "impact of a given situation must be assessed against current and future intentions, given the routines at our disposal" (*ibid*:14-15). This focus on intent opened up a rich ontological vein.

A decade before his (2009a) blue-print paper, Lambert (1999a, b & c) had spelt out the main outlines of his approach to information fusion. In order to accomplish higher-level information fusion, Lambert's ATTITUDE agents need to build up templates (routines) that reflect their world. Whether or not we adopt a process or substance based view of the world, machine communication with humans will only be possible if machines can identify, represent and incorporate into their calculations the phenomena that we regard as the constitutive *facts* of our human situation.[17] The major development since 1999 has focused on providing the formal specification of the objects and relations that constitute the various levels of meaning that we use to make sense of our situation. The aim of this endeavor is to create a conceptual framework that can provide a precise logical representation of the flow of intent that moves human action, and likewise enable machines to represent and participate in communicating this flow of intent in a manner that makes sense to human agents.

Superficially, this would appear to signal a return to the manifest realism and linguistic concerns that shaped Classical AI, however this is not the case. Implicit in every natural language are metaphysical intuitions; as in the case of the European tradition, these intuitions are not resolved into a coherent and rationally complete world view. On the contrary, semantic tensions and contradictions invade every area of human understanding. Whitehead's response to this situation was to urge philosophers and scientists to self-consciously and critically examine their

---

[16] The following examples illustrate the respective physical, functional and intentional modes of prediction: predicting where the snooker ball will go by observing the alignment of the cue and ball; predicting the cause of a car breaking down by following a trouble shooting guide-book; predicting how a business competitor will behave, given their commitment to making profits and their rational disposition to take advantage of available opportunities (See Lambert 2003c:14; Dennett 1971).

[17] Facts are termed 'events' in the ATTITUDE architecture and combinations of events are referred to as scenarios. 'Situations', in Lambert's technical sense, can refer to any combination of scenarios and events. An ATTITUDE avatar, in expressing situation awareness, is an inference engine oriented to situations. Each agent is programmed (through the routines in its long term memory) to deal with a certain types of objects and relations. A formal theory specifies the logical dependency between these objects and relation types. Routines (embodying the agent's internal representation of patterns of change in its world) must be specified that are capable of identifying both significant potential events and the agent's potential involvement with these events. Combinations of these events will generate all the possible situations that the agent could identify or anticipate (Lambert 1999b:117).

metaphysical presuppositions, and seek to create a more comprehensive and coherent framework of meaning. Mathematics was an essential partner in this endeavor. This was the issue that divided Whitehead and Dewey as related in chapter three (§3.4). The issue is fundamental to the construction of Lambert's ontology.

Dewey had argued that Whitehead needed to choose between a 'genetic-functional' and a 'mathematical-formal' interpretation of first principles. He approved of Whitehead's endeavor to create a critical philosophy of experience, but rejected Whitehead's attempt to transcend this fertile experiential ground via increasingly abstract and formalized structures of meaning, that aspire toward "the kind of structure exhibited in pure mathematics" (Dewey 1951:657). He feared that this entailed reifying meaning; ripping truth from its real experiential grounds, and making it a captive of a logical construction. In his own words, he feared that in attempting to formalize his metaphysical system, that Whitehead was substituting "abstract logical connectedness" for "concrete existential temporal connectedness" (Dewey 1951:658). Whitehead, on the other hand, believed that symbolic logic was an ally in the quest for more satisfying modes of understanding. He believed that due to vagueness and variations of meaning "conventional English was the twin sister to barren thought", and that the "method of algebra embodies the greatest discovery for the partial remedy of defective language" (Whitehead 1937:96). His speculative metaphysics was grounded in experience, but he sought to derive from experience general principles a more adequate system of ideas, in terms of which every element of experience could be interpreted. Symbolic logic was an essential ally in this process; through logic it was possible to examine the patterns underlying the accidental forms of experience. Belief in the power of mathematics to disclose patterns that escape unaided human intuition was essential to Whitehead's orientation.[18]

Whitehead's faith in the power of algebra (or symbolic logic – he used the terms interchangeably) to remedy the shortcomings of natural language was tied up with developments that he had helped define. The essential method of algebra, he argued, what that of selecting a few basic connectives, such as 'and', 'or', 'equivalent to',

[18] Half a century later the chief architect of the modern theory of social systems and Whitehead's most influential sociological disciple, Talcott Parsons, had the same argument with Alfred Schutz and Clifford Geertz. Toward the end of his life, Parsons came under a sustained attack by those defending the notion that our understanding of human life is constrained by natural language understandings of human experience. Parsons' critics argued that if an anthropologist or sociologist seeks to understand a community or society, they must do so within the meanings that the subjects in question employ; they must respect the unique intentionality of the people, as framed in their system of understanding. Parsons, on the other hand, followed the stance of the 'hermeneutics of suspicion', arguing that the theorist should seek a more adequate system of concepts than those implicit in the natural languages of the people concerned. Parsons, who had devoted his life to defining a simple scheme of concepts for dealing with all forms of social action, was firmly committed to the idea that human understanding could advance, and that developing formal models (though at a preliminary stage in sociology) was a powerful tool for promoting this advance. "The study of everyday life", he wrote, "simply cannot be the key to the understanding of the human condition. . . . Somehow or other it seems to be extraordinarily difficult for modern social thought to get away from the proposition that the developmental origins have a greater significance and a greater soundness than the subsequent evolutionary developments [in our modes of understanding]" (Parsons 1981:36). Those that take the side of Geertz and Schutz argue that a more adequate understanding of the human condition involves a descriptive multiplicative approach that catalogues the ever-expanding diversity of forms of communal self-understanding. Those following Parsons pursue a prescriptive reductive approach that seeks to discern within the diverse accidents of human history unifying principles that ideally can be framed in mathematical terms.

'more than', 'less than', and then applying them to symbols indicating 'real variables' to build patterns and patterns of patterns. According to Whitehead, the "basic connectives are the relevant mathematical-formal principles", while the "real variables are the unspecified accidental factors" (Whitehead 1937:97). According to Whitehead the power of algebra lies in its capacity to generate meaningful compositions ("apart from composition there is no meaning"); it allows us to symbolically examine patterns generated by combining connectives and real variables. Whitehead was acutely aware that this process was fallible; that we cannot "complete the approximation of disengaging the principles from the accidents of their exemplification", but he remained convinced that symbolic logic would provide a vehicle for partially remedying the deficiencies of language and enable meaningful analysis of areas extending far beyond the traditional focus of mathematical analysis (Whitehead 1937:97-9). Lambert's ontology builds upon this algebraic tradition; it contains a relatively small number of primitive relations, constants and functions, and carefully defines further key concepts on the basis of these primitives.

In *EMWC*, Lambert demonstrated the power of mathematics to generate meaningful compositions. However, while his representation of the time-space dimension of pattern was tightly constrained, attaching qualitative meaning to those bare patterns of process involved an element of improvisation. The choice of Mephisto primitive terms was arrived at by Nowak and Lambert and rests on "a presumption that a small number of constructs is sufficient to describe . . . the various kinds of phenomena that constitute our reality" (Lambert 2009b:11). What was required, Nowak and Lambert decided, was an extension of the primitive terms defined in the Metaphysical ontology, so as to provide interpretive models that are as mathematically rigorous as the metaphysical model, but capable of dealing systematically with those domains of phenomena in our world that shape human action. That is, capable of dealing with those levels that Husserl referred to as 'material ontologies'.[19] Lambert, together with Nowak, has attempted to realize this ambition through what they refer to as the 'Mephisto' framework (Lambert & Nowak 2006).

In ontological terms Mephisto is reductionist rather than multiplicative, prescriptive rather than descriptive, and formal rather than informal (Lambert, Saulwick et.al. 2008:4). The Mephisto conceptualization assumes that the world is made up of processes and that these processes should be approached through five different levels of abstraction, representing metaphysical, physical, functional, intentional and social processes. Together, these levels form a hierarchy of dependent concepts that provide a logically precise language for reasoning about the entities and activities that define our world. The focus of each layer of this conceptual hierarchy is as follows:

[19] Lambert's colleague, Chris Nowak has suggested that Lambert's physical, functional and intentional layers correspond to ontological levels that have been defined as the physical, symbolic and cognitive (Nowak 2003:663). This would suggest parallels between the formal and material ontologies of Edmund Husserl and Lambert's system, with its formal foundation in process metaphysics, and a superstructure of material ontologies. In a loose sense formal ontology concerns analytic a priori judgments (eg. formal eidetic domains such as defined by abstract parts and wholes and relations of connection between them), and material ontology concerns synthetic a priori judgments (eg. empirical domains such as referred to by traditional classifications such as body (physical things), mind (intentional acts), and spirit (content of symbolic objects and sense of cultural forms of representation). For a lucid introduction to Husserl's approach to ontology, see Smith (2007:ch.4).

- The *Metaphysical Layer* introduces foundational concepts like existence, identity, space and time. This allows the machine to identify fragments of the environment of interest and to do so with respect to their spatiotemporal parts.
- The *Environmental Layer* introduces environmental properties and relations to the metaphysical parts. This allows the machine to ascribe attributes like temperature and weight to individuated parts, while identifying some parts as ocean, others as land, and so on.
- The *Functional Layer* considers the functionality of identified physical parts. Principal functional relations in a military context are the ability to sense, move, strike, attach (includes carry), inform and transform. These are sufficient to characterise: surveillance and reconnaissance, weapons, logistics, communications and engineering capabilities.
- The *Cognitive Layer* adds cognitive relations to the identified physical and functional spatiotemporal parts. The attribution of beliefs, intentions and other mental states is performed at the Cognitive Layer.
- Finally, the *Social Layer* introduces social constructs between the cognitive individuals. Concepts like authority and enemy prevail at the Social Layer (Lambert 2009a:9).

In terms of facilitating the function of a multi-agent system, the additional layers of the Mephisto ontology are a mechanism for adding content to the bare abstractions of the metaphysical level's representation of the world. The metaphysical layer is the extensional ontology of relational events that make up the extensive continuum (the connected regions of the brick-wall universe). These are Whitehead's actual entities stripped of qualitative content and regarded as bare interconnected regions; this dimension is the fundamental ground of uniformity upon which all else depends. Building upon the formal ontology of existence, the remaining levels are roughly correlated with the traditional material ontological domains associated with body, ego and society. The environmental and functional layers of Mephisto together prescribe potential physical properties, relations and functional capabilities to spatio-temporal patterns of process. The cognitive level adds cognitive relations to these physical process parts, such as beliefs and intentions. Finally, the social level introduces a social dimension, defined primarily in terms of groupings of individuals and specifications of the relations that can hold between them. Thus, successive levels of the Mephisto ontology introduce concepts necessary to define different kinds of 'affordances' – behavioral, cognitive and social – that perspectivally qualify patterns of process in the metaphysical Universe, thereby defining the environment of action for the Mephisto agents.

Thus, in the same sense outlined in *EMWC* with respect to process worlds, a Mephisto theory is a formal (mathematical) specification of these various abstract domains (ie. each level is specified in the same logically rigorous manner to constitute a model of a particular domain), and these theories play a crucial role in facilitating reasoning and the semantic interpretation of statements about each domain. Lambert's aim with Mephisto is to produce explanations that make sense to humans, hence his goal is to identify key abstract dimensions of the domain encompassed by human explanations and then formalize these primitive concepts in logic, so as to make the meanings precise and capable of being "implemented within a machine with a logical reasoner" (Lambert, Saulwick et.al. 2008:42). Figure 8.3 provides examples of type of constants found in each domain, and the properties, relations and functions used to define each level of analysis.

Thus Mephisto is a computational conceptualization in the sense outlined in chapter 4.1; as such it facilitates a family of Mephisto constructions. A Mephisto implementation is a computational implementation of a Mephisto theory in some software; a Mephisto agent employs the Mephisto ontology in its implemented form, and a Mephisto society is a society of Mephisto agents using the Mephisto ontology to understand their world and communicate with each other. In other words, if the Mephisto ontologies are used to define and implement the beliefs (routines) stored in Lambert's ATTITUDE agent's long-term memories then these agents can form a Mephisto society (Lambert & Nowak 2006).

	level				
	Metaphysical	Physical	Functional	Intentional	Social
domain	metaphysical processes	physical processes	functional processes	intentional processes	social processes
constants	UNIVERSE, NOTHING	RED-SEA	HMAS-ADELAIDE	US-PRESIDENT	UN
properties	exists, spatial, temporal,	solid, liquid, gas, land, water, air,	operational, disrupted, neutralised, destroyed,	individual, routine, belief,	individual, group, nation, government, coalition,
relations	identical, part-of, overlaps, before, connects, north, west,		senses, informs, attached, strikes, moves	achieves, performs, succeeds, fails, intends, desires, believes, expects, anticipates,	ally, enemy, neutral, owns, possesses,
functions	join, meet, complement, space, time, distance,	temperature, weight, length,	price, cost,	truthness, trustworthiness,	reliability,

Fig. 8.3 An early version of the Mephisto ontology (Nowak 2003: 663).

The various levels of the Mephisto conceptualization reflect a pragmatic process of classification driven by the need to categorize chunks of the world process in a formal scheme that could be implemented and trialed in a FOCAL demonstration of higher-level information fusion. As related in Nowak and Lambert (2005), the Mephisto ontology was first implemented using off-the-shelf products.[20] A screen shot of the OilEd ontology editor used to implement the Mephisto ontology in the OIL ontology framework is shown in figure 8.4. The resultant Mephisto ontology was used to interface with FOCAL's virtual animated advisers, enabling the laboratory to

[20] Nowak (2003) and Nowak and Lambert (2005) describe how the Mephisto conceptualization was implemented using the OIL ontology framework with SHIQ as the underlying description logic; they note that the OilEd ontology editor was used when constructing each level of the Mephisto ontology, and the Racer framework used for computing the description logic.

demonstrate the FOCAL system's capacity to draw upon the Mephisto ontology to formally represent various queries, decompose them into component questions that can be directed to relevant data bases or sensors, and reason about the information obtained to form a reply to the query.

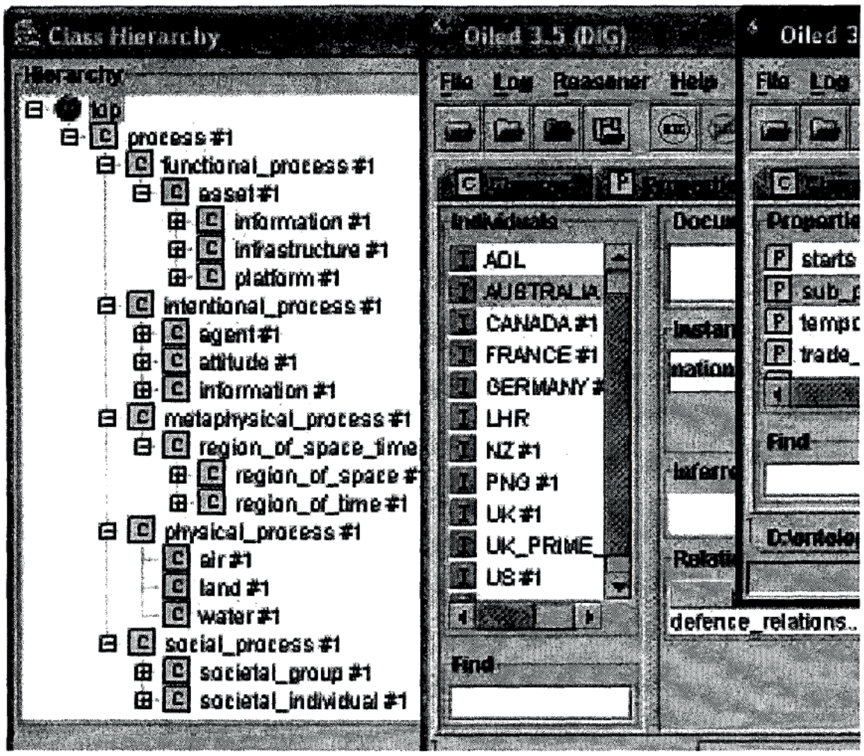


Fig. 8.4 Screen shot of Mephisto ontology in OilEd (Nowak 2003: 661)**.**

As is evident in figures 8.3 and 8.4, metaphysical processes remain fundamental and each domain is viewed as a fragment of process. Although the ontology is mathematically rigorous, it is not an endeavor to provide a rationally coherent and comprehensive account of the world, in the Cartesian sense. Rather, it provides a logical framework within which to formulate routines. These routines can be totally misguided in an epistemic sense; as in the case of *EMWC*, the aim is not realize the Cartesian ideal of rational omniscience but to instead capture the functionality of a habit oriented (and fallible) 'commonsense'. Thus, the Mephisto ontology is an extension of the commitment to process foundations expressed in *EMWC*. As will be seen in the next sections, the other main elements of Lambert's Post-Classical approach are equally important in his blue-print for information fusion.

8.22 Post-Classical representation, belief and mind: engineering semantic and cognitive machines.

Lambert's fusion system assumes that the world is made up of processes (events and scenarios) that can be represented using routines. Process fragments of the world exhibit patterns of change that can sometimes be matched against some combination of routines in human or machine memories. To accomplish this match-up between external event and template, the situations which routines represent must be defined in terms of the formal theories that constitute the Mephisto ontology. These theories provide the precise mathematical structure used to transform a well-formed sentence into a meaningful statement and similarly, to transform data picked up by sensors into meaningful classifications. The theories perform the role of Tarski's meta-language, as in the 'it is snowing' example discussed in the previous chapter. For example, if the fusion machine has syntactic tokens such as 'exist', 'moves', 'missile', 'target' and 'tower' for describing the data relating to missiles picked up by radar, formal theories with logics can be used to attach meanings to these tokens, so that the agent knows that the missile is targeting a tower, and can make further valid inferences.[21]

This way of endowing symbols with meaning draws inspiration from Barwise and Perry's (1983) situation theory but, as in the case of *EMWC*, modifies it to include both the agent and theorist perspectives. As represented in figure 8.5, Lambert's vision of how fragments of process acquire meaning is clearly continuous with his previous work, and with the Whiteheadian approach to semantics; the similarities with figures 3.4 and 6.2 in the previous chapters are obvious. The agent is able to draw on a cognitive template from its long term memory to arrive at a situation assessment which, from the theorist perspective is a formal theory inside the agent (with an extensional reference) but from the agent perspective is a direct experience of the missile targeting the tower. According to Lambert, this approach avoids the semantic problems associated with Frege's morning star/evening star problem and Putnam's twin earth problem (see my chapter sections 2.4 and 6.2) and is at the core of his implementation of semantic machines within the FOCAL framework.

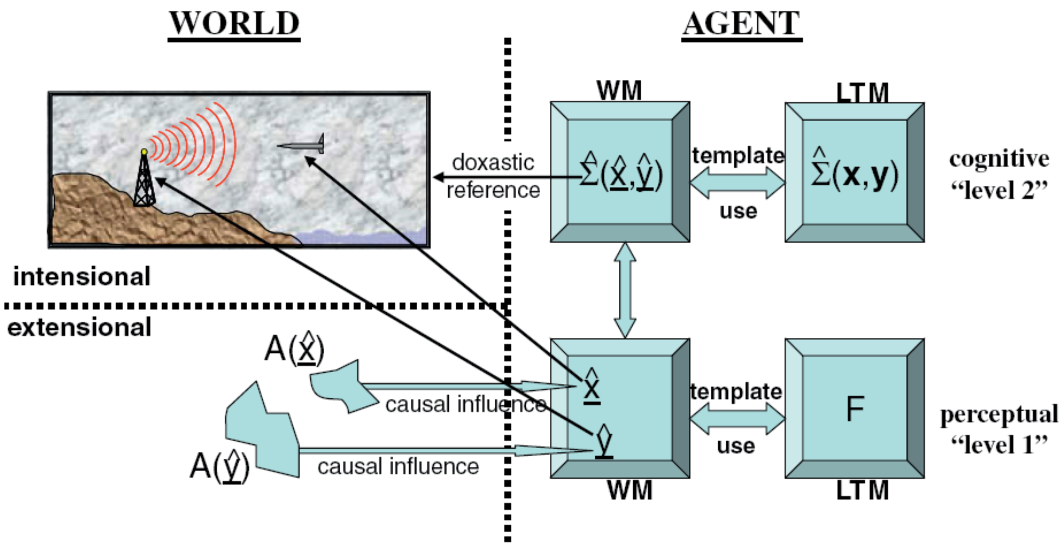


Fig. 8.5 Content defined by intensional reference (Lambert 2009a:11)

[21] Lambert (2009a) provides a very compact statement of the basic arguments presented in *EMCW*. He links together the formal theories in his ontology, his semantics, and the ATTITUDE cognitive agents and shows how they underpin situation and impact assessments. The role of formal theories is crucial. Lambert notes that a "formal theory *M*, together with logic $\langle L, \models\rangle$, can provide the basis for machine-based meaning. A formal theory *M* can be used to introduce a theory identifying what is meant by tokens like missile and targeting. By insisting that the machine only has interpretations $\Re$ that make *M* true, i.e. $\models_{\Re} M$, the machine can use the theory *M* to constrain allowable interpretations, and so constrain allowable referents for $\Re$(missile) and $\Re$(targeting)" (Lambert 2009a:10).

Lambert's semantic theory enables a machine to meaningfully represent situation and impact assessments, but a cognitive agent is needed to perform the comprehension and projection tasks at the heart of situation and impact assessments. Thus, as in *EMWC* Lambert's 2009a fusion paper next turns to task of outlining a multi-agent architecture designed to be situated in and interact with its environment. Thus, having reiterated the key elements of his Post Classical vision, from his process metaphysical foundations through to his philosophy of belief, he arrives at the point where he advances his suggested architecture for engineering a multi-agent mind with a *commonsense* capacity for dealing with the world or, in his current terminology, a 'cognitive machine'. This machine takes the form of his ATTITUDE multi-agent system (MAS).

8.23 Cognitive Machines

The ATTITUDE cognitive model sketched in Lambert (2009a) is detailed in Lambert (2003b), and goes back to Lambert (1999b) and these in turn were based upon the ρbot architecture outlined in *EMWC*. As in the case of ρbot, the ATTITUDE multi-agent system extends Lambert's automation principle through to its ultimate end: that of engineering machines that we are able to interact with as if they were humans. Just as we predict and explain human behaviour "by ascribing mental states like beliefs, desires and expectations to individuals; and making a presumption of rationality about their behaviour", so the ATTITUDE model applies this approach in the 'attitude programming' of software with propositional attitudes and propositional instructions (Lambert 2009a:12).[22]

The aim of attitude programming is to combine propositional goals and propositional attitude instructions, such that the successful execution of a routine from its start state through to its final state should mean the attainment of the routines' propositional goal. ATTITUDE routines (alluding to both human 'routines' and the computer idea of 'subroutines') are implemented as finite state machines (FSM)[23]; routines "are

[22] Although its initial store of routines is coded by a human programmer, Lambert envisages the ATTITUDE agent learning from its interactions with its environment or from other agents so as acquire new routines; how it executed this learning process would depend upon the learning routines stored in its long-term memory. Lambert notes that an ATTITUDE agent "can store whatever beliefs its routines contextually tell it to store, and these can be about anything real or imaginary, itself, its past performance or its future performance" (Lambert 2009a:13).

[23] Finite State Machines (FSMs) are a computational form of a state transition network. They have been used broadly in the video game industry. For example, a complex FSM system was the basis of the AI in *Warcraft III*, while other games have used FSMs to implement interactive plots, where the user chats with an avatar, using multiple choice prompts, to determine the way the game develops. Each FSM has a set of states that follow a certain path. A state has transitions to other states, which are triggered by events or actions within a state. An example of a simple FSM (see figure below), has as an initial state, a person before a closed door; they then open the door (action) which creates a new state. In this new state they can then close the door, returning the system to its original condition. Transition conditions are events, while the entry action is an event. Agre and Horswill (1997) use FSMs to implement their situated approach to life-world analysis.

template recipes for behaviour that are dynamically clipped together like lego™ blocks via pattern matching" (Lambert 2009a:12). Thus, as in *EMWC* the ATTITUDE agent's quest to satisfy its volitions is essentially a pattern matching exercise:

> Volitions are responsible for an agent's behavior and arise through interactions with the environment, interactions with other agents, or through autonomous internal goal setting by the agent. Intentions are independent volitions, while desires are subordinate volitions to other desires or intentions. When attempting to satisfy volitions, an agent will select a routine with a matching goal and begin execution of that routine in the matched context. This mechanism provides a dynamic lego™ block like clipping together of executing routines when cascading subordinate desires arise. To execute a routine, each successive propositional attitude instruction step(s) in a routine gives rise to an internal mental task(s) that typically modifies working memories and competes for execution time against concurrently executing routine steps (Lambert 2009a:13).

One aspect of the *EMWC* approach to propositional instructions that has become more important in the ATTITUDE agent architecture concerns the shift of orientation toward distributed applications and the interaction of autonomous agents in 'social environments'. Rather than think in terms of implementing a single agent ρbot, ATTITUDE routines are oriented toward facilitating cooperation between agents. This builds upon the *EMWC* approach, but interprets the subjects in propositional attitude instructions as external agents.

In *EMWC* propositional attitude instructions had the form <subject> <attitude> <proposition> (see chapter section 7.11). The key idea was that the subject in propositional attitude instructions is not restricted to observations about their own or other agent's dispositions; with subtle changes in language an observation can be changed to an instruction.[24] In ATTITUDE, the subjects (agents) can be explicitly addressed (as in ρbot) or they can be variables, meaning that the choice of subject in the routine can be contextually instantiated. As a result, agent routines may lead to patterns of interaction between agents, which may result in enduring patterns of social

---

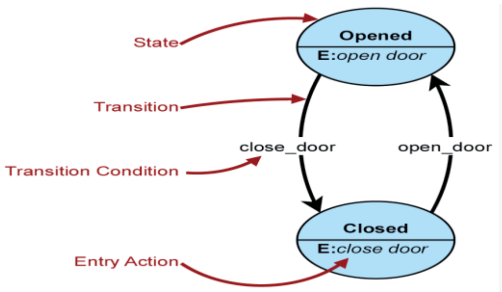


Diagram from Wikipedia - http://upload.wikimedia.org/wikipedia/en/9/94/Fsm_example_annotated.svg

[24] For example, a propositional attitude observations about Fred such as

Fred believes that the sky is blue

can be transformed into propositional attitude instructions to Fred like

Fred believe that the sky is blue (Lambert 2009a:12).

Instructions then lead to interactions between agents. For example, if agent Fred executed the instruction **Bill desire α** in one of Fred's cognitive routines, for some proposition α, then the effect is that the desire **Bill desire α** is be sent to Bill with Fred as the return address (Lambert 2009b:12). In operation, the ATTITUDE agent can store whatever beliefs its routines tell it to store, and these can be about anything, so that if a human agent asks the ATTITUDE agent what it believes, expects, anticipates, can do, desires, intends, sense or effects, it can tell you; alternatively it can instruct other ATTITUDE agents what to believe, expect, etc.

interaction. Hence, by mediating the flow of intent within each agent propositional attitudes serve as a mechanism for controlling cognitive behaviour, but they also provide a mechanism for instigating social behaviour between agents. Hence, Lambert notes that:

> Cognitive routines are thus recipes for cognitive behaviour, but when two or more agents execute cognitive routines that interact, then those same cognitive routines become social routines that guide interaction. Each player's cognitive routine is a contributing piece to a broader social routine (Lambert 2009b:12).

The possibility of agents interacting to create wider societies characterized by routine patterns of action is thus a latent capability of the *EMWC* approach to mind, but in the context of Lambert's model for information fusion, the social dimension acquires a greater significance. As indicated in figure 8.6, the ATTITUDE agent architecture has the same essential structural and functional layout as the ρbot agent architecture (see chapter 7, figure 7.2), with a control system over-seeing the operation of the long term memory (with its store of routines), and a working memory for dealing with the immediate context of interactions.[25] Again, summarizing this section of Lambert's (2009a) paper would simply reiterate for the most part the ρbot architecture outlined in the previous chapter. The key feature remains the notion of a dynamic environment, conceptualized in terms of processes that can be individuated in terms of patterns of activity that can be captured by agents as routines and stored as memories. The functioning of an agent begins with the initialization of a set of cognitive routines that establish anticipations in the agent as to what the world might be like. Thereafter, as Lambert notes,

> . . . the agent's behaviour will depend substantially on its interaction with the external world – either through its sensors or through communication with other agents. And it is the cognitive routines that are invoked when these anticipations are met that drive the changes to the various working memories and how the agent responds. Humean custom remains the driver (Lambert 2009b:11-2).

Thus Lambert's vision of cognitive agents builds upon his vision of a strategy for engineering commonsense, as outlined in *EMWC*. Process grounded routines remain fundamental; however, the horizon of action has been considerably enriched through the Mephisto vision. This has involved formally recognizing various ontological

---

[25] As noted previously, in 2009 Lambert began development of ATTITUDE TOO, a further refinement of the ATTITUDE architecture discussed in this chapter. This work has not yet been published, however in a communication to the author, Lambert (2009b) has stated that:

> It [the ATTITUDE TOO architecture], *inter alia*, additionally allows for single agents to be distributed across multiple platforms (anywhere in the world). . . . In general, ATTITUDE TOO is to provide the basis for cognition in interacting teams of virtual advisers (Lambert 2009b:11).

The system is now coded in Prolog, and its functions have been enhanced in a number of ways, several of which involve the operation of memory. Lambert notes that:

> One transformation from *EMWC* is the expansion of LTM in ATTITUDE TOO. It now comprises:
> - Semantic Memory (houses Mephisto constructs);
> - Epistemic Memory (mostly simple basic knowledge about the world of the agent e.g. which country borders which country, where the airports are, etc.);
> - Episodic Memory (holds all the cognitive routines) (Lambert 2009b:11).

Another significant innovation involves the modification of the memory process to include a Long Term Working Memory (LTWM) mediating between Working Memory (WM) and Long Term Memory (LTM). There are now wake and sleep functions associated with LTM, with wake loading the contents from LTM into LTWM, and sleep loading the contents of LTWM into LTM. The aim of these functions is to endow the agent with a "persistent life" (Lambert 2009b:12).

dimensions of agent action, including the social context. As will become evident in section 8.3 of this chapter, recognition of the social level has opened up considerations that extend beyond the cognitive considerations that were the focus of *EMWC*.

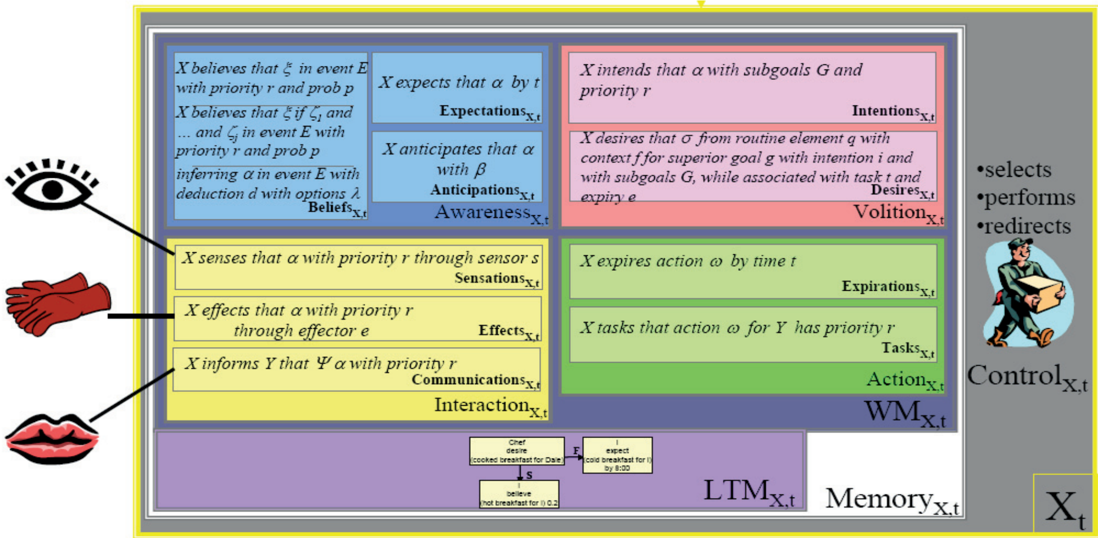


Fig. 8.6 ATTITUDE cognitive model (Lambert 2009a:12).[26]

Harnessing the capabilities of the ATTITUDE MAS to the Mephisto ontology is an ongoing project. Obviously a key challenge lies in developing the formal theories that agents like the ATTITUDE MAS use to give meaning to their world. Lambert's finite state machines are an implemented version of the state transition graphs described in *EMWC*, and these lego block constructions will only succeed in bringing about the anticipated transitions if their structure is imitating the pattern of causal influences that holds sway in the real world. It is not by coincidence that Lambert calls his model of data fusion a 'State Transition Data Fusion' model (STDF). Just as with ρbot, it is this dynamic interaction between the internal representation of state transitions and real world events that mediates the process of embedded interaction at the heart of Lambert's high-level fusion system.

8.24 The state transition data fusion (STDF) model

---

[26] Lambert and Nowak provide the following commentary to explain this diagram:

> Long-term memory houses cognitive routines that can be selected and executed. The execution of a routine R by cognitive individual X gives rise to changes in awareness, volition, interaction and internal action working memories within X. Awareness includes beliefs, expectations and anticipations.
> Volition includes independent intentions and potentially nested dependent desires. Intentions therefore signify a greater commitment than desires. Interactions include perceiving the world, effecting outcomes in the world, and the ability to inform other cognitive individuals. The internal mental actions occur when a cognitive individual X has volition to satisfy expression α, identifies routines whose behaviour can achieve α, and performs some of those routines, resulting in changes to working and possibly long-term memory (Lambert and Nowak 2008:25).

The STDF model (Lambert 2006b; 2007b; 2009a; 2009c) was developed to show how the approach to representation and belief, and the associated model of mind, discussed above, could be applied to the task of information fusion. While the JDL model represented object, situation and impact assessment as discrete processes, Lambert's STDF model stresses their interdependence.[27] The STDF model represents the object of assessment at each fusion level (objects, situations, scenarios) in terms of states, and transitions between those states (see figures 8.7, 8.8 and 8.91).

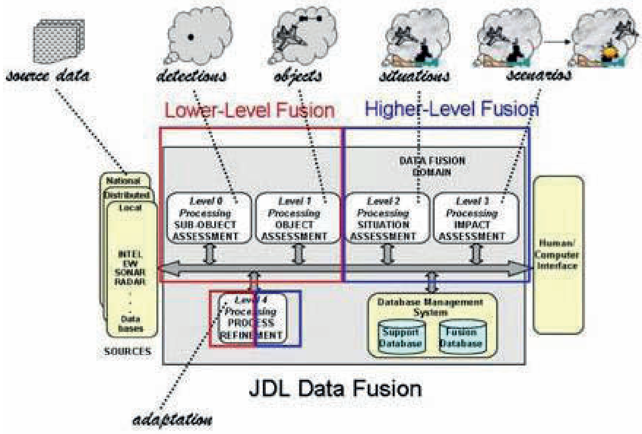


Fig. 8.7 The STDF levels of data fusion and the JDL model (Lambert et.al. 2009: 7).

As in *EMWC*, states and transitions between states are modelled mathematically, essentially by building upon and supplementing strategies outlined in that earlier work.[28] Higher level forms of analysis are essentially subsuming lower level templates of states under more encompassing state representations. This can be illustrated in terms of the *EMWC* examples provided in previous chapters; the object level is analogous to identifying the state of a light switch, at the next level this switch becomes part of a situation, such as a thief entering a darkened room; and then at the next level this situation becomes part of a scenario, such as possible outcomes of the thief finding or failing to find the light switch. With each level the new state is defined by shifting the focus from object instances, to situation instances to scenario instances.

Thus, STDF model envisages fusion as essentially the same process carried out at different levels focusing on larger chunks of process, including processes that have not yet taken place. At each level of assessment a networked, distributed fusion machine senses current states in the world using its systems of sensors (discriminators) and transfers this data to an observation process. This process draws

[27] Lambert states that the STDF Model is based on two premises:

1. "At each level of fusion the world can be assessed in terms of states and transitions between those states. For example, level 1 states are *state vectors*; level 2 states are *states of affairs*; level 3 states are *scenario states*; with *objects*, *situations* and *scenarios* being the sets of respective transited states over time".
2. "At each level of fusion, the fusion process adheres to the *same pattern of behaviour* but the *nature of the content changes*. . . . For example, registration is *coordinate registration* at level 1; *semantic registration* at level 2; and *situation assessment* at level 3" (Lambert 2009c:7).

[28] While the STDF model is grounded in Lambert's earlier work, it is far more sophisticated; for example, it is able to deal with probabilistic uncertainties in incoming data by treating the partial states of affairs as combinations of complete possible worlds for which a probability distribution exists. Mathematical procedures are detailed in Lambert (2007).

upon a prediction process that accesses stored representations of previous states to make comparisons between the predicted and actual situation. Control then transfers to an explanation process that creates new representations of the world (see the general model in figure 8.8). Data fusion therefore concerns "the prediction, observation and explanation of state transitions in the real world"; the fusion process "observes to explain; explains to predict; and predicts to observe" (Lambert 2009a:14).

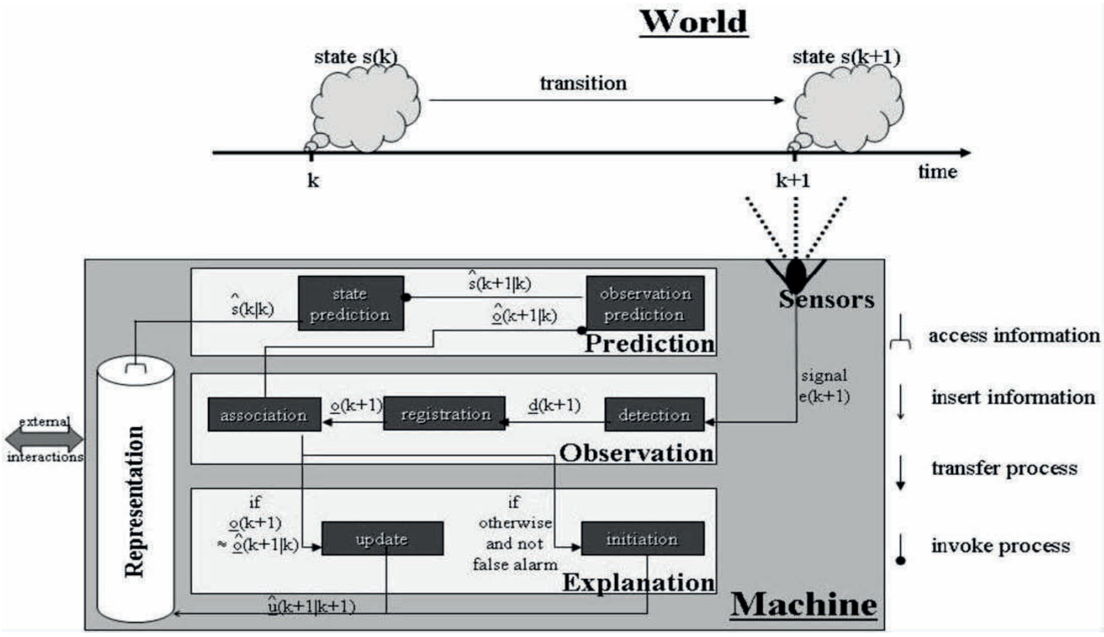


Fig. 8.8 The general form of the state transition data fusion model (Lambert 2009a:14).

What follows is a simple outline of the levels of assessment that make up the STDF model (see figure 8.9 for a formal description of these levels); higher level assessments build upon the input from lower levels. Thus:

1. Object assessments are primarily oriented toward measurable objects located in the world; sensors discriminate data and the object assessment level translates this into numerically based representations. "Changes in the measurable properties over time are taken to reflect objects as transitions in object instances over time".
2. Situation assessments build upon object assessments to present a semantic interpretation of the world (drawing upon the Mephisto ontology). Situations are represented "as sets of sets of statements about the world" and "changes in the statements over time are taken to reflect situations as transitions in situation instances over time".
3. Impact assessments build upon the situation assessment and draw upon routines to anticipate and evaluate potential scenarios unfolding from the present situation. Scenarios are expressed symbolically as sets of situations; as with the *EMWC* analysis of states and state instances, "changes in the statements over time reflect scenarios as transitions in scenario instances over time" (Lambert 2009a:13-17).

Assessment	State s(k)	Transition $\{s(t) \mid t \in \text{Time} \,\&\, t \leq k\}$
object assessment $\underline{\underline{\hat{u}}}(k)$	Each s(k) is a <u>state vector</u> $\underline{u}(k)$ explained by $\underline{\hat{u}}(k\|k)$.	Each $\{s(t) \mid t \in \text{Time} \,\&\, t \leq k\}$ is an <u>object</u> $\underline{\underline{u}}(k) = \{\underline{u}(t) \mid t \in \text{Time} \,\&\, t \leq k\}$ explained by $\underline{\underline{\hat{u}}}(k) = \{\underline{\hat{u}}(t\|t) \mid t \in \text{Time} \,\&\, t \leq k\}$.
situation assessment $\underline{\hat{\Sigma}}(k)$	Each s(k) is a <u>state of affairs</u> $\Sigma(k)$ explained by $\hat{\Sigma}(k\|k)$.	Each $\{s(t) \mid t \in \text{Time} \,\&\, t \leq k\}$ is a <u>situation</u> $\underline{\Sigma}(k) = \{\Sigma(t) \mid t \in \text{Time} \,\&\, t \leq k\}$ explained by $\underline{\hat{\Sigma}}(k) = \{\hat{\Sigma}(t\|t) \mid t \in \text{Time} \,\&\, t \leq k\}$.
impact assessment $\underline{\hat{S}}(k)$	Each s(k) is a <u>scenario state</u> $S(k) = \{\Sigma(n) \mid n \in \text{Time} \,\&\, n \leq \partial(k)\}$ explained by $\hat{S}(k) = \{\hat{\Sigma}(n\|k) \mid t \in \text{Time} \,\&\, t \leq \partial(k)\}$.	Each $\{s(t) \mid t \in \text{Time} \,\&\, t \leq k\}$ is a <u>scenario</u> $\underline{S}(k) = \{S(t) \mid t \in \text{Time} \,\&\, t \leq k\} = \{\{\Sigma(n) \mid n \in \text{Time} \,\&\, n \leq \partial(t)\} \mid t \in \text{Time} \,\&\, t \leq k\}$ explained by $\underline{\hat{S}}(k) = \{\{\hat{\Sigma}(n\|t) \mid n \in \text{Time} \,\&\, n \leq \partial(t)\} \mid t \in \text{Time} \,\&\, t \leq k\}$.

Fig. 8.9 Comparison of object, situation and impact assessment (Lambert 2009a:18).

Thus each level of the STDF model provides a more encompassing assessment of the events taking place in the world, from object, to situation, to scenario assessments. By focusing on states and their transitions and generalizing a model of sensor fusion to all of the levels of the fusion process, STDF promotes a unifying framework for data fusion. According to this model, situation awareness arises from object, situation and impact assessments, with these elements of data fusion corresponding in functional terms to the human actions of perception, comprehension and projection. Scenario prediction is obviously the most difficult challenge, as it involves assessing intent, capability and awareness of the various agencies.

Lambert defines a machine agent's intent as "the future states of affairs that the agent is instructed to achieve by effecting its environment", while its awareness of the world is "represented by its current situation assessment", and its capability options are defined as "any activity that has the ability to transition the state of the world from one state to an anticipated state" (Lambert 2009a:16). Capability, awareness and intent as previously noted (figure 8.1) are the action trinity that lie at the heart of command and control. Lambert notes that his notion of capability is understood more broadly than in the normal military sense, such that considering the effect of a capability option involves consideration of the reaction of other agents, and its broader impact upon the environment of the action. Hence, capability options involve an awareness of opportunities, or affordances, for achieving a particular end.[29]

[29] The notion of 'affordance' can be generalized from Gibson's (1979) restricted sense to encompass the action potential of a situation; awareness of a capability option is closely related to the notion of affordances, when viewed from Lambert's situated perspective. See also, Lambert & Scholz (2005) on capability options in the context of UC^2. Heinze (2003) explores the notion of affordances in the context of modeling intention recognition.

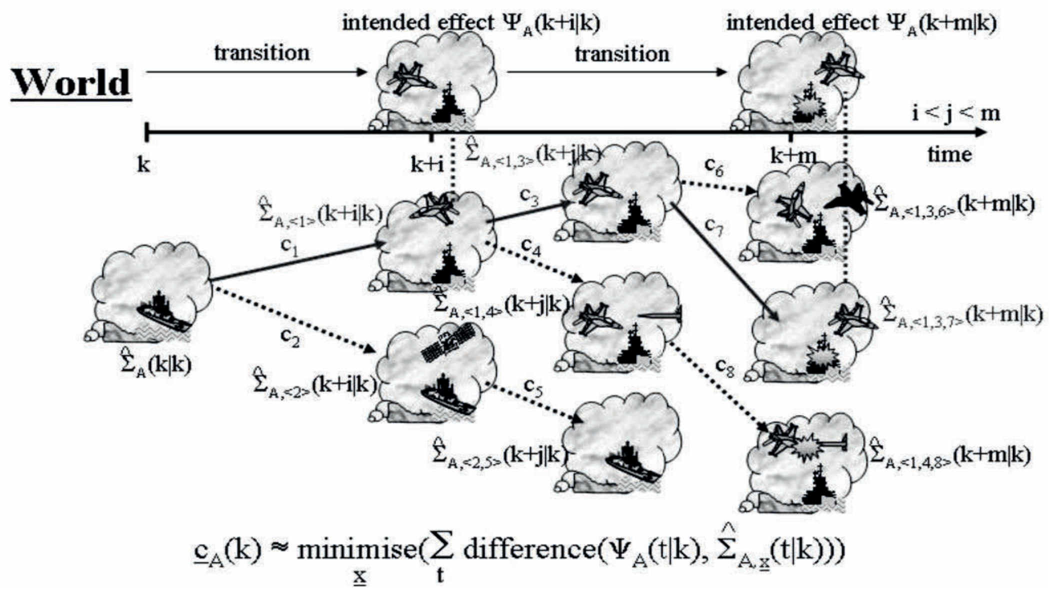


Fig. 8.91 An impact assessment (Lambert 2009a:17).

Because agent intent covers both human and machine agents, Lambert argues that the impact assessment level of situation assessment exercises a top-down controlling effect on the lower levels of information fusion, “for what one observes in part depends upon what one looks for, and what one looks for is shaped by one’s intent” (Lambert 2009a:18). He notes that his “goal or intent driven approach to data fusion” is out of step with the bulk of the data fusion community, but argues that a goal driven view of data fusion is necessary when multiple intents come into play, and the fusion system is engaged in a reciprocal interaction between the data and intent levels of information fusion. In this ‘Kantian’ form of situation assessment, the machine agent’s intent determines the data that it collects and the situation assessments that it forms. This goal driven vision of information fusion is central to the social level of the Mephisto ontology, and has far reaching implications. It recognizes that what the observers sees is dependent upon the standpoint of the observer, such that, for example, a Western soldier and a traditional Afghani warrior may see the same situation very differently. This awareness has consequences for estimating intent, and capability, and hence is fundamental to the operation of the UC^2 system.

A review of the social level of the Mephisto system and some comments on its implications for the operation of a UC^2 hybrid society will conclude my analysis of Lambert’s progress in realizing his Post-Classical vision.

8.3 The social level of Mephisto

The Mephisto ontology is a work in progress, and Lambert and Nowak acknowledge that the social level is only partially formalised.[30] Nevertheless, the work carried out provides a good indication of the vision guiding the current conceptualization of the social domain, and the potential of the Mephisto system for interpreting complex

[30] Lambert has stated that there “is no well defined theory in place” dealing with social interactions; all “the Mephisto framework offers are constructs for developing such a theory” (Lambert 2009b:13).

social situations that provide a meaningful context for fusion-based planning and responses to those situations.[31]

The STDF model envisages object, situation and impact assessments that are continually updated in response to information about unfolding events. As noted previously, Mephisto characterizes the world through five layers of domain concepts and associated interpretable symbols. Social events typically involve the five layers concurrently. Unlike computer game simulations of scenarios, Mephisto provides categories for machine agents to interact with the outside world; just as a vehicle's satellite navigation aids interact in real time with the Global Navigation Satellite System in order to locate the driver and provide information and directions to a desired destination, so Mephisto provides semantic categories that enable the fusion system to deal with information inputs and queries concerning complex geo-political strategic situations, and equally complex military goals.[32] As Lambert and Nowak observe, even a simple military scenario is likely to involve:

- nations and conflicts;
- physical geography;
- moving objects;
- military equipment with certain capabilities;
- civilian maritime and air traffic;
- masked intents – a military "chess game" (Lambert and Nowak 2008:2).

Mephisto provides a capability to represent such a situation, enabling ATTITUDE agents to reason about likely scenarios and interact with human agents in a meaningful fashion.

Lambert and Nowak state that the social theory level is the most important and challenging of all the theories that comprise Mephisto, "for it embraces all the theories of the lower levels, and in complex domains (domains where social interaction comes into play) it provides the highest level view of the domain" (Lambert & Nowak 2006:5).[33] Theory in this sense refers to a formal theory; "it clarifies the conceptualisation and facilitates the implementation of a Mephisto ontology based agent society" (*ibid*:1). It does this by first engaging in a meta-theoretical reflection on the domain area, in order to formulate a conceptual framework, and then formalizes this framework using formal axioms and definitions

---

[31] In 2008 a review panel concluded that:

> Mephisto is presently only partially formalised, so one drawback is that additional development is required. Its strength is its philosophical rigour and its emphasis toward developing sound and complete formal theories where possible. This makes it amenable to computational implementation and provides clarity of interpretation, but at a price of being further removed from natural language, which is more suited to a descriptive and multiplicative approach, rather than Mephisto's prescriptive and reductionist style (Lambert, Saulwick, et.al.2008:42).

[32] For detailed examples of such scenarios see Lambert (2007), Lambert and Nowak (2008).

[33] As previously noted, a Mephisto agent is an agent the employs a Mephisto ontology, and a Mephisto society is made up of Mephisto agents. Each level of the Mephisto ontology is based upon a small set of primitive elements and a theory that formally defines these elements. Examples of concepts include:

Social: ally, agrees, possesses, commands, …
Cognitive: believes, expects, prefers, perceives, …
Functional: senses, strikes, informs, moves, …
Environmental: air, water, upland, outer_space, …
Metaphysical: exists, identical, before, connects, …

expressed in the language of first-order logic.[34] Implementation of a theory thus rests upon a conceptualization. Conceptualization, in this sense, is analogous to what sociologists such as Talcott Parsons described as 'theory', in the broad sense of specifying a set of concepts adequate for describing the social objects and processes that constitute the social domain. Because of the pivotal role of social theory in the Mephisto system, and the dependence of theory on conceptualization, creating an adequate conceptualization of society is of great importance to the Mephisto system.

Lambert has stated that the key idea underpinning his Mephisto notion of social agents is derived from the *EMWC* notion that the subjects referred to in propositional attitude instructions are agents; either a reflexive reference to "I", a specific reference to another agent ("Fred"), or a contextual reference to another agent ("someone"). Thus cognitive agents could send requests to other agents that they perform some action; hence, propositional attitude instructions "not only served as the mechanism for behaviour within an agent, but also served as a mechanism for behaviour between agents" (Lambert 2009b:12). As a result of such interaction it is possible for cognitive routines to become part of a broader social routine. Lambert understands these broader social routines as patterns of coordination arising from the "flow of intent between the cognitive entities" due to a subject other than "I" featuring in propositional attitude instructions. Once agents have accepted an instruction to desire some routine, then that routine becomes part of their habitual behaviour and opportunistic collaboration may result - such as is found on the soccer field between members of team passing the ball to each other in the hope of creating the opportunity for a shot at goal.

Thus, Lambert and Nowak presuppose that the social level can be understood entirely in terms of interactions between individuals; in their words:

> The Mephisto Social Layer is understood in terms of *the flow of intent* between cognitive individuals (Lambert and Nowak 2008:29).

In formal terms this means that the social layer of Mephisto has virtually no primitive terms; all the concepts are defined by building upon primitive concepts associated with the cognitive layer. Thus social concepts are built up from ATTITUDE programming language concepts, such as 'believe', 'desire' and 'intend'. Lambert and Nowak comment that:

> The Mephisto Social Layer is therefore characterised in terms of social consequences based on cognitive individuals. Sociology, by contrast, tends to focus more on group interactions independently of the cognitive individuals involved. Mephisto could be expanded to cater for this (Lambert and Nowak 2008:29).

The motivation defining the Mephisto conceptualization of society is that of developing a military command and control system, where command is defined as the exercise of human will to accomplish a mission, and "control extends beyond command in that the controller disseminates a plan (routine) for achieving intent, and monitors and corrects execution of that plan" (*ibid*:34). A plan is interpreted as a cognitive routine which, by engaging other agents through propositional instructions,

[34] Lambert argues that the development of a computational semantic formal theory involves three steps:

> (a) *philosophical*, in which the conceptualisation of the domain of discourse is specified;
> (b) *mathematical*, in which the formal structure of that philosophy is specified; and
> (c) *computational*, in which a computational implementation of that mathematical theory is specified (Lambert 2006:1).

gives rise to social routines. The flow of intent between agents is mediated by agreements or conflicts. Agreements, Lambert contends, are the basis of social cohesion. In Mephisto agreements are understood in terms of a Common Law model of offer and acceptance. Lambert and Nowak note that their reference to Common Law "is not in the law *per se*, but in the naturally occurring social agreements that the law is seeking to represent. An offer involves informing one's own intent for another to that other, while acceptance involves informing compliance to that intent" (Lambert and Nowak 2008:30).

This logic is extended through to the definition of groups, of roles within groups, and of agency exercised on behalf of groups. Possessions and ownership are defined in terms of control by individuals and conflicts arise when mutually exclusive intentions collide. Communication, negotiation, collaboration and agreement between agents are the primary form of activity, within a social context defined by notions such as possession and ownership, trading, trading contracts, agreements, conflict, hostility and threat. Hence "a theory of agreement and conflict is a core of Social Theory" (Lambert & Nowak 2006:6). This theory is expressed via carefully defined concepts, and these formal definitions and axioms provide the strict formal constraints for machine reasoning about the social domain.[35]

Thus, the social level of Mephisto could play a key role in enabling a fusion system to "represent and model functional entities, such as organizations and assets, modelling their intentions, behaviours and business processes" (Perugini et.al. 2003). Examples of organizations or assets could include, at the micro end of the scale, individual agents, friendly or enemy army units, cargo ships and aircraft, and commercial freight companies through, at the macro end of the scale, to social collectives such as friendly or hostile nation states, multi-national organizations, and so on. The social level of Mephisto involves developing formal definitions of those factors that contribute to different levels of trust and social cohesion. Ideally, Mephisto's social theory would provide the conceptual framework for modelling the politics and economy of an Iraqi village, and linking this micro domain to the macro realm of global finance and politics, as reflected in corporate deals and international diplomacy surrounding the operation of Iraq's oil fields. Linking the micro and macro levels of social analysis to questions of trust and social cohesion would also mean contending with the role of cultural and religious factors. This is obviously a significant challenge.

---

[35] Logical constraints specify the meaning of each concept, and the implementation of these constraints determines how the machine interprets each concept, facilitating "both numerical calculation and abstract symbolic reasoning" (Lambert 2009c:8). For example, the following propositions provide a flavour of how, using this formal framework, a failure to negotiate trading contracts can lead to conflict.

$$\text{owns}(x, y, z) \text{ iff}_{\text{def}} \text{ possesses}(x, y) \ \&\ \text{legal-contract}(z)$$

$$\begin{aligned}&\text{trades}(u, v, w, x, t_1, t_2, y) \text{ iff}_{\text{def}}\\ &\quad \text{owns}(u \bullet t_1, v \bullet t_1) \ \&\ \text{owns}(w \bullet t_1, x \bullet t_1) \ \&\\ &\quad \text{owns}(u \bullet t_2, x \bullet t_2) \ \&\ \text{owns}(w \bullet t_2, v \bullet t_2) \ \&\\ &\quad \text{legal-contract}(y)\end{aligned}$$

$$\text{trades}(u, v, w, x, t_1, t_2, y)\ ,\ \text{agrees}(u, w, y)$$

$$\begin{aligned}&\neg\text{agrees}(u, w, y) \rightarrow\\ &\rightarrow \text{conflict}(u, w, y) \rightarrow\\ &\rightarrow \text{hostile}(u, w) \ \&\ \text{threat}(u, f(y)) \ \&\ \text{threat}(w, g(y))\end{aligned}$$

$$\text{threat}(x, y) \text{ iff}_{\text{def}} \text{ capable}(x) \ \&\ \text{hostile}(x, y)$$ (Lambert & Nowak, 2006)

8.31 The basis of machine societies: Lambert's Legal Agreement Protocol (LAP)

At the basis of Lambert's model of society is the notion of a flow of intent between individuals, resulting in agreements and conflicts. This flow of intent is mediated by offers that involve informing others of one's own intent, and acceptances which involve informing the agent making the offer of your compliance with their intent. According to Lambert, all forms of social cooperation depend upon some preliminary form of agreement. Even the players in the soccer team alluded to above would have to make some form of agreement prior to the "players exercising their social routines" (Lambert 2009b:13). At the core of societies are shared or interlocking routines that mediate a flow of intent, however prior to these routines "we have to establish the connection for the flow of intent to occur"; this, Lambert states, "is where the Legal Agreement Protocol comes in" (*ibid:*13.). Mephisto's social layer is based upon LAP agreements, but the technical detail of how this LAP agreement process occurs is "largely masked in the Mephisto framework" (Lambert and Nowak 2008:30). These details have been spelt out in Lambert and Scholz (2005), and were the subject of Don Perugini's (2006) doctoral thesis.

Don Perugini, a participant in the FOCAL program, showed in his doctoral thesis how ATTITUDE agents could be used to model the process of negotiating and arriving at agreements between those requesting and those providing a service (Perugini 2006). Lambert and Perugini have built upon the Extended Net Contract Protocol (ENCP) developed by Klaus Fischer and others (Fischer et.al. 1996). Fischer demonstrated that transportation planning problems, in an open domain, can be solved using an Extended Net Contract Protocol; this protocol provides a method for an agent to solicit offers of assistance from a group of agents to achieve part or all of the agent's intent. Members of the group may respond with offers, and the soliciting agent then selectively joins these offers together, lego block style, until they have a viable strategy for realizing their intent (Fischer et.al. 1996). Perugini's "Provisional Agreement Protocol" (PAP) builds upon this model and enables machine agents to negotiate the best solutions to complex logistical tasks (Perugini et.al, 2003). Fischer, it is interesting to note, subsequently became one of the key contributors to the Socionics program (Schillo, Fischer, & Klein 2001; Fischer et.al. 2005; Hahn et.al.2007). Lambert has promoted the "Legal Agreement Protocol" (LAP) as his preferred agent social agreement protocol. It is a strengthened version of PAP that models legal contractual agreements that can apply between people, between machine agents, or between people and machine agents (Lambert & Scholz 2005).

In the PAP and LAP models of social coordination, agents must communicate information about their intent, and their offer of service. If requests and offers are to be matched up successfully, they must share an understanding of the required *function*, as well as the "*what*, *where*, *when* and *how*" that specify the place, time and type of service requested (Perugini 2006:48-9). Perugini has suggested that the development of ontologies is essential to agent communication in open systems, and envisages a MAS architecture in which agents would interact using a separate ontology to specify each component of the agent's goals (such as function, what, who, when, where, etc.). This, he argues, would result in a similar ontology to Lambert and Nowak's Mephisto schema, with different layers of Mephisto dealing with questions relating to the different ontological domains.

Lambert has commented that his "proposal to use the Legal Agreement Protocol that appears in Contract Law was not motivated by a desire to have legally enforceable agreements (though that can be a bonus) or to ensure that these agreements were rational". Rather:

> The point is that the flow of intent between two cognitions can result in an agreement . . . or it can result in conflict, in which case the competing intents battle for supremacy. Agreement and conflict thus become fundamental constructs for addressing the social level. Agreements are the glue by which societies form. Conflicts are the mechanism for fragmenting societies (Lambert 2009b:13).

The LAP proposal is Lambert's solution to the problem of achieving social coordination in a hybrid society. Lambert suggests that his UC^2 system will consist of pool of human and machine agents that, depending on the circumstances, can be demanding or supplying capability. To maintain flexibility without succumbing to anarchy, Lambert envisages "full contract law agreements between agents, be they human or machine"; but notes that significant problems remain to be solved. Thus,

> Additional social policies will be required if individuals are to trade self-benefit for the benefit of a collective. This could involve, for example, a multi-scale networking constraint to ensure efficiency within a collective. It could involve a prioritising of intent according to the rank of the members of the collective. It could have social constraints that govern how to trade membership and the quality of decisions with the time available. These issues will vary with the nature of the collective. (Lambert & Scholz 2005:34)

Thus Lambert's initial foray into the social realm is guided by a conceptualization that revolves around a flow of intent; capability for achieving intent must be grounded in accurate situation and impact assessment. For machine agents to contribute to this increased capability, their contractual negotiations must contend with the complexities of the real world. As Lambert notes, the "social agreement protocols and constraints have to contend with both dynamic intra and inter social group consequences" that impact upon the achievement of intent; issues such as trust and social cohesion "will be vital in situations of high uncertainty" (Lambert and Scholz 2005:35).

8.32 The UC^2 vision of hybrid societies

In the previous sections of this chapter I have described how Lambert's adaptation of his Post-Classical paradigm to the information needs of a military organization gave rise to his vision of 'Ubiquitous Command and Control' or UC^2; I have also provided a broad outline of his design for an information fusion system capable of realizing the UC^2 vision. In the UC^2 vision of Network Centric Warfare, machine agents act as personal assistants to human agents, linking them into a networked information fusion system; this system acts as a distributed Command and Control information and communications hub, introducing a more flexible element into the military's traditional organizational structure. The enhanced flexibility is a result of the UC^2 'awareness-intent-capability' trinity, and the role that machine agent's play in structuring and supplementing the flow of intent, awareness and capability within the military organization.

Crucial to the UC^2 vision is the notion that these human and machine agents are part of a networked hierarchical organizational structure. The aim of the UC^2 system is to exploit machine capabilities; ideally, machine communications could provide human agents with powerful tools to enhance their situation awareness, and act as a multiplier of intent and capability. In part this would be accomplished by cognitive agents acting as gate-keepers (decreasing the 'chaff' of communications that potentially arise between human agents when the need arises to evaluate and respond to a rapidly changing situations); in part by cognitive agents acting as research assistants and facilitators (processing, evaluating and communicating information that is tailored to the role of the recipient, such that it sharpens their awareness of their situation, speeds the flow of their intent and enhances their capability to carry out their responsibilities). The latter proactive role of machine agents leads beyond the vision of intelligent personal assistants to that of teams of machine agents.

Lambert's vision of virtual teams, inspired in part by developments such as 'Wikipedia' and 'ebay', is to harness the potential of machine agents in order to maximize the diverse capacities and expertise of a large networked society. The UC^2 platform can support teams in various configurations with various capabilities assembled as the need arises. In the UC^2 framework both human and machine agents are engaged in developing an awareness of situations, forming goals, negotiating plans and creating the organizational structures necessary to implement those plans 'on the fly'. Potentially, the Australian military and key strategic allies, if linked by such a system, could mobilize public and private expertise and resources in order to enhance awareness, intent and capability needed to address situations of concern.

Lambert envisages the real-life organizational assembly process as a mirror of the information fusion assembly process, as suggested by his lego™ model of state transitions. Negotiations fly backward and forth between situated agents, human and machine, in a manner analogous to the automated bidding process in ebay. The fusion process is not just about assembling information about a situation but about assembling military capabilities so as to respond to situations. Different types of situations would be rehearsed, and these would involve different modes of human-machine interaction. Responding to a civil disaster, for example, a maritime rescue, could be a rapid, largely machine led response, while a terrorist alert might require far greater levels of human input and evaluation. In both cases the flow of intent between human and machine agents, and resultant forms of hybrid social collaboration, could lead to the development of hybrid social routines for responding to different types of scenarios – routines that involve a symbiotic relationship between machine and human decision making and goal attainment processes.

As the military's real world capability becomes dependent upon the fusion process, military capability will become increasingly synonymous with fusion capability. Fusion capability is not just about integrating information; it is about assembling people and machines into effective teams, tailored to accomplishing specific tasks. Fusion capability flows through the various phases of strategic situation analysis and response, into the domain of practical action. In Lambert's words, if a need arises:

> Fusion capability is then assembled like lego™ blocks on the fly as required, through contracted agreements that propagate between available agents through the network at that time, with social policies and protocols to manage social coordination. Fusion expertise is then dynamically contracted from the available pool of expertise, not prescribed in advance. This accommodates a network in which the constitution of agents is dynamic and in which the fusion

> product can be arrived at by a combination of competitive and collaborative processes. (Lambert 2009a:19-20)

Lambert's UC^2 ideology is primarily oriented towards enhancing the capabilities of the military organization, and wider political, social and cultural ramifications of this ideology are largely bracketed. However it would appear incontestable that the very same factors that lead Lambert to associate information with enhanced awareness and capability are also responsible for introducing greater measures of contingency into society.[36] Within this dynamic environment, Lambert's vision is oriented toward enabling the military to realize its mandated functions in a more efficient and effective manner. Gone is Kubrick and Clarke's vision of a solitary electronic brain; instead there is a networked world of autonomous agents – a benign version of the Matrix – facilitating collaboration between human agents and supplementing the traditional hierarchical leadership model with virtual teams, assembled for specific tasks, using a collaborative management style. Lambert suggests that this amounts to the creation of a new type of social order: in the language of Socionics he is envisaging the emergence of hybrid societies.[37]

8.33 Critical comments: how adequate is Lambert's conceptualization of the social domain?

Lambert's vision of society is an extension of the notion of commonsense cognition detailed in *EMWC*, however, in situating his cognitive agents in real social contexts the sociological shortcomings of the Post-Classical paradigm vision of cognition become more apparent. In brief it revolves around the under-theorized notion of routines, and the more general failure to deal with the structured environments of social action. However, these shortcomings are not inherent in the paradigm itself; lego blocks can be used to build all manner of constructs and as I have indicated in the above discussion, though sociological theory has played little part in shaping the author's vision of the social level, there is no in-principle reason that prevents this from occurring. However, there is a considerable cultural gap between sociology and computer science that imperils appreciation of what both sides have to offer, and vitiates the chances of productive collaboration.

Standing in the way of collaboration are deep-rooted prejudices. On the one hand the sociological theorist is familiar with long discursive explorations of subject domains that tease out a multitude of theoretical positions; the absence of such analysis and the bare formal language of Mephisto would tend to alienate most sociologists. On the other hand, those with an engineering background tend to find long discursive reviews

---

[36] The social implications of information fusion technology would obviously be profound; every bit as 'game-changing' for our society as the advent of the automobile.

[37] In Lambert's words: "The UC^2 framework also encourages a new social order by not discriminating between human and machine agents, and so accommodates mixed initiative strategies under which both people and machines can initiate and propagate intent, . . ." (Lambert 2009a:20). In suggesting 'no discrimination' between human and machine agents, Lambert's motivations are utilitarian rather than socio-political. Given the differences between human and machine capabilities, costs, availability and so on, there would obviously be task differentiation, however, selection of individual people or machine agents for specific roles would be based upon the flow of intent passing in a non-discriminatory fashion between members of the hybrid society.

of theoretical debates alienating; they are more accustomed to looking at short papers that provide bare formal details specifying what something does and how it works. Thus, for the sociologist, there is a risk of dismissing Lambert and Nowak's formulations prematurely without considering the power that stems from the logical simplicity and rigor of the Mesphisto approach. There is a need for sociologists to get inside the mindset of the engineer and reframe ideas in sociological terms; what for example, does the 'flow of intent' signify from a Luhmannian systems theory perspective? A more considered evaluation of the social level of Mephisto should acknowledge that the privileged status that Lambert accords the social level, from the context of computer science and that disciplines efforts to engineer an artificial intelligence capability, represents something of a revolution; the challenge for sociologists is not simply to be arm-chair critics of this revolution, but to be participants. While taking a critical stance, the social theorist needs to recognize that Lambert and Nowak have taken some very significant steps toward unlocking the social implications and potential of the Post-Classical paradigm.

Lambert's stress upon the importance of the social layer of the Mephisto ontology is a consequence of two factors, both of which have pushed his Post-Classical paradigm away from Classical AI's agent-centric, cognitively oriented view of information 'processing', toward an organization-centric, socially oriented view of information 'processes'. Firstly, the UC^2 context of information fusion envisages a hybrid society, made up of human and machine agents, working together in a collaborative fashion enacting complementary routines to achieve the goals of the military organization. This necessitates a deep understanding of not only how individual's function, but how groups of people and organizations function. In other words the focus shifts from a psychological to sociological frame of analysis. Secondly, by recognizing the importance of intent, both in terms of estimating that action of other agents, and as a top-down factor in shaping awareness, the STDF model of information fusion implicitly directs attention toward the shared customs that define communal and organizational patterns of action and in combination shape the intent of individuals within a group. Thus, at a strategic level, the STDF model shifts the focus away from a naïve 'physical data' driven view of information that assumes language is a neutral filter behind which lies one 'manifest reality', toward a constructivist view of reality, that recognizes that 'information' is in large measure defined by the socially constrained patterns of activity in which agents participate.[38]

An organization-centric view recognizes that the importance of information is defined by the social context of communications, and that context is defined in large part by the function and structure of the organization. As a social organism, what we perceive is what the customary forms of organization and understanding peculiar to our society has conditioned us to perceive.[39] Hence, what constitutes 'useful information' is defined by the social location and identity of the agent. Information theorists have been slow to see the significance of this social dimension. *Recognizing that*

---

[38] As Whitehead argued, even science hangs on a slender thread largely constituted by the human imagination (see chapter three).

[39] Broadly speaking, this entails recognizing (following Whitehead, Bateson, Gibson et.al.) that what an organism perceives in its environment as information are those 'differences that make a difference' from the standpoint of the organism; that is, perceived affordances are a consequence of the organisms' historically conditioned habitual dispositions interacting with an environment that is structured by diverse constituent societies, each likewise guided by habitual patterns of activity.

information has a social dimension, means acknowledging that belief is not just a cognitive process; rather, it is a social process, intimately tied to the dynamic functioning of organizations.[40]

The Mephisto framework's explicit recognition of the social level of analysis represents a significant development in the evolution of the Post-Classical AI paradigm. In *EMWC* Lambert had addressed the theme of intelligent action primarily through the notion of routines viewed from the standpoint of cognitive science and the neurophysiology of the brain. The focus was largely on ρbot the individual; there was no explicit mention of a society of ρbot agents or of the way in which societal 'life-worlds' define the environments of human agent interaction. Yet, as Agre and Horswill (1997) have argued, a cursory consideration of the acquisition of routine behaviour immediately takes us from consideration of the isolated individual into the social context of behaviour and the role that societal structures play in shaping routine behaviour. This social dimension is an under-theorized aspect of Lambert's original Post-Classical paradigm. Lambert begins to address this neglect in his reflections upon the way in which his ATTITUDE agents might implement his UC^2 vision of a hybrid society.

At the heart of Lambert's UC^2 ideology, and its Post-Classical vision of a hybrid society of human and machine agents, is the social theory that defines the social level of his Mephisto ontology. Clearly, Lambert's Post-Classical approach to AI constitutes a significant advance over Classical AI's disembodied, solipsistic, overly rational approach to mind. However, there is a disturbing flavor to the LAP proposal; it suggests a contractual model of social solidarity forged by individual agreements more reminiscent of neo-classical economics than a process vision of action as emergent from the nexus of unit act and society. Mephisto's social level, by focusing on individual agent intent, and the role of individual consent in promoting social solidarity or conflict, creates a strong flavor of methodological individualism and suggests a classical *homo-economicus* view of war.[41]

While Lambert's model of commonsense is inspired by a vision of human action that is at odds with the neo-Cartesian image of the rational agent, bent upon maximizing outcomes in accordance with some calculus of self-interest, the decision to base the sharing of routines upon a process of negotiation and agreement means that consideration of the complex intersection of biological, psychological and socio-cultural factors that condition and constrain individual choices receive scant consideration.[42] This is perhaps in part due to the implemented examples of the social

---

[40] The social dimension of information has been central to the 'organizational semiotics' approach to information inspired by Ronald Stamper (see Stamper 2001; Gazendam & Liu 2005). While social theorists, from the time of the founding fathers through to the present time, have been stressing the social relativity of human ideas and understanding, the implications of this are only slowly being absorbed within other disciplines.

[41] Cramer has analysed a recent trend toward explaining war in these terms, in part due to the growing influence of rational choice theories in the social sciences. Cramer argues:

> . . . that rational choice theories of conflict typically lay waste to specificity and contingency, that they sack the social and that even in their individualism they violate the complexity of individual motivation, razing the individual (and key groups) down to monolithic maximizing agents (Cramer 2002:1846).

[42] Lambert does acknowledge that families are excluded from his vision of social solidarity, and that his model has the capacity to model various forms of solidarity, including those grounded in biological

level. The potentials of the Mephisto platform might be better demonstrated if more consideration was given to the role of the family in society, for example, and the part that a range of structural factors play in shaping that family environment and its routine forms of life.

A potential advantage of the Mephisto system is that of accommodating a wide range of human motivations, as exemplified in diverse constellations of behavioral and socio-cultural routines. The Post-Classical paradigm is a significant alternative to the 'game-theoretic', and rational choice approaches to society. While rational choices and contractual arrangements play an important part in human society they work in conjunction with other powerful forces. Empirical studies have demonstrated many times over, and in many different cultural contexts, that rational choice models of human behavior have a poor capacity to explain human action.[43] The tendency of economists to view rational instrumental action as the paradigm of all human action is another example of the same flawed Cartesian presuppositions that have underpinned Classical AI. Nowhere is the importance of underlying models better illustrated than in economics; millions of dollars are invested annually in processing massive amounts of quantitative data to generate predictions that are only as viable as the presuppositions built into their underlying models. If Lambert is to avoid this scenario, and equip his fusion system with a capacity to make realistic assessments of the human situation the conceptualization and formalization of the cognitive and social levels of Mephisto may benefit from further consideration of sociological efforts to conceptualize social action.

Lambert's approach toward theorizing the bonds that join people into social collectives would appear to have neglected sociologically formulated approaches to explaining the formation of social groups. Durkheim, Weber, Mead and their sociological successors are absent voices. In his critique of Classical AI Lambert framed a persuasive critique of overly rationalistic reconstructions of human action. Lambert's agreement based model of social action would enable computer agents to shape new social organizations that are integrated into the wider social body, and form part of its web of legal obligations and responsibilities; however this focus on

---

drives and those based upon shared systems of cultural beliefs. This capacity is rooted in the logical constructs mediating his system. Hence:

> Within the implemented ATTITUDE model the **approves** and **disapproves** predicates are applied as success and fail constructs to control the execution of cognitive routines. In the Mephisto model **approves** and **disapproves** provide a basis for some absolute value judgements, like good and bad; **succeeds** and **fails** provide a basis for other absolute value judgements, such as true and false; while the **prefer** predicate provides scope for representing relative value judgements. Emotional behaviour by a cognitive individual can be cast in terms of heightened preference for lower level limbic system like routines, including fighting, feeding, fleeing and sexual intercourse. . . .[These constraints form the basis for defining the flow of intent that create and sustain social groups]. Social groups of interest are typically identified by the agreements and conflicts that unite and differentiate them respectively, with biological families being a notable exception. [Such groups could include culturally defined identities such as religions] (Lambert and Nowak 2008:26,29).

[43] See, for example, critical perspectives on rational choice theory in the work edited by Coleman, J. S. and Fararo, T. J. (1992); also the 'Introduction', in Archer, M. S. & Tritter, J. (2000) and Archer's chapter in the same work, 'Homo economicus, homo sociologicus, homo sentiens' (Archer, 2000). For an empirical assessment of rational choice theory, based upon cross-cultural studies, see Henrich, J., et al. (2005) "'Economic man' in cross-cultural perspective: Ethnography and experiments from 15 small-scale societies". *Behavioral and Brain Sciences*, 28, 795-855.

agreements creates the appearance that society is based on something analogous to the social contract theory famously pilloried by David Hume.

Contrary to contract theorist's faith in the importance of the freely given consent of the governed for the formation of societies, Hume argued that history discloses an entirely different mechanism underpinning the formation and dissolution of states.[44] A cursory glance at history discloses a bloody trail of wars of conquest, accidental evolution of constitutional frameworks, periodic reversion to the rule of fear and terror, economic booms and busts driven by greed and fear, and religious and cultural movements driven by emotionally charged beliefs, that express aesthetic, ethical and spiritual visions of what could be. Macro-societal factors at times appear to be more influential than individual freedom, such that at times the individual's capacity for rational calculation and negotiation play only a marginal role in the processes that shape human history. As Whitehead recognized, the macro and the micro are interwoven in every act and rational judgments are a form of routine operative in a small percentage of our actions; hence the polarity between methodological individualism and methodological holism is an artifact of an inadequate conceptualization, rather than a reflection of experienced reality.

Recognition that rational, instrumentally oriented human activity is but one dimension of human social action should provoke caution in according rational negotiation and legal contracts the pivotal role in defining the social dimension of MAS activity. As Boden (2008) has observed, simply because a MAS-based simulation has multiple agents engaged in communicative interaction does not mean that agents are truly social. As argued in my earlier review of AI (see chapter 1), a socially situated approach to MAS should look to human society for a deeper understanding of social action. On purely pragmatic grounds, it is worth considering the range of theories that sociologists have developed in an attempt to conceptualize social action. This is the position that advocates of the Socionics research paradigm have adopted.

8.4 Mephisto, Socionics and Process Philosophy

The development of software such as Mephisto is an interdisciplinary process that relies upon a productive collaboration across traditional academic boundaries. Sociology and computer science are not traditional bed-fellows.[45] However, since the

---

[44] David Hume famously lampoons the idea of an historical social contract in his 1748 essay, "Of the Original Contract": "The face of the earth is continually changing, by the increase of small kingdoms into great empires, by the dissolution of great empires into smaller kingdoms, by the planting of colonies, by the migration of tribes. Is there anything discoverable in all these events, but force and violence?" (Hume 1994, 190).

[45] As noted above, one of the primary obstacles to collaboration between computer scientists and sociologists is the cultural chasm that separates the disciplines; a divide that finds expression in the contrasting methodologies and philosophical presuppositions that frame disciplinary research methodologies and theoretical reflection. It is difficult for a sociologist to take a computer scientist's use of social metaphors seriously if the use of social concepts amounts to no more than the anthropomorphic projection of human qualities onto a system that bears only the most superficial of resemblances to human society. Likewise, computer scientists remain blithely indifferent to much sociological theorization, as long as sociologists fail to clarify and make explicit the various conceptual and logical presuppositions that inform their theoretical models of society. However, this situation changes when sociologists and computer scientists begin to utilize the same conceptual framework to describe and explain the workings of both human and machine agent-based societies. At least this is the

late nineties, a convergence of interests arising from the role of social theory in inspiring MAS architecture, and the role of MAS software as a vehicle for modelling society, has begun to generate the outlines of a new field of study. Following Malsch (2001), Fischer et.al. (2005), Nickles & Weiss, (2005), etc., I have referred to this emerging field as Socionics. Socionics is an attempt to fuse together the horizons of both the computer sciences and sociology around a shared interest in the potential of MAS for social simulations; a specific focus of Socionics is the purposeful interaction of human and machine agents in virtual social worlds and the design of the hybrid human-computer based social organizations that support this purposeful interaction. As previously noted, Lambert and Perugini have indirectly, via the work of Fischer, created a link to the work carried out within Socionics, but they have not yet embraced the Socionic methodology. Such an encounter, I believe, could benefit both Lambert’s paradigm, and the emerging discipline of Socionics.

While the immediate goal of this current study is to promote an encounter between the philosophy of Whitehead and Lambert’s Post-Classical paradigm, with the aim of interpreting Lambert’s work in the light of Whitehead’s philosophy, and vice versa, interpreting Whitehead’s philosophy in the light of Lambert’s Post-Classical paradigm, the larger goal is that of using a Whiteheadian reading of Lambert's project as a vehicle for contributing to the development of the Socionic paradigm. Various scenarios explored by Lambert in the development of his Mephisto project provide concrete illustrations of the type of technology envisaged by Malsch and other pioneers of the Socionics program. Thus Lambert's agents, for example, have many of the characteristics of Malsch's personal electronic assistants (PEAs) in that they provide support for military personnel to achieve their goals, by engaging in the task of co-coordinating and facilitating the collection of data, analysis, decision-making and implementation related to military operations (Müller et.al.,1998). The potential contribution of Lambert and his team to the socionics program goes beyond that of creating hybrid systems that exemplify in concrete form the socionic domain of research. His Mephisto conceptualization also brings to the fore a philosophical dimension that practitioners of the socionics methodology have not explicitly addressed.

The process philosophy based conceptual framework of Mephisto illustrates the fundamental, but often overlooked importance of philosophical foundations, both for computer science and social theory. Thomas Fararo, one of the doyens of mathematical sociology, and a founder of computational sociology, has long been advocating a process philosophical foundation for the social sciences (Fararo, 1989; 2001). He has argued convincingly that most of the theoretical streams that make up the tangled delta of sociological theorization can be inter-related within a process framework. He has also played a central role in developing various mathematical approaches to modelling social theories. Andrew Abbott, in his advocacy of a process based approach to social theory, argues for a more radical over-hall of existing theories. However, he also recognizes a relational linkage between the various competing schools of thought that have shaped the sociological tradition, and views the process philosophical tradition as an appropriate conceptual framework to recast the sociological tradition. (Abbott, 2001; 2006; 2007)

---

hope that drives my attempt to reframe computer science and social theory in terms of Whitehead’s metaphysical system.

Lambert's Mephisto framework, by drawing upon the concepts of process philosophy, points the way to a mode of thought that is capable of bridging the interdisciplinary gulf between computer science and sociology. It offers the outlines of a conceptual framework within which sociology and computer science can collaborate and develop theories that encompass the complex issues involved in the integration of machine-based and human societies. If viewed from the context of Whitehead's metaphysical system, the process metaphysical presuppositions the Post-Classical paradigm and their expression within the Mephisto ontology provide a promising basis upon which to build a more adequate encounter between computer science and sociological theory.

Thus, from a socionic perspective, Lambert's process conceptualization, at least at a philosophical level, provides a promising framework for reflecting upon computational and sociological theories of agent-based societies. Moreover, his approach toward providing a rigorous mathematical formalization and implementation of situated action, invites the sociologist to consider the possibility of drawing upon Lambert's conceptual framework to transform discursively formulated sociological theories into the formal language of computer science, and vice versa, to draw upon sociological theories to flesh out Lambert's theory of society.

8.5 Conclusion

This chapter has described the progress made by Lambert and his team in further developing and implementing the Post-Classical AI vision outlined in *EMWC*. I have described how Lambert, under the auspices of DSTO and a program aimed at developing automated information fusion, has been able to continue to refine and develop his ρbot inspired multi-agent system, and to incorporate the resultant ATTITUDE agents into his Post-Classical Mephisto framework. In my analysis of these developments I have focused on the social level of the Mephisto ontology, as this level is crucial to the fusion system and its role in the creation of UC^2 inspired hybrid societies.

The work carried out by Lambert and his team, under the auspices of the FOCAL project, have provided a glimpse of the hybrid societies that will emerge as fusion technology matures. I have suggested that the social level of his Mephisto conceptualization needs further development, and concluded my account of post-*EMWC* developments by suggesting that the Mephisto program might benefit from engaging with sociological theory, in the fashion advocated by proponents of Socionics. The challenge runs in both directions. Socionics, if it is to foster a mutually intelligible conversation between sociologists and computer scientists, must confront philosophical issues.

For many sociologists already engaged in MAS based social simulation, my emphasis on clarifying philosophical foundations is an unnecessary digression. However Lambert's work constitutes a decisive response to those who want to get on with the job of programming and running simulations while dismissing concerns about the ontological foundations of computer science and social theory as 'overly philosophical'.[46] As *EMWC* demonstrated, philosophy is central to designing the architecture of a MAS, and as I indicated at the end of my analysis of *EMWC*, more work at the philosophical level is required if Lambert's Post-Classical paradigm is to serve as a vehicle for conceptually framing and clarifying the particular attributes that machine and human agents bring to the creation of a hybrid society.

To that end, I propose to develop a three way Socionic conversation between Whitehead's process philosophy, Lambert's Post-Classical paradigm, and Talcott Parsons' sociological theory. Parsons, as a pivotal figure in modern sociological theory, provides a point of access into key developments in modern social theory. At the same time, as I have discussed elsewhere his work has a strong linkage to Whitehead's process foundations, (Dawson 1991). A conversation between Whitehead's process philosophy, Lambert's computational theory, and neo-Parsonian social theory could build a new foundation for Socionics; one based upon a coherent philosophical bridge between the concerns of computer science and the preoccupations of sociology.

[46] Although there are sociologists who emphasize the importance of philosophical presuppositions, and are keenly aware of there relevance to methodological questions, such as Abbott (2001, 2006 & 2007) and Fararo (1989 & 2001), their efforts tend to be ignored or dismissed as 'overly philosophical'. See, for example, Manzo's (2007) critique of Abbott (2007). Computer science, likewise, has a long tradition of largely ignoring even well-informed philosophically oriented reflections on the nature of computation (Agre 1997b).

Conclusion

In the introduction to this work I suggested that the label *high-level information fusion* failed to convey the significance of the technology that is currently being developed by the DSTO team led by Dale Lambert. In the course of this work I have tried to show that this technology rests upon a significant shift in theoretical perspective; in fact, nothing less than a paradigm change in the field of artificial intelligence. Just as the theoretical change from Newtonian physics to the New Physics was precipitated by profound changes to the philosophical presuppositions that had framed the Newtonian scientific views of the Cosmos, so too Lambert's work in information fusion rests upon a process philosophy inspired transformation of the presuppositions that informed Classical AI.

Paradigm changes in the sciences do not occur rapidly nor do their implications disclose themselves readily. Just as we can walk the soil of another nation and not appreciate that the people of that society and culture see and experience their world differently, so likewise it is possible to walk through the technical details of a new scientific paradigm without appreciating the different world that it signifies. This applies to the detail of Lambert's information fusion system. Hence, in this work I have devoted a good deal of attention to constructing a Whitehead inspired metaphysical horizon of meaning to frame my interpretation of Lambert's Post-Classical paradigm and reinforce the distinctiveness of his process philosophy inspired presuppositions.

Twenty years ago, when Brendan Rogers awakened my mind to the possibility of developing a computational model of Whitehead's Creative process, he urged me not to worry about the formal side of things, and instead focus on achieving a clear conceptualization. This advice, while perhaps appropriate in that situation, was fudging the truth; in reality, as Whitehead urged in his debate with Dewey, philosophical and mathematical modes of thought should walk hand-in-hand.

Symbolic logic is more than a fancy dress for the conceptual body of a theory; it challenges us to think more clearly about the essentials of a conceptualization. Hence, in the preceding chapters I devoted a good deal of attention to showing how Lambert appropriates the formal tools of symbolic logic to express his Post-Classical vision of AI. My aim was to echo the message of Whitehead's (1906) pioneering article on mathematical visions of the material world, and show that the same logical tool-kit can be used to express radically different metaphysical visions. Whitehead demonstrated this by contrasting Newtonian and Leibnizian visions of the space-time continuum; Lambert demonstrates the same truth by contrasting Classical and Post-Classical visions of mind. The challenge in the case of both Whitehead's (1906) and Lambert's (1996) use of symbolic logic is to see through the formal definitions and axioms to the underlying model and appreciate how that model is giving expression to a metaphysical vision.

It should be clear by now that Lambert's information fusion technology is an effort to appropriate the tools of Classical AI and harness them to a different vision of human commonsense. I have shown how Lambert justifies this move by a critique of Classical AI's philosophical foundations; I have argued that Whitehead's criticisms of the bifurcated Cartesian notion of representation reinforce Lambert's critique. Read together, their work reinforces the message that to have any chance of engineering commonsense it is first necessary to get the metaphysical presuppositions right. As Whitehead argued, those that contend that they have no need for metaphysics labour under a delusion, as our "handling of scientific concepts is controlled by the diffused metaphysical concepts of our epoch" (Whitehead 1933:183). Ideas of mind and matter reflect metaphysical presuppositions. In Whitehead (1926) this was demonstrated with respect to physics but it is also true with regard to cognitive science and AI. Whitehead's metaphysical system provides a novel standpoint from which to relate an abstract understanding of the human capacity for commonsense judgements to machine mediated judgement. Lambert's blue-print for information fusion has begun the process of harnessing this novel understanding of the mind's capabilities and created the promise of a new source of purposeful action.

A transformed notion of representation was a key element in motivating Lambert's reformulated notion of commonsense action. I have suggested that Whitehead's metaphysical system supplements and supports Lambert's notion of representation. As argued in chapters two and three the Classical AI approach to representation was motivated by Aristotelian and Cartesian presuppositions and sentiments. It was assumed that Natural language mirrored the ontological structure of the world, and that machines could replicate our knowledge of the world by logically manipulating linguistically framed propositions that express facts about the world. Lambert's Post-Classical AI suggests an alternative strategy; one based upon the notion of detecting and recording repetitive patterns of activity in an environment.

Rather than build upon Aristotle's notion of a static mind-substance mediating the process of representation, the Post-Classical paradigm follows Whitehead in asserting that "process itself is the actuality" (Whitehead 1933:318). Whitehead's socially oriented view of representation is articulated in terms of the metaphysical categories used to define the generic societal processes constitutive of Reality. Hence, 'society' for Whitehead, assumes a metaphysical significance akin to substance in Aristotle's philosophy. I have shown that Lambert gives a central place to a similar factor in his Post-Classical explanation of the Universe; order arises in the flux of process from *self-repeating patterns of influence.* Lambert states that *this notion is his central intellectual invention*, as fundamental to his philosophy as substance is to Aristotle's philosophy. These self-repeating patterns of influence are fundamental to Lambert's original Post-Classical notion of representation.

Lambert's notions of routine or custom builds upon this notion of representation. I have shown that from Whitehead's standpoint enduring patterns of activity, including agents and their environments, are custom-defined societies; our capacity to objectify and interact meaningfully with these societal environments is likewise born of customs. Language rests upon a historical legacy of customs; from ephemeral cultural fashions to deep-rooted genetically structured dispositions, such customs are the medium that couple our awareness, intent and capability with the environments that we inhabit. These routine sets of expectations and anticipations provide a context for action; they provide a lure and a constraint upon action, disposing the agent to go down one path or another. Thus, routines and customs have a similar functional role in Whitehead's and Lambert's philosophies.

On the basis of the central role that they grant custom, I have shown that the congruence between Whitehead's metaphysical system and Lambert's computational paradigm extends through into the domain of belief. Lambert builds upon the mathematical representation of the process world initiated by Whitehead and shows how qualitative feelings and the subjective process of concrescence can be represented in formal terms. To believe is to classify. Lambert's semantics of success is defined in terms of patterns of activity that unfold in accordance with expectations. Such pattern-matching activity defines machine belief.

In the journey from a Classical to a Post-Classical understanding of the role of logic in intelligent systems, and the subsequent development of a Post-Classical blue-print for high-level information fusion, the focus of analysis has shifted from logical operations in isolated mechanical minds, to communications between networks of human and machine agents collectively constituting hybrid social organizations. As reflected in Lambert's work, this has progressively transformed AI from a discipline concerned with abstract symbol systems, to a discipline grounded empirically in the broader context of human action: human society.

Human society is the ground of Lambert's semantics of success. The semantics of Post-Classical AI are defined by successful classification: to classify is to believe. However this success is mediated by the machine agent's participation in the real world of purposeful human action. Valid classifications are those that realize the desired real world consequences. Lambert's semantics of success is mediated by the adequacy of the routines mediating agent participation in real world situations, and complex communication feedback loops involving both human and machine agents. Thus, a classification process defines the beliefs that express machine awareness; from awareness comes intent, and from the flow of intent comes system capability.

For Lambert this flow of intent, whether transmitted by human or machine agents, is magnifying system capability only if the underlying awareness is valid. This validity is ultimately a function of routines and the purposive dispositions that these routines define; thus the ultimate grounds of the systems capabilities are routines that are well-fitted to the fusion system's environment. Lambert's Mephisto ontology is crucial to an adequate representation of the system environment.

My concluding observations in the previous chapter focused on the social level of the Mephisto ontology; this ontology gives expression to presuppositions concerning the nature of human society. I argued that the conceptualization of society created by Lambert and Nowak constitutes a useful beginning but that much work is still required. To date the conceptualization has been driven by a vision of society that has not tapped into the conceptualizations developed by social theorists. Although Lambert recognizes that the social domain is vital to the development of his system, it is the least developed level of his ontology.

I have argued that in order to develop an adequate specification of the social domain it is necessary to go beyond Lambert's cognitively inspired view of commonsense action. While Whitehead's process philosophy points the way, it alone does not provide the conceptual resources. As Talcott Parsons observed, while Whitehead's generic categories provide a meta-theoretical framework for conceptualizing human society, they are inadequate for dealing with the specificities that determine the how, when, where and what of who connects with whom. What is required is a more detailed examination of the way that social environments motivate and constrain the flow of intent. This, I have argued, signals the need for Lambert's Post-Classical paradigm to take a Socionic turn.

Just as Lambert reframed the conceptual context of AI by philosophically redefining notions such as existence, representation, belief and mind, I now wish to extend the Post-Classical paradigm to better encompass the notions of society and the civilizing processes that are shaping our societies. These broader issues have been part of the dormant background to Lambert's Whitehead-Hume vision of custom mediated 'commonsense'. The Mephisto system, in appropriating and extending the presuppositions of the Post-Classical paradigm has disturbed these sleeping giants, and they now have to be accommodated explicitly within the Post-Classical vision of a socially situated AI.

To achieve this aim, I hope to extend this work by engaging Lambert's Post-Classical AI in a three sided Socionic dialogue with Post-Parsonian sociological theory and Whitehead's process philosophy. The base of this Socionic triangle - its foundation - is philosophical. My strategy will be to reinterpret both Lambert's and Parsons' work within a conceptual framework based on Whitehead's metaphysical cosmology. Whitehead's philosophy is an ideal partner for facilitating the encounter between Lambert and Parsons. In this work I have laid out meta-theoretical connections between the work of Whitehead and Lambert. I have previously demonstrated that similar links tie Whitehead's metaphysical system to Parsons' theory of society (Dawson 1991). The opportunity now exists to complete the triangle.

The prospects for a flow of ideas between Parsons and Lambert are manifold, beginning with their existential process foundations and proceeding through to their focus on the flow of intent/communications that steer social action. In addition, Parsons' ubiquitous influence on the development of twentieth century social theory means that his work provides a point of entry into the central theoretical debates that have shaped modern social theory. By engaging with the work of Parsons, Lambert's Post-Classical paradigm has the opportunity to engage with the wider sociological field of empirical and theoretical research. This is not the place to explore the ways in which this engagement could advance the Mephisto conceptualization; however the potential benefits of recruiting the insights of a century of scientific effort directed toward conceptualizing society should be obvious, and have a particular relevance in illuminating the broader questions that inspired this work.

A final word on those broader questions; in my introduction I suggested that high-level information fusion needs to be understood in relation to both narrow and broad understandings of the role of Logic and Reason in the formation of Western Civilization. I believe that sociological theory provides an essential perspective on these broader questions and, as this work has argued, that a narrow logical understanding of 'commonsense reasoning' must ultimately confront these broader issues. This, I have argued, points the way toward the Socionic future of AI. Many other developments are contributing to this Socionic future, including theoretical, technological and social imperatives. However information fusion, as outlined in this work, signals a unique and powerful convergence of these forces. Given incremental progress in the effort to engineer effective high-level information fusion software, it would appear that we are moving toward the threshold of a technological revolution that will have an impact upon our collective lives of at least the order of magnitude occasioned by the spread of electricity grids, or the proliferation of automobiles and highways. Already machine intent is playing a significant role in steering transactions on Global financial markets; however machines are blind. If the blindness of these machines is not to be a metaphor for our progress into the coming era of semantic machines then the human sciences and the computer sciences must make a more concerted effort to join forces in shaping our situation awareness.

As Lambert argues, s*ituation awareness is always a state of human awareness*; we determine the constraints that define the semantic intent of machines. Since humans invented the art of writing the development of new technologies for communicating semantic intent have played a crucial role in shaping human history. While symbol systems have always played a vital role in mediating commercial transactions they have also played a vital cultural role in promoting broader reflections upon our human situation and communicating those reflections to others. Development of ontology driven MAS represents a new type of opportunity to represent a vision of the human situation. A Whiteheadian approach to AI could mediate the development of a new generation of ontology driven computer games, in which a flow of human and machine generated intent interactively builds a narrative horizon that all game participants, human and machine, are responding to and attempting to shape. Lambert's high-level information fusion technology could therefore be at the core of a new form of cultural production.

Consideration of the broader cultural implications of high-level information fusion technology has brought this work to an intellectual frontier where the concerns of the human sciences and the computational sciences decisively join forces. Lambert’s work suggests a promising strategy for engineering agents and social situations that enable us to simulate and explore our human dispositions and motivations; one which shifts the focus away from bottom-up physical determinants of action to the top-down role of ideas. When the role that horizons of meaning play in determining social processes comes to the fore, the focus of attention shifts towards those ideas and ideals that have inspired and shaped our thinking. This, according to my Whiteheadian reading of Lambert’s Post-Classical paradigm points the way to a profound challenge: that of enlisting machine agents in the *Adventures of Ideas* that have shaped human history. Why not engineer an ontological system that reflects the ideal of Reason in the deep sense honoured by Whitehead and Husserl in their reflections upon the nature of the civilizing process?

BIBLIOGRAPHY

Abbott, A., 2001. *Time Matters,* Chicago: University of Chicago Press.

Abbott, A. 2005. "Ecologies and Fields", unpublished appendix to "Linked Ecologies", *Sociological Theory*, 23:3, pp.245-274. Available at: http://home.uchicago.edu/~aabbott/Papers/BOURD.pdf

Abbot, A. 2005. “History, culture, and computation: Tough questions about the relationship between models and human behaviour”, in *Threat Anticipation: Social Science Methods and Models, The Joint Threat Anticipation Center Workshop*, April 7-9, The University of Chicago, viewed 2 June 2008. Available at: http://jtac.uchicago.edu/conferences/05/.

Abbott, A. 2006. "The Concept of Order in the Processual Sociology", in Morrisey, R. (ed.) 2006, *Cahiers Parisiens Parisian Notebooks,* No. 2., University of Chicago, pp. 315-344.

Abbott, A. 2007. "'Mechanisms and Relations': A Response to the Comments" *Sociologica* No. 2, (Italian Journal of Sociology on line). Available at: http://www.sociologica.mulino.it/journal/article/index/Article/Journal:ARTICLE:99

Ackerman, C., and Parsons, T. 1967 “The Concept of ‘Social System’ as a Theoretical Device,” in G. J. DiRenzo, ed., *Concepts, Theory, and Explanation in the Behavioral Sciences.* New York: Random House.

Allen, B. 2004. *Knowledge and Civilization,* Boulder, Colorado: Westview Press.

Agre, P. 1993. “The symbolic worldview: Reply to Vera and Simon”, *Cognitive Science,* 17 (1): 61-69.

Agre, P. 1995. “The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project”, *Stanford Humanities Review,* 4(2), pp.1-19.

Agre, P. 1997a. *Computation and Human Experience*, Cambridge: Cambridge University Press.

Agre, P. 1997b. "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI", in Bowker, G., Gasser, L., Star, L., and Turner, B., (eds.), 1997 *Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide*. New Jersey: Erlbaum. Available at: http://polaris.gseis.ucla.edu/pagre/critical.html

Agre, P. and Chapman, D. 1987. “Pengi: An Implementation of a Theory of Activity”. *Proceedings of AAAI 1987*: 268-272.

Agre, P. and Chapman, D., 1990. "What Are Plans For?" in Maes, P. (ed.), 1990 *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*, Cambridge, MA: MIT Press, pages 17-34.

Agre, P. and Horswill, I. 1997. "Lifeworld Analysis", *Journal of Artificial Intelligence Research* 6(1), pages 111-145.

Anderson, M. L. 2003. “Embodied Cognition: A Field Guide”, *Artificial Intelligence,* Vol. 149: pp. 91-130.

Anderson, R. and Lawrence, C. (eds.) 1987. *Science, Medicine and Dissent: Joseph Priestley (1733-1804),* London: Wellcome Trust.

Andrews, T. Broughton, M.and Estival, D. 2006. “Implementing an Intelligent Multimedia Presentation Planner using an Agent Based Architecture”. *Workshop on Effective Multimodal Dialogue Interfaces.* IUI Conference. Sydney, 29 January 2006.

Archer, M. 2000. 'Homo economicus, homo sociologicus, homo sentiens', in Archer M. and Tritter J. (eds.) 2000. *Rational Choice Theory: Resisting Colonization,* Routledge, London.

Archer M. and Tritter J. (eds.) 2000. *Rational Choice Theory: Resisting Colonization,* Routledge, London.

Arnason, J. P. 2001. “Civilizational Patterns and Civilizing Processes”, *International Sociology*, Vol. 16, No. 3. pp. 389-407.

Arnason, J. P. 2003. *Civilizations in Dispute: Historical Questions and Theoretical Traditions*. Leiden and Boston: Brill.

Arsenault, D. 2005. “Abstract of Dynamic Range: When Game Design and Narratives Unite”, in *Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play*. Available at: http://www.gamesconference.org/digra2005/

Ashmore, J. 1974. “Essence in recent philosophy: Husserl, Whitehead, Santayana”, *Philosophy Today,* Vol. 18, No. 3 / 4, pp. 198-210.

Axelrod, R. 1997. “Advancing the Art of Simulation in the Social Sciences”, Published in Rosario Conte, Rainer Hegselmann and Pietro Terna (eds.), *Simulating Social Phenomena,* Berlin: Springer, pp. 21-40.

Baader, F., Bauer, A., Baumgartner, P., Cregan, A., Gabaldon, A., Ji, K., Lee, K. & Schwitter, R. 2008. “A Novel Architecture for Situation Awareness”, NICTA publication. Available at: http://nicta.com.au/__data/assets/pdf_file/0019/15472/AI-08-submission-long-version.pdf.

Baez-Barranco, J-A. Stratulat, T. and Ferber, J. 2006. “A Unified Model for Physical and Social Environments”, paper delivered at E4MAS 2006: The Third International Workshop on Environments for Multiagent Systems, held at the Future University-Hakodate, Japan, May 8, 2006. Available at: http://www.cs.kuleuven.ac.be/~distrinet/events/e4mas/2006/

Baillie-de Byl, P. 2004. *Programming Believable Characters for Computer Games,* Hingham, Massachusetts: Charles River Media.

Baldoni, M. Boella, G. and L. van der Torre, 2006. “Modelling the Interaction Between Objects: Roles as Affordances”, paper to be presented at First International Conference on Knowledge Science, Engineering and Management (**KSEM'06**). August 5-8, 2006, Guilin, China. Available at: http://agamemnon.uni.lu/ILIAS/vandertorre/papers/ksem06.pdf.

Banach, D. M. 1987. *Representing and Knowing,* Southern Illinois University at Carbondale, Ph.D. Dissertation.
Available at: http://www.anselm.edu/homepage/dbanach/RepresentingandKnowing/

Bartle, R. A. 2004. *Designing Virtual World,* Berkeley, California: New Riders.

Barwise, J. and Perry, J. 1983 *Situations and Attitudes*. Cambridge, MA.: The MIT Press.

Baum, R. C. and Lidz, V. M. 1976. “Metatheory: Introduction”, in Loubser, J. J., Baum, R. C., Effrat, A. and Lidz, V. M. (eds.) 1976 *Explorations in General Theory in Social Science,* New York: Free Press, 26-34.

Baynes, K., Bohman, J., & McCarthy, T. (eds.) 1987. *After Philosophy: End or Transformation?* Cambridge, Mass: The MIT Press.

Bellah, R. N. 2005. "What is Axial about the Axial Age?", *Archives Europeennes de Sociologie*, 46(1): 69-89.

Bechtel, W. and Abrahamsen, A. 2005. “Explanation: A Mechanistic Alternative”, *Studies in History and Philosophy of the Biological and Biomedical Sciences,* Vol.36, pp.421-441.

Bell, D. 1999. “The Revolution of Moore and Russell: A Very British Coup?” in O’Hear, A. (ed.) *German Philosophy Since Kant*, Cambridge: Cambridge University Press, pp.193-208.

Bennet, J. B. 1975. “Husserl’s *Crisis* and Whitehead’s Process Philosophy”, *The Personalist*, Vol. 56, pp.289-300.

Berners-Lee, T. & Fischetti, M. 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*, San Francisco, CA: Harper.

Bickhard, M. H. 1993. Representational Content in Humans and Machines. *Journal of Experimental and Theoretical Artificial Intelligence*, *5*, pp.285-333.

Bickhard, M. H. 2004. “Process and Emergence: Normative function and representation”, *Axiomathes* 14: 121–155.

Bickhard, M. H. 2008. “Social Ontology as Convention”, *Topoi* ,27:139–149.

Bickhard, M. H. 2009a. “Interactivism: introduction to the special issue”, *Synthese,* 166: 449–451.

Bickhard, M. H. 2009b. “The interactivist model”, *Synthese,* 166:547–591.

Bickhard, M. H. 2009c. ‘Interactivism: A Manifesto”. *New Ideas in Psychology*, 27: 85–9.

Block, N. 1998. “Semantics, conceptual role”, in E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*, vol. 8, pp. 652-657. Available at - http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/ConceptualRoleSemantics.html#anchor696459

Bloodsworth, P., Greenwood, S., 2005. "COSMOA an ontology-centric multi-agent system for co-ordinating medical responses to large-scale disasters." *AI Communications,* Vol. 18 (3) pp. 229-240.

Boden, M. A. 2006. *Mind as Machine: A History of Cognitive Science*, 2 vols. Oxford: Oxford University Press.

Bogost, I. 2006. *Unit Operations: An Approach to Videogame Criticism,* Cambridge, Massachusetts: The MIT Press.

Bourdieu, Pierre. 1981. "Men and machines". In *Advances in Sociological Method and Methodology,* ed. K. Knorr-Cetina and A.V. Cicourel. London: Routledge and Kegan Paul, pp. 304-17.

Bourdieu, P. 1998. *Practical Reason: On the Theory of Action,* Stanford, Calif.: Stanford University Press.

Bourdieu, P. 2003. *Firing Back: Against the Tyranny of the Market 2,* New York: The New Press.

Bourdieu, P. & Wacquant, L. J. D. 1992. *An Invitation to Reflexive Sociology,* Chicago: The University of Chicago Press.

Bourricaud, F. 1981. *The Sociology of Talcott Parsons,* Chicago: The University of Chicago Press.

Bracken, J. 1991. *Society and Spirit: A Trinitarian Cosmology.* Cranbury: Associated UP.

Braeckman, A. 1985. “Whitehead and German Idealism: A Poetic Heritage”, *Process Studies*, Vol. 14, pp.265-286.

Bratko, I. 2001. *PROLOG, Programming for Artificial Intelligence*, 3rd ed., Addison-Wesley.

Bratman, M. E., Israel, D. J. and Pollack M. E., 1988. "Plans and Resource-Bounded Practical Reasoning." *Computational Intelligence* 4(4): 349-355.

Brauer, W., Nickles, M., Robatsos, M., Weiss, G., Lorentzen, K., 2001. "Expectation-Oriented Analysis and Design", *AOSE 2001*: 226-244. Available at: http://citeseer.ist.psu.edu/494231.html

Brent, E. and Thompson, G. A. 1999. "Sociology: Modeling Social Interaction with Autonomous Agents", *Social Science Computer Review*, Vol. 17, No. 3, pp. 342-357.

Breuvart, J.M. 2006. "Husserl and Whitehead, mathematicians, on intentionality", paper delivered at Process Philosophy Conference, Salzberg, 17 July 2006. Available at: http://www2.sbg.ac.at/whiteheadconference/abstracts/Breuvart-HUSSERL%20AND%20WHITEHEAD.doc.pdf.

Brom, C. 2005. "GAL: A framework for large artificial environments inhabited by human-like intelligent virtual agents". Available at: http://urtax.ms.mff.cuni.cz/ive/public/docs/ive_TechRep_200503.pdf.

Brooks, R. A. 1990. "Elephants Don't Play Chess". In Maes, P. (ed.) 1990 *Designing Autonomous Agents*. Cambridge, MA: MIT Press.

Brooks, R.A., 1991a. "How to build complete creatures rather than isolated cognitive simulators," in K. VanLehn (ed.), *Architectures for Intelligence*, Hillsdale, NJ: Lawrence Erlbaum Assosiates, pp. 225-239.

Brooks, R. 1991b. "Intelligence Without Representation", *Artificial Intelligence Journal* 47, pp. 139–159.

Brooks, R. 1991c. "Intelligence Without Reason", *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, Sydney, Australia, (August 1991), pp. 569–595.

Broughton, M., Carr, O., Taplin, P., Estival, D., Wark, S., & Lambert, D. 2002. "Conversing with Franco, FOCAL's Virtual Adviser", DSTO publication, available at: http://hdl.handle.net/1947/2318

Brown, G. S. 1979. *Laws of Form* (2nd edn.) New York: E. P. Dutton.

Brown, J. 1982. "Poetry and the Possibility of Theology: Whitehead's Views Reconsidered", *Journal of the American Academy of Religion*, Vol. 50 No. 4, pp. 507-520

Browning, D. & Myers, W. T. 1998. *Philosophers of Process,* New York: Fordham. University Press.

Campbell, R. 2009. "A process-based model for an interactive ontology", *Synthese* 166:453–477.

Capek, M. 1964. "Simple Location and Fragmentation of Reality", *The Monist*, 48:195-218.

Cardelli, L. & Wegner, P. 1985. “On understanding types, data abstraction, and polymorphism”, *ACM Computing Surveys*, 17/4, pp.471-522.

Carley, K. M. 2001. “Computational Approaches to Sociological Theory”, in Turner, J. (ed.) *Handbook of Sociological Theory*, New York: Kluwer Academic, pp. 69-84.

Castelfranchi, C. 2006. “Cognitive Architecture and Contents for Social Structures”, in Sun, R. (ed.) 2006 *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Cambridge: Cambridge University Press, pp. 355-392.

Castronova, E. 2005a. *Synthetic Worlds: The Business and Culture of Online Games*, Chicago: University of Chicago Press.

Castronova, E. 2005b. "On the Research Value of Large Games: Natural Experiments in Norrath and Camelot" (Dec. 2005). CESifo Working Paper Series No. 1621. Available at SSRN: http://ssrn.com/abstract=875571

Cederman, L-K. 2005. “Computational Models of Social Forms: Advancing Generative Process Theory”, *American Journal of Sociology,* Vol. 110, No. 4, pp. 864-93.

Chait, R., Sciarretta, A., Lyons, J., Barry, C., & Shorts, D., 2007. *A Further Look at Technologies and Capabilities for Stabilization and Reconstruction Operations,* Defense & Technology Paper No. 43, published by the National Defense University Center for Technology and National Security Policy, Fort Lesley (J. McNair, Washington, DC). http://www.ndu.edu/ctnsp/Def_Tech/DTP%2043%20SROII.pdf

Chaitin, G. 2005. *Meta Math! The Quest for Omega*, NY: Pantheon Books.

Chaitin, G. 2006. “The limits of reason,” *Scientif c American*, Vol.294, No.3, pp. 74-81.

Chaitin G. J. 2007. *Thinking About Godel and Turing: Essays on Complexity, 1970-2007*, Singapore: World Scientific,

Chang, P. H-M., Chen, K-T., Chien, Y-H., Kao, E., and Soo, V-W. 2005. “From Reality to Mind: A Cognitive Middle Layer of Environment Concepts for Believable Agents”, in Weyns, D. and Parunak, H.V.D. and Michel, F., (eds.), *Environments for multi-agent systems*, *Lecture Notes in Computer Science*, vol. 3374, pp.57-73.

Clancey, W. 1994. “Review of *The Invention of Memory: An New View of the Brain*, by Israel Rosenfield”, in Clancey, W. Smoliar, S. W. and Stefik, M. J. (eds.) *Contemplating Minds: A Forum for Artificial Intelligence,* Cambridge, Massachusetts: The MIT Press, pp. 387-430.

Clancey, W. 1997. *Situated Cognition: On human knowledge and computer representations.* Cambridge: Cambridge University Press.

Clancey, W. 1999. *Conceptual Coordination: How the Mind Orders Experience in Time,* Mahwah, N.J.: Lawrence Erlbaum Associates.

Clancey, W. Sierhuis, M, Damer, B. and Brodsky, B. 2006. “Cognitive Modeling of Social Behaviors”, in Sun. R, (ed.) 2006 *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Cambridge: Cambridge University Press, pp. 151-185.

Clark, A. 2001. *Mindware: And Introduction to the Philosophy of Cognitive Science,* Oxford: Oxford University Press.

Clark, A. 2004. *Natural-Born Cyborgs*. Oxford: Oxford University Press.

Clark, A. & Chalmers, D. 1998. "The Extended Mind," *Analysis* Vol. 58/1: 7-19. Available at http://www.philosophy.ed.ac.uk/staff/clark/pubs/TheExtendedMind.pdf

Clarke, B. L. 1981. “A Calculus of Individuals Based on Connection”, *Notre Dame Journal of Formal Logic*, 22: 204-218.

Cobb, J. B. 1971. “The ‘Whitehead Without God’ Debate: the Critique”, *Process Studies*, Vol. 1, No. 2, pp.91-100.

Cohn, A. G., and Varzi, A. C. 2003, “Mereotopological Connection,” *Journal of Philosophical Logic*, Vol.32: pp.357-90.

Coleman, J. S. and Fararo, T. J. (eds.) 1992. *Rational Choice Theory: Advocacy and Critique,* Newbury Park, Ca.: SAGE Publications.

Combs, N. 2004a. “The intelligence in the MMOG: from scripts to stories to directorial AI”, paper presented at the Other Players conference, Center for Computer Games Research, IT University of Copenhagen, Denmark. 6-8 December, 2004. (http://www.itu.dk/op)

Combs, N. 2004b. “Artificial intelligence and the narrative experience in the virtual world”, *On the Horizon*, Vol. 12, No. 3, pp. 117-122.

Conte, R.; Edmonds, B.; Moss, S. and Sawyer, R. K. 2001. “Sociology and Social Theory in Agent Based Social Simulation: A Symposium”, *Computational & Mathematical Organization Theory,* 7:183-205.

Copleston, F. 1985. *A History of Philosophy*, Vol. 8, New York: Doubleday.

Cornwell, J., O'Brien, K., Silverman, B. G., & Toth, J., 2003. “Affordance theory for improving the rapid generation, composability, and reusability of synthetic agents and objects.” *Proceedings of the Twelfth Conference on Behavior Representations in Modeling and Simulation (BRIMS, formerly CGF)*, SISO. 12-15 May 2003. Available at: http://www.seas.upenn.edu/~barryg/BRIMS2003_AffordanceTheory.doc.

Cramer, C. 2002. “Homo Economicus Goes to War: Methodological Individualism, Rational Choice and the Political Economy of War”, *World Development* Vol. 30, No. 11, pp. 1845-1864.

Crannell M., Howard N., Norwood G., Tait A. 2005. "A C2 system for Winning hearts and minds: tools for Confrontation and Collaboration Analysis", *Proceedings of the Tenth International Command and Control Research and Technology Symposium*, McLean, VA, 13-16 June 2005. (2005). http://www.andrewtait.com/papers/ccrts2005.pdf

Cull, P. 2007. "The mathematical biophysics of Nicolas Rashevsky", *BioSystems,* Vol.88, No.3, pp.178-184.

Cutter, A. 2005. "Peace building: a literature review", *Development in Practice*, Volume 15, Number 6, pp.778-784.

Davidsson, P. 2002. "Agent Based Social Simulation: A Computer Science View", *Journal of Artificial Societies and Social Simulation* vol. 5, no. 1. Available at: http://www.soc.surrey.ac.uk/JASSS pp. 1-5.

Davis, D. N. 2004a. "Why do anything? Emotion, affect and the fitness function underlying behaviour and thought", paper presented at *Affective Computing AISB*, University of Leeds, UK. Available at: http://www.dcs.hull.ac.uk/NEAT/dnd/papers/aisb04.pdf.

Davis D.N. 2004b. "Affect and Affordance: Architectures without Emotion", AAAI2004, Spring Symposium, Stanford, USA.

Davis D.N. 2008. "Linking perception and action through motivation and affect", *Journal of Experimental & Theoretical Artificial Intelligence*, Vol. 20, No.1, pages 37 - 60.

Davis D.N. & Lewis, S.C. 2003, "Computational Models of Emotion for Autonomy and Reasoning", *Informatica* 27(2):159-165.

Dawson, A. S. 1991. *Parsons and Whitehead: A study of the relationship between social theory and metaphysics*, (unpublished M.A. thesis), Bundoora, VIC: La Trobe University.

Dawson, A. S. 2008. "Universal Algebra", in Weber, M. & Desmond, W. 2008 *Handbook of Whiteheadian Process Thought*, [Vol. 2], Frankfurt: Ontos Verlag, pp.67-86; 149-156.

Dennett, D. C. 1971. "Intentional Systems", *The Journal of Philosophy,* Vol. 68, No. 4. pp. 87-106.

Dennett, D. C. 1991. *Consciousness Explained,* London: Penguin.

de Regt, H. W. 2006. "Wesley Salmon's Complementarity Thesis: Causalism and Unificationism Reconciled?", *International Studies in the Philosophy of Science*, 20:2, pp.129 – 147.

Devlin, K. 1997. *Goodbye Descartes: The end of logic and the search for a new cosmology of mind,* New York: John Wiley.

Devlin, K. 2006. “Situation theory and situation semantics”, in Woods, J. (Ed.), *Handbook of the History of Logic*, vol. 7., Elsevier, pp. 601-664. URL = http://www.stanford.edu/~kdevlin/HHL_SituationTheory.pdf

Dewey, J. 1931. *Philosophy and Civilization*, New York: Putnam.

Dewey, J. 1951. “The Philosophy of Whitehead”, in Schilpp, P. A. (ed.) 1951 *The Philosophy of Alfred North Whitehead*, La Salle: Open Court, pp.643-661.

Dewey, J. 1958a. [1925] *Experience and Nature* (Revised) New York: Dover.

Dewey, J. 1958b. [1934] *Art as Experience*, New York : Capricorn Books.

Doran, P. 2006. “Ontology Reuse via Ontology Modularisation”. In *Proceedings of KnowledgeWeb PhD Symposium 2006,* (KWEPSY2006) Budva, Montenegro, 17th June 2006. Available at: http://www.l3s.de/kweb/kwepsy2006/FinalSubmissions/kwepsy2006_doran.pdf

Dourish, P. 2004 [2001]. *Where the Action Is: The Foundations of Embodied Interaction,* Cambridge, Massachusetts: MIT Press.

Dowden, B. & Swartz, N. 2006. “Truth: Tarski’s Semantic Theory”, *The Internet Encyclopedia of Philosophy* http://www.iep.utm.edu/t/truth.htm#H4

Dray, A. Perez, P. Jones, N. Le Page, C. D'Aquino, P. White, I. and Auatabu, T. 2006. “The AtollGame Experience: from Knowledge Engineering to a Computer-Assisted Role Playing Game”. *Journal of Artificial Societies and Social Simulation,* 9(1)6. Available at: http://jasss.soc.surrey.ac.uk/9/1/6.html .

Dreyfus, H.L 1972. *What Computers Can't Do: A Critique of Artificial Reason*. New York: Harper & Row.

Dreyfus, H. L. 1992. *What computers still can’t do: A critique of artificial reason*. Cambridge, MA: MIT Press.

Dreyfus, H.L. 2007 “Why Heideggerian AI failed and how fixing it would require making it more Heideggerian”, *Philosophical Psychology,* 20(2), 247-268.

Drummond, J. J. 1995. "Moral Objectivity: Husserl's sentiments of understanding", *Husserl Studies*, 12/2:165-183.

Dummett, M. 1994. *Origins of Analytic Philosophy*, Cambridge, Mass: Harvard University Press.

Edling, C. R. 2002. "Mathematics in Sociology", *Annual Review of Sociology*, August 2002, Vol. 28, Pages 197-220.

Egenfeldt-Nielsen, S. 2005. *Beyond Edutainment: Exploring the Educational Potential of Computer Games*, Ph. D. thesis, IT-University of Copenhagen. Available at: http://itu.dk/people/sen/egenfeldt.pdf

Eisendrath, C. 1971. *The Unifying Moment: The Psychological Philosophy of William James and Alfred North Whitehead*, Cambridge, Mass.: Harvard University Press.

Eisenstadt, S. N. 1986. "The Axial Age breakthroughs: Their characteristics and origins", in Eisenstadt, S. N. (ed.), 1986 *The Origin and Diversity of Axial Age Civilizations.* Albany, N.Y.: SUNY Press, pp.1-28.

Eisenstadt, S. N. 2004. "Interdisciplinary reflections on the civilizational dimensions of modernity", in Camic, C. & Joas, H. (eds.) *The Dialogical Turn: New Roles for Sociology in the Post-Disciplinary Age*, Lanham: Rowman & Littlefield, pp.275-295.

Elias, N. 1978. *What Is Sociology?* translated by Stephen Mennell and Grace Morrissey. New York: Columbia University Press.

Elias, N. 2000. *The Civilizing Process: Sociogenetic and Psychogenetic Investigations*, (Revised Edition). Translated by E. Jephcott, edited by. E. Dunning, J. Goudsblom and S. Mennell. Oxford: Blackwell Publishers.

Elster, J. 1989. “Social Norms and Economic Theory”, *Journal of Economic Perspectives*, 3 (4): 99-117.

Elster, J. 1994. “Rationality, emotions, and social norms”, *Synthese*. 98, pp.21-49.

Elster, J. 1996. “Rationality and the Emotions”, *The Economic Journal*, 106 (438), pp.1386-1397.

Emery, D. & Bassett, D. 2005. "SOSCOE-The Glue That Holds FCS Together". *Army AL & T* September-October 2005, pp. 21-24. Available at: http://www.dsci-usa.com/document/Emeryarticle.pdf

Emmet, D. 1996 “Whitehead”, *Philosophy,* 71/275: 101-115.

Endo, H. 1984. “The metaphysics of time in Husserl and Whitehead”, in *Proceedings, International Conference-Process and Reality, East and West.* (Tokyo: The Japan Society for Process Studies, 1984).

Endo, H. 1990. “From Lorenzen through Husserl to Whitehead”, *Process Studies,* 19/4:235-239. Available at http://www.religion-online.org/

Endsley, M. R. 1995. “Toward a Theory of Situation Awareness in Dynamic Systems”, *Human Factors* Vol. 37 No. 1 pp.32 – 64.

Epstein, J. M. & Axtell, R. 1996. *Growing Artificial Societies: Social Science from the Bottom Up,* Cambridge, Mass.: MIT Press.

Estival, D. Nowak, C. & Zschorn. A. 2004. "Towards Ontology-based Natural Language Processing". *RDF/RDFS and OWL in Language Technology: 4th Workshop on NLP and XML (NLPXML-2004), ACL 2004*, Barcelona, Spain.

Fararo, T. J. 1989. *The Meaning of General Theoretical Sociology: Tradition and Formalization*. New York: Cambridge University Press.

Fararo, T. 1997. "Reflections on Mathematical Sociology", *Sociological Forum*, Vol, 12 No. 1, pp.73-101.

Fararo, T., 2001. *Social Action Systems: Foundation and Synthesis in Sociological Theory.* Westport, CT: Praeger.

Farrell, L. P. 2004. "'Information Fusion' Key to Winning Wars", *National Defense Industrial Association (NDIA) – President's Corner - May 2004* http://www.ndia.org/Content/NavigationMenu/Resources1/Presidents_Corner2/May_2004.htm

Favell, A. 2005. "The Sim-Sociologist: Using Computer Games to Teach Sociological Theory", *Teaching with Technology.* Available at: http://www.college.ucla.edu/edtech/interviews/favell.htm

Feigenbaum, E. and Feldman, J. (eds.) 1963. *Computers and Thought*, New York: McGraw-Hill.

Ferber, J. 1999. *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence,* Harlow, England: Pearson Education.

Fischer, F. & Nickles. M. 2006. "Computational Opinions". *Procs. of the 17th European Conference on Artificial Intelligence* (ECAI'06), 2006.

Fischer, K. Muller, J.P. & Pischel, M. 1996. "Cooperative transportation scheduling: an application domain for DAI", *International Journal of Applied Artificial Intelligence* 10 (1): 1–33.

Fischer, K. Florian, M. & Malsch, T. (eds.), 2005. *Socionics*. Lecture Notes in Computer Science, Vol. 3413. Springer-Verlag.

Fish, J. S. 2004. "The Neglected Element of Human Emotion in Talcott Parsons's *The Structure of Social Action*", in *Journal of Classical Sociology*, 4/1:115-134.

Fish, J. S. 2005. *Defending the Durkheimian Tradition: Religion, Emotion and Morality.* Aldershot: Ashgate.

Fitzpatrick, M. 1987. "Joseph Priestley and the Millenium", in Anderson, R. and Lawrence, C. (eds.) 1987, *Science, Medicine and Dissent: Joseph Priestley (1733-1804),* London: Wellcome Trust.

Fitzgerald, J. A. 1979. *Alfred North Whitehead's Early Philosophy of Space and Time*, Washington, D.C.: University Press of America.

Fletcher, J. 1997. *Violence and Civilization: An Introduction to the work of Norbert Elias,* Cambridge: Polity Press.

Floyd, J. 2001. "Number and Ascriptions of Number in Wittgenstein's *Tractatus*", in *Future Pasts: The Analytic Tradition in Twentieth Century Philosophy.* Edited by Floyd, J. and Shieh, S., Oxford: Oxford University Press, pp. 145-191.

Floyd, J. 2005. "Wittgenstein on Philosophy of Logic and Mathematics", in *The Oxford Handbook of Mathematics and Logic.* Edited by Shapiro, S., Oxford: Oxford University Press, pp. 75-128.

Ford, L. S. 1987. "Recent interpretations of Whitehead's Writings", *The Modern Schoolman,* LXV (November, 1987): 47-59

Ford, L. S. 1997. "Nobo's Eternal Realities and the Primordial Decision", *Process Studies,* 26/3-4: 205-217.

Ford, L. S. 2000. *Transforming Process Theism,* Albany: State University of New York Press.

Frasca, G. 2001. *Videogames of the Oppressed: Videogames as a means for Critical Thinking and Debate,* MA thesis, Georgia Institute of Technology.

Fraser, M. 2006. "The ethics of reality and virtural reality: Latour, facts and values", *History of the Human Sciences,* 19/2: 45-72.

Frege G.1956 [1918]. "The Thought: A Logical Inquiry", *Mind*, New Series, Vol. 65, No. 259. (Jul., 1956), pp. 289-311.

Frisina, W. G. 2002. *The Unity of Knowledge and Action*, Albany, N.Y.: SUNY Press.

Froehner, T. Nickles, M. & Weiß. G. 2004. "Towards modeling the social layer of emergent knowledge using open ontologies". *ECAI Workshop on Agent-Mediated Knowledge Management* (AMKM, Workshop Notes pp. 10-19). Available at: http://www7.in.tum.de/~weissg/weissg-pub.html

Froehner, T. Nickles, M. Weiss, G. Brauer, W. & Franken R.2005. "Integration of ontologies and knowledge from distributed autonomous sources". *Special issue on Knowledge Representation, reasoning and management* (ed. by U. Furbach), Künstliche Intelligenz, No. 1/2005, pp. 18-23. Available at: http://www7.in.tum.de/~weissg/weissg-pub.html

Fron, J. Fullerton, T. Morie, J. & Pearce, C. (aka Ludica) 2007. "A Game of One's Own: Towards a New Gendered Poetics of Digital Space." *Digital Arts and Culture 2007*, Perth, Australia.

Gat, A. 2006. *War in Human Civilization,* Oxford: Oxford University Press.

Gay, V. P. 1992. *Freud on Sublimation: Reconsiderations*. Albany: State University of New York Press.

Gazendam, H. & Liu, K. 2005. “The Evolution of Organisational Semiotics -A Brief Review of the Contribution of Ronald Stamper”, in Joaquim Filipe & Kecheng Liu (Eds.). *Studies in organisational semiotics*. Dordrecht: Kluwer Academic Publishers. Available at: http://www.orgsem.org/papers/00.pdf

Gazendam, H. 2007. “Interaction of Simulated Actors with the Environment” in Charrel, P-J. & Galarreta, D. (eds.) 2007 *Project Management and Risk Management in Complex Projects: studies in organisational semiotics,* Netherlands: Springer.

Gee, J. P. 2003. *What Video Games have to Teach us about Learning and Literacy,* New York: Palgrave Macmillan.

Gee, J. P. 2005a. *Why Video Games are Good for the Soul: Pleasure and Learning.* Australia: Common Ground.

Gee, J. P. 2005b. “What would a state of the art instructional video game look like?” *Innovate* Vol.1, No. 6, pp. 1-10. Available at: http://innovateonline.info/

Geertz, C. 1973. *The Interpretation of Cultures.* New York: Basic Books.

Gerhardt, U. 2002. *Talcott Parsons: An Intellectual Biography*, Cambridge: Cambridge University Press.

Ghattas, K. 2008. "New approach for US army manual", *BBC News*, Washington Friday, 8 February 2008. http://news.bbc.co.uk/2/hi/americas/7236054.stm

Gibson, J. 1979. *The Ecological Approach to Visual Perception*. New Jersey: Houghton Mifflin.

Gibson, J. J. 1982. “More on Affordances” On-line memo taken from: E.S. Reed & R. Jones (Eds.) (1982). *Reasons for Realism*. Chapter 4.9, Parts III (pp. 406 - 407) and IV (pp. 407 - 408). Hillsdale, NJ: Lawrence Erlbaum Associates. Available at: http://huwi.org/gibson/moreaff.php

Giffin, R.E., 2002. “Superstitious Rituals: Naïve Inductivism in Command and Control Doctrine: Its Causes, Consequences and Cures”, *Proceedings of the 7th International Command and Control Research & Technology Symposium (ICCRTS),* Quebec City, 16-20 Sept.

Giffin, R.E., Reid, D.J., 2003a. “A Woven Web of Guesses, Canto One: Network Centric Warfare and the Myth of the New Economy”, *Proceedings of the 8th ICCRTS*, Washington DC, 17-19 June.

Giffin, R.E., Reid, D.J., 2003b. “A Woven Web of Guesses, Canto Two: Network Centric Warfare and the Myth of Inductivism”, *Proceedings of the 8th ICCRTS*, Washington DC, 17-19 June.

Gilbert, N. 1999. “Simulation: A New Way of Doing Social Science”, *American Behavioral Scientist,* Vol. 42, No. 10, pp.1485-1487.

Gilbert, N. 2004. "Agent-Based Social Simulation: Dealing with Complexity**".** Available at: http://www.complexityscience.org/NoE/ABSS-dealing%20with%20complexity-1-1.pdf

Gilbert, N. 2005. "Agent-based social simulation: dealing with complexity". Available at http://www.complexityscience.org/NoE/ABSS-dealing%20with%20complexity-1-1.pdf

Gilbert, N. Besten, M. den, Bontovics, A. Craenen, B.G.W. Divina, F. Eiben, A.E. Griffioen, R. Hvzi, G. Lorincz, A. Paechter, B. Schuster, S. Schut, M.C. Tzolov, C. Vogt, P. and Yang, L. 2006. "Emerging artificial societies through learning." *Journal of Artificial Societies and Social Simulation*, 9(2). Available at: http://www.soc.surrey.ac.uk/JASSS pp.1-21

Gilbert, N., Terna P., 2000. "How to build and use agent-based models in social science", *Mind & Society*, Vol.1, No. 1, pp. 57-72. Available at: http://web.econ.unito.it/terna/deposito/gil_ter.pdf.

Gilbert, N. and Troitzsch, K. G. 1999. *Simulation for the Social Scientist,* Buckingham, Philadelphia: Open University Press.

Gier, F. 1976. "Intentionality and Prehension", *Process Studies,* No. 6. pp. 197-213.

Giunchiglia, F., Giorgini, P., Blanzieri, E., and Birukou, A., 2009. "A Formal Definition of Culture" in Sycara, K. P., Gelfand, M., and Abbe, A., (eds.) 2009 *Proceedings of W:24 Modeling Intercultural Collaboration and Negotiation (MICON), at International Joint Conference on Artificial Intelligence (IJCAI'09),* Pasadena CA, July 13 2009. Available at: http://www.cs.cmu.edu/~softagents/papers/IJCAI-09-MICON-Proceedings-final.pdf

Goldspink, C & Kay R. 2004. "Bridging the Micro-Macro Divide: a new basis for social science", *Human Relations*, Vol 57 (5).

Goldspink C. & Gilbert N. (eds.) 2008. *Proceedings of the AISB 2008 Symposium on Agent Cognitive Ability and Orders of Emergence (Vol.6).* Available at http://www.aisb.org.uk/convention/aisb08/

Gonzalo, F. 1999. "Ludology meets narratology: similitudes and differences between (video) games and narrative". Available at: http://www.ludology.org/articles/ludology.htm

Gonzalo, F. 2003. "Ludologists love stories, too: notes from a debate that never took place". Available at: http://www.ludology.org/articles/Frasca_LevelUp2003.pdf

Gordon M. R. 2008. "After Hard-Won Lessons, Army Doctrine Revised", *New York Times*, (February 8, 2008). Available at: http://www.nytimes.com/2008/02/08/washington/08strategy.html?ex=1360213200&en=fbd2aa2e9f83a945&ei=5088&partner=rssnyt&emc=rss

Gotts, N. M., Gooday, J. M., and Cohn, A. G. 1996. "A connection based approach to common-sense topological description and reasoning", *The Monist*, 79/1: 51-75.

Griffiths, P. E. 2004. "Towards a Machiavellian theory of Emotional Appraisal", in Cruse, P. & Evans, D. (eds.) *Emotion, Evolution and Rationality,* Oxford: Oxford University Press, pp.89-105.

Griffiths, P.E. & Scarantino, A. 2007. "Emotions in the Wild: The situated perspective on emotion", in P. Robbins and M. Aydede (eds.) *Cambridge Handbook of Situated Cognition,* Cambridge: Cambridge University Press.

Grob, A. 1973. *The Philosophic Mind: A Study of Wordsworth's Poetry and Thought*, Columbus: Ohio State University Press.

Gruber, T. R. 1995. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing", *International Journal of Human and Computer Studies*, 43(5/6), 907–928.

Gruber, T. R. 2008. "Ontology", in Ling Liu and M. Tamer Özsu (Eds.), 2008. *Encyclopedia of Database Systems*, Springer-Verlag. Available at: http://tomgruber.org/writing/ontology-definition-2007.htm

Guarino, N. 1998. "Formal Ontology and Information Systems", in N. Guarino (ed.), *Formal Ontology in Information Systems.* Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15.

Guy-Vuillème, A. 2004. *Simulation of Nonverbal Social Interaction and Small Groups Dynamics in Virtual Environments,* Ph. D thesis, Lausanne : EPFL.

Guye-Vuillème, A. Thalmann, D. 2000. "Requirements for an Architecture for Believable Agents", Proc. Agents 2000, pp. 48-49. Available at: http://vrlab.epfl.ch/Publications/publications_index.html

Guye-Vuillème, A. Thalmann, D. 2001. "A high-level Architecture For Believable Social Agents", *VR Journal*, Springer, Vol.5, pp. 95-106. http://vrlab.epflHadot, P. 1993 *Plotinus or the Simplicity of Vision*, Chicago: University of Chicago Press.

Hadzic, M. and Chang, E. 2005. "Use of ontology-based multi-agent systems in the biomedical domain", in Dascalu, S. (ed), *18th International Conference on Computer Applications in Industry and Engineering (CAINE),* pp. 141-145, Hawaii, Nov 09 2005.

Hahn, C.; Fley, B.; Florian, M.; Spresny, D. and Fischer, K. 2007. "Social Reputation: A Mechanism for Flexible Self-Regulation of Multiagent Systems", *Journal of Artificial Societies and Social Simulation* (*JASSS*), Vol. 10 (1).

Hales, D. & Edmonds, B. 2004. "Sociologically Inspired Engineering". Available at: http://cfpm.org/papers/sie/

Hales, D. 2005. “Engineering with Sociological Metaphors: Examples and Prospects”. Available at: http://www.delis.upb.de/paper/DELIS-TR-0201.pdf

Harman, P. M. 1998. *The Natural Philosophy of James Clerk Maxwell,* Cambridge: Cambridge University Press.

Harrington, A. 2008. "A Sociology of the Demonic?: Alfred Weber's Conception of `Immanent Transcendence'", *Journal of Classical Sociology* 2008; 8; 89.

Hart, J. G. 1992. *The Person and the Common Life: Studies in Husserlian Social Ethics,* Dordrecht: Kluwer.

Hartshorne, C. 1950. “Whitehead’s Metaphysics”, in Lowe, V., Hartshorne, C., and Johnson. A. H. 1950, *Whitehead and the Modern World; Science, Metaphysics, and Civilization: Three Essays on the Thought of Alfred North Whitehead.* Boston: Beacon Press, pp.25-41.

Hartshorne, C. 1965. “The Development Of Process Philosophy”, in Browning D. (ed.) 1965 *Philosophers of Process,* New York: Random House Inc, pp. v – xxii. In the revised second edition [Browning D. & Myers, W.T. (ed.) 1998 *Philosophers of Process*, New York: Fordham University Press], pp.391-407.

Haugeland, J. 1985. *Artificial Intelligence: The Very Idea,* Cambridge, Massachusetts: The MIT Press.

Hayes, P.J.1979. “The Naïve Physics Manifesto”, in Michie D (ed.), *Expert Systems in the Microelectronic Age*, Edinburgh: Edinburgh University Press, pp.242-70.

Hayes, P.J.1985. “The Second Naïve Physics Manifesto”, in Hobbs, J.and Moore, R.C., (eds), *Formal Theories of the Commonsense World*, Norwood: Ablex, pp.1-36.

Hedstrom, P. 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press.

Hedstrom, P. (with Richard Swedberg), 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press.

Heft, H. 2001. *Ecological Psychology in Context: James Gibson, Roger Barker, and the Legacy of William James’s Radical Empiricism*, Mahwah, N.J.: Lawrence Erlbaum Associates.

Heidegger, M. 1977. “The question concerning technology”, in Krell, D. F. (Ed.), *Martin Heidegger: Basic Writings*, San Francisco: HarperCollins Publishers, pp.283-317.

Heimann, P. M. 1978. “Voluntarism and Immanence: Conceptions of Nature in Eighteenth Century Thought”, *Journal of the History of Ideas*, 39:271-83.

Heinze, C., 2003. *Modelling Intention Recognition for Intelligent Agent Systems*. Melbourne: PhD Thesis, Department of Computer Science and Software Engineering, University of Melbourne.

Helmhout, M. 2006. *The Social Cognitive Actor: A multi-actor simulation of organisations*, Ridderkerk: Labyrint Publications.

Helmhout, M. Gazendam, H. & Rorna, R. 2008. "Control over emergence", in Goldspink C. & Gilbert N. (eds.) 2008 *Proceedings of the AISB 2008 Symposium on Agent Cognitive Ability and Orders of Emergence* (Vol.6), pp.1-8. Available at http://www.aisb.org.uk/convention/aisb08/

Hempel, C. G. 1965. *Aspects of Scientific Explanation*. New York: Free Press.

Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, McElreath R, Alvard M, Barr A, Ensminger J, Henrich NS, Hill K, Gil-White F, Gurven M, Marlowe FW, Patton JQ, Tracer D., 2005. "'Economic man' in cross-cultural perspective: Ethnography and experiments from 15 small-scale societies". *Behavioral and Brain Sciences*, 28, 795-855Hexmoor, H. Venkata, S. G. and Hayes, D. 2006 "Modelling social norms in multi-agent systems", *Journal of Experimental and Theoretical Artificial Intelligence,* Vol. 18, No. 1, pp. 49-71.

Henry, G. C. 1993. *Forms of Concrescence: Alfred North Whitehead's Philosophy and Computer Programming Structures,* Lewisburg: Bucknell University Press.

Henry, G. C. and Geertsen, M. 1986. "Whiteheadian Philosophy and Prolog Computer Programming", *Process Studies,* vol.15, no. 3: pp.181-91.

Henry, G. C. and Valenza, R. J. 1993. "Whitehead's Early Philosophy of Mathematics", *Process Studies*, 1/22, pp. 21-36.

Henry, G. C. and Valenza, R. J. 2008. "Computer Science: Form without Content", in Weber, M. & Desmond, W. 2008 *Handbook of Whiteheadian Process Thought*, [Vol. 2], Frankfurt: Ontos Verlag, pp.193-203.

Hexmoor, H. Venkata, S. G. and Hayes, D. 2006. "Modelling social norms in multi-agent systems", *Journal of Experimental and Theoretical Artificial Intelligence,* Vol. 18, No. 1, pp. 49-71.

Hodges, W. 2005a. "Model Theory", Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (on-line edition, June 2009). Available at: http://plato.stanford.edu/entries/model-theory/

Hodges, W. 2005b. "First-order Model Theory", Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (on-line edition, June 2009). Available at: http://plato.stanford.edu/entries/model-theory/

Hodges, W. 1997. *A Shorter Model Theory.* Cambridge: Cambridge University Press.

Howells, E. 1977. “Hume, Shaftsbury and the Peirce-James Controversy”, *History of Philosophy*, 15:449-62.

Hoyningen-Huene, N. Kirchlechner, B. and Beetz M. 2008. “GrAM: Reasoning with Grounded Action Models by Combining Knowledge Representation and Data Mining”, in Rome, E., Hertzberg, J., Dorffner, G. (eds.) 2008 *Towards Affordance-Based Robot Control* , Springer-Verlag, Berlin; pp. 47-62.

Hume, D. 1777. *An Enquiry concerning Human Understanding*. Nidditch, P. N. (ed.), 3rd. ed. (1975), Clarendon Press, Oxford.

Hume, D. 1994. *Political Essays,* (edited by Knud Haakonssen). Cambridge: Cambridge University Press.

Huntbach, M. M. & Ringwood, G. A. 1999. *Agent-Oriented Programming: From Prolog to Guarded Definite Clauses.* Berlin: Springer.

Husserl, E. 1969. *Formal and Transcendental Logic.* Translated by Dorion Cairns. The Hague: Martinus Nijhoff.

Husserl, E. 1970a. *Experience and Judgement: Investigations in a Geneaology of Logic.* Translated by Churchill, J. S. & Ameriks, K. Evanston: Northwestern University Press.

Husserl, E. 1970b. *Logical Investigations (Vol. I & II)* Translated by Findlay, J. N. London: Routledge.

Husserl, E. 1970c. *The Crisis of the European Sciences and Transcendental Phenomenology: And Introduction to Phenomenological Philosophy.* Translated by David Carr. Evanston: Northwestern University Press.

Innocenti Badano, B. M. 2008. *A Multi-agent Architecture with Distributed Coordination for an Autonomous Robot*, Doctoral Thesis, Girona: Universitat de Girona.

James, W. 1912. *Essays in Radical Empiricism,* New York and London.

James, W. 1981 [1890]. *The Principles of Psychology*, [vol. 1 & 2] Cambridge, Mass.: Harvard University Press.

Järvinen, A. 2008a. *Games without Frontiers: Theories and Methods for Game Studies and Design*, University of Tampere, Finland: Acta Electronica Universitatis Tamperensis; 701 [Ph.D. thesis]. Available at: http://acta.uta.fi/pdf/978-951-44-7252-7.pdf

Järvinen, A. 2008b. "All Rules Lead to Players", Doctoral defense speech by Aki Järvinen, presented at the University of Tampere, Finland, March 8th, 2008. Available at: http://www.gameswithoutfrontiers.net/g2/2008/03/25/all-rules-lead-to-players-defense-speech/#more-66

Jaspers, K. 1953. *The Origin and Goal of History,* London: Routledge and Kegan Paul.

Joas, H. 1985. *G. H. Mead: A Contemporary Re-examination of his Thought*, Cambridge, M.A.: MIT Press.

Joas, H. 1996. *The Creativity of Action,* Chicago: The University of Chicago Press.

Joas, H. 2000. *The Genesis of Values,* Cambridge, UK: Polity Press.

Joas, H. 2003. *War and Modernity,* Cambridge: Polity Press.

Joas, H. 2008. *Do we need Religion: On the experience of Self-Transcendence,* Boulder, CO: Paradigm Publishers.

Johnston, P. 1989. *Wittgenstein and Moral Philosophy,* London: Routledge.

Juul, J. 2005. *Half-Real: Video Games between Real Rules and Fictional Worlds,* Cambridge, Massachusetts: The MIT Press.

Keeton, H. C. S. 1984. *The Topology of Feeling: Extensive Connection in the Thought of Alfred North Whitehead, its Development and Implications* [Ph.D Thesis], Berkeley, California: Graduate Theological Union.

Kersten, F. 1969. "Franz Brentano and William James", *History of Philosophy*, 7:177-91.

Kitayama, S. Mesquita, B. & Karasawa, M. 2006. "Cultural Affordances and Emotional Experience: Socially Engaging and Disengaging Emotions in Japan and the United States", *Journal of Personality and Social Psychology*. 2006 Nov Vol 91(5) 890-903

Klein, A. 2008. "Future Combat – The complex crux of wireless warfare" *Washington Post*, Jan 24, 2008. Available at: http://www.washingtonpost.com/wp-dyn/content/article/2008/01/23/AR2008012303695.html

Kohout, L. J. and Pinkava, V. 1980. "The Algebraic Structure of the Spencer Brown and Varela Calculi", *International Journal of General Systems Theory*, 6: 155-171.

Kokar, M. Matheus, C. J. & Baclawski, K. 2009. "Ontology-based situation awareness", *Information Fusion*, 10(1): 83-98.

Kokar, M. & Ng, G. W. 2009. "High-level information fusion and situation awareness", *Information Fusion*, 10(1):2-5.

Kuhn, T. 1962. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.

Kultgen, J. H. 1973. "Intentionality and the Publicity of the Perceptual World", *Philosophy and Phenomenological Research*, Vol. 33, No. 4, pp. 503-513.

Kultgen, J. H. 1981. “Development and criticism of a behaviorist analysis of perception”, *Philosophy and Phenomenological Research*, Vol. 41, No. 4, pp. 465-486.

Lambert, D. A. 1996[1]. *Engineering Machines With Commonsense: Representation Revisited: An Essay in the Foundations of Artificial Intelligence* (unpublished Ph D thesis), Bedford Park, S.A.: The Flinders University of South Australia.

Lambert, D.A. 1999a. "Ubiquitous Command and Control", *Proceedings of the 1999 Information, Decision and Control Conference, Adelaide, Australia*, pp. 35–40, IEEE. Proceedings of the 10th International Command and Control Research and Technology Symposium (ICCRTS), MacLean, VA, June 13–16, 2005.

Lambert, D.A. 1999b. "Advisers with Attitude for Situation Awareness", *Proceedings of the 1999 Workshop on Defense Applications of Signal Processing*, pp. 113–118, edited by A. Lindsay, B. Moran, J. Schroeder, M. Smith and L. White, La Salle, Illinois, USA.

Lambert, D. A. 1999c. "Assessing Situations", *Proceedings of the 1999 Information, Decision and Control Conference, Adelaide, Australia*, pp. 503–508, IEEE.

Lambert, D. A. 2001a. "Opening up the future battlespaces through computer games", in *Journal of Aerospace and Defense Industry News*, April 13th, 2001.

Lambert, D. A. 2001b. "Situations for Situation Awareness". *Proceedings of the 4th International Conference on Information Fusion, Montreal, Canada.*

Lambert, D. A. 2001c. "An Exegesis of Data Fusion". [paper published in 2003; see under Lambert 2003]

Lambert, D. A. 2002a. "DSTO and Virtual Reality – Fact Sheet", DSTO publication.

Lambert, D. A. 2002b. “'Adelaide Command Centre to Lift Fog of War' – Transcript of ABC broadcast” 10/6/2002.

Lambert, Dale A. 2003a. "Grand Challenges of Information Fusion". In *Proceedings of the Sixth International Conference on Information Fusion*. Cairns, Australia. pp. 213–219.

Lambert, D.A. 2003b. "Automating Cognitive Routines", In *Proceedings of the Sixth International Conference on Information Fusion,* July 2003, Cairns, Queensland, Australia. pp. 986–993.

Lambert, D. A. 2003c. "An Exegesis of Data Fusion", in Edited L. Reznik and V. Kreinovich. *Soft Computing in Measurement and Information Acquisition*, [*Studies in Fuzziness and Soft Computing*], Berlin and Heidelberg: Springer Verlag; pp. 68-75.

[1] In Lambert’s publications the date cited for this work is 1995; however, in the copy obtained by the author from the Flinder’s University, S.A., the date that the thesis was submitted is listed as 31st of May 1996. Hence, I have used the latter date in my citations and in this bibliography.

Available at:
http://www.springer.com/cda/content/document/cda_downloaddocument/ch6.pdf?SGWID=0-0-45-373598-p2307611

Lambert, D.A 2006a. "Formal Theories for Semantic Fusion" *9th International Conference on Information Fusion, 2006. ICIF '06.* July 2006 Page(s):1 - 8

Lambert, D.A 2006b. "A Unification of Sensor and Higher Level Fusion", *International Conference on Information Fusion, 2006. ICIF '06.* 9th July 2006 Page(s):1 – 8

Lambert, D. A. 2007a. "Tradeoffs in the Design of Higher- Level Fusion Systems", *Information Fusion, 2007 [10th International Conference on] (9-12 July 2007*) pp. 1-4. Available at: http://www.isif.org/fusion07CD/Fusion07/Panels/Panel%201-05.pdf

Lambert, D. A. 2007b. "STDF Model Based Maritime Situation Assessments", *Information Fusion, 2007 [10th International Conference on] (9-12 July 2007*) pp. 1-8. Available at: http://www.isif.org/fusion07CD/Fusion07/pdfs/Fusion2007_1269.pdf

Lambert, D.A. 2009a. "A blueprint for higher-level fusion systems", *Information Fusion*, 10, pp.6-24.

Lambert, D.A. 2009b. "Changes since *EMWC*", unpublished document, July 2009, pp.1-14.

Lambert D. A. & Nowak, C. 2006. "Mephisto I: Towards a Formal Theory", Paper presented at Australasian Ontology Workshop, Hobart, 2006. Available at: http://crpit.com/confpapers/CRPITV72Lambert.pdf

Lambert, D.A., & Nowak, C., 2008. "The Mephisto Conceptual Framework", *Technical Report, DSTO-TR-2162*, Defence Science & Technology Organisation, Australia. URL = http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/10021/1/DSTO-TR-2162%20PR.pdf.

Lambert, D. A. & Relbe, M. G. 1998. "Reasoning with tolerance", *Second International Conference on Knowledge-Based Intelligent Electronic Systems, 21-13 April, 1998, Adelaide, Australia*. Editors, Jain, L. & Jain, R. IEEE, pp.418-427.

Lambert, D. A., Saulwick, A., Nowak, C., Oxenhan, M., & O'Dea, D., 2008 *An Overview of Conceptual Frameworks: Technical Report*. Edinburgh, South Australia: Defence Science and Technology Organisation. Available at: http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/10022/1/DSTO-TR-2163%20PR.pdf

Lambert, D.A, & Scholz, J.B. 2005. A Dialectic for Network Centric Warfare, *Proceedings of the International Command and Control Research and Technology Symposium (ICCRTS)*. Available at: www.dsto.defence.gov.au/attachments/ A_Dialectic_for_NCW-final_public_release.pdf

Lambert, D. A. & Scholz, J. 2007. "Ubiquitous command and control", *Intelligent Decision Technologies,* Volume 1, Number 3, pp.157 – 173.

Lambert, D. A., Wark, S., Blackman, C., Bossé, E., Pigeon, L., & Hinman, M. 2009. "A Coalition Approach to Higher-Level Fusion", paper presented at 12th International Conference on Information Fusion (Fusion 2009) on 6-9-July-2009 Seattle, Washington. Available at: http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA506230

Lango, J. W. 1994. "Review: Granville Henry, Forms of Concrescence", *Transactions of the Charles S. Peirce Society,* 30, no. 3: pp.727-38.

Lango, J. W. 2000. "Whitehead's Category of Nexus of Actual Entities", *Process Studies*, 29/1, pp.16 -42. Available at - http://www.religion-online.org/showarticle.asp?title=3006

Lansing, J. S. 2002. "*Artificial Societies* and the Social Sciences", *Artificial Life,* Vol. 8, pp.279-292.

Laszlo, E. 1966. *Beyond Scepticism and Realism. A Constructive Exploration of Husserlian and Whiteheadian Methods of Inquiry.* The Hague: Martinus Nijhoff.

Latour, B. 1999. *Pandora's Hope: Essays on the Reality of Science Studies,* Cambridge, M.A.: Harvard University Press.

Latour, B. 2005. "What is the Style of Matters of concern? Two Lectures on Empirical Philosophy", available at: http://www.bruno-latour.fr/articles/index.html

Lederman, R. and Johnston, R. B. 2008. "Manual Collaboration Systems: Decision Support or Support for Situated Choices", In *Collaborative Decision Making : Perspectives and Challenges*, Edited by Zarate, P., Belaud., J.P., Camilleri., G, Ravat, F., Amsterdam: IOS Press, pp. 333-343.

Lewis, S. C. 2004. *Computational Models of Emotion and Affect*, PhD thesis, Computer Science, University of Hull, July 2004.

Ley, R. D. and Johnson, L. E. 1990. "Selecting Social Goals: Alternative Concepts of Rationality", *American Journal of Economics and Sociology,* Vol. 49, No. 4, pp.469-481.

Lingard, D.M., & Lambert, D.A.. 2008. "Evaluation of the effectiveness of machine-based situation assessment - preliminary work", Scientific Commons: http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/9656/1/DSTO-TN-0836%20PR.pdf

Little, E. G., & Rogova, G. L. 2009. "Designing ontologies for higher level fusion" *Information Fusion* 10(1):70-82.

Liu, K. 2000. *Semiotics in information systems engineering*. Cambridge Univerisity

Press, Cambridge, England.

Llinas, J., Bowman, C., Rogova, G., Steinberg, A. and White, F. 2004. "Revisiting the JDL data fusion model II", in P. Svensson and J. Schubert (Eds.), *Proceedings of the Seventh International Conference on Information Fusion (FUSION 2004).* Available at: http://www.fusion2004.foi.se/papers/IF04-1218.pdf

Lorenz, K. 1982. "Kant's doctrine of the a priori in the light of contemporary biology", in Plotkin, H. C. (ed.) *Learning, Development and Culture.* Chichester: John Wiley, pp.121–143.

Lowe, V. 1941. "William James and Whitehead's Doctrine of Prehensions", *The Journal of Philosophy*, 38:113-126.

Lowe, V. 1951. "The development of Whitehead's Philosophy", in Schilpp (1951) pp.15-124.

Lowe, V. 1975. "A. N. Whitehead and his Mathematical Goals: a Letter of 1912", *Annals of Science*, 32:85-101.

Lowe, V. 1985. *Alfred North Whitehead: The Man and His Work,* Vol. 1. 1861-1910. Baltimore: John Hopkins University Press.

Lowe, V., Hartshorne, C., and Johnson. A. H. 1950. *Whitehead and the Modern World; Science, Metaphysics, and Civilization: Three Essays on the Thought of Alfred North Whitehead.* Boston: Beacon Press.

Loyal, S. & Quilley, S. (eds.) 2004. *The Sociology of Norbert Elias*, Cambridge: Cambridge University Press.

Lubbe, H. 1978. "Positivism and Phenomenology", in Luckmann, T. (ed.) 1978 *Phenomenology and Sociology*, Harmondsworth: Penguin, pp. 90-118.

Lucas, G. R., Jr. 1989. *The Rehabilitation of Whitehead: An Analytic and Historical Assessment of Process Philosophy*. Albany: State University of New York.

Luck, M. McBurney, P. and Priest, C. 2004. "A Manifesto for Agent Technology: Towards Next Generation Computing", *Autonomous Agents and Multi-Agent Systems,* vol. 9, no. 3, pp. 203-252.

Luck, M, McBurney, P. Shehory, O. and Willmott, S. 2005. *Agent Technology: Computing as Interaction. A Roadmap for Agent Based Computing,* AgentLink III. Available at: http://centria.di.fct.unl.pt/events/easss08/proceedings/01-EASSS08Intro1.pdf.

Luck, M., McBurney, P. and Gonzalez-Palacios, J. 2006. *Agent-Based Computing and Programming of Agent Systems*, in Bordini, R., Dastani, M., Dix, J. and El Fallah Segrouchni, A., Eds. *Programming Multi-Agent Systems, Lecture Notes in Artificial Intelligence*, 3862. Springer. Available at: http://www.ecs.soton.ac.uk/~mml/papers/promas06.pdf

Luhmann, N. 1986. *Love as Passion: The Codification of Intimacy*. Cambridge: Polity.

Luhmann, N. 1995 *Social Systems*. Stanford, California: Stanford University Press.

Luhmann, N. 1996. “The Sociology of the Moral and Ethics”, *International Sociology*, Vol. 11, No. 1, pp. 27-36.

Luhmann, N. 2000. *Art as a Social System,* Stanford, California: Stanford University Press.

Mach, E. 1905. *Erkenntnis und Irrtum. Skizzen zur Psychologie der Forschung.* Leipzig: J. A. Barth; (translated by P. McCormack & P. Foulkes, 1976, as *Knowledge and error: Sketches on the psychology of enquiry,* Dordrecht: Reidel).

Macy, M.W., Willer, R., 2002. “From Factors to Actors: Computational Sociology and Agent Based Modeling", *Annual Reviews Sociology,* 28: 143-166.

Malsch, T. 2001. “Naming the Unnameable: Socionics or the Sociological Turn of/to Distributed Artificial Intelligence”, in *Autonomous Agents and Multi-Agent Systems 3/2001*, pp.155-187. Available at: http://www.tu-harburg.de/tbg/Deutsch/Mitarbeiterinnen/Thomas/thomas-down.htm

Malsch, T. and Schleider, C. 2004. “Communication without Agents? From agent-oriented to communication-oriented modeling”. In: Lindemann, G.; Moldt, D.; Paolucci, M. (Hg): *Regulated Agent-Based Social Systems. First International Workshop, RASTA 2002*, Bologna, Italy, July 2002. Revised Selected and Invited Papers. Berlin, Heidelberg, New York: Springer-Verlag, pp. 113-133.

Malsch. T., and Schulz-Schaeffer, I., 2007. "Socionics: Sociological Concepts for Social Systems of Artificial (and Human) Agents", *Journal of Artificial Societies and Social Simulation* vol. 10, no. 1. Available at: http://jasss.soc.surrey.ac.uk/10/1/11.html

Malsch, T. & Weiss, G. 2000. “Conflicts in Social Theory and Multiagent Systems: On importing Sociological Insights into Distributed AI”, in Tessier, C. Chaudron, L. & Müller, H. J. (eds.) *Conflicting Agents: Conflict Management in multi-agent systems.* Dordrecht: Kluwer.

Manninen, T. 2004. *Rich interaction model for game and virtual environment design,* Ph. D Thesis, Department of Information Processing Science, University of Oulu, P.O.Box 3000, FIN-90014, University of Oulu, Finland.

Manzo, G. 2007. "Mechanisms and Relations", *Sociologica* No. 2., (Italian Journal of Sociology on line). Article available at: http://www.sociologica.mulino.it/journal/article/index/Article/Journal:ARTICLE:97

Marx, K. 1990 .*Capital, Volume I*. Trans. Ben Fowkes. London: Penguin.

Mateas, M. 2002. *Interactive Drama, Art and Artificial Intelligence,* Ph. D thesis, Carnegie Mellon University.

McCarthy, J. 1987. "Generality in artificial intelligence", *Communications of the ACM* 30(12): 1030-1035. Available at: http://www-formal.stanford.edu/jmc/index.html

McCarthy, J. 1989. "Artificial intelligence, logic, and formalizing common sense". In *Philosophical Logic and Artificial Intelligence*, R. H. Thomason, ed., 161-190. Dordrecht, The Netherlands: Kluwer. Available at: http://www-formal.stanford.edu/jmc/index.html

McCarthy, J. and Hayes, P.J., 1969. "Some philosophical problems from the standpoint of Artificial Intelligence", in Meltzer, B. and Michie, D. (eds.) 1969 *Knowledge Representation 4*, Edinburgh: Edinburgh University Press, pp.463-502.

McClelland, K. A. and Fararo, T. J. (eds.) 2006. *Purpose, Meaning and Action: Control Systems Theories in Sociology,* New York: Palgrave MacMillan.

McCulloch, W.S. & Pitts, W. 1943. "A logical calculus of the ideas immanent in nervous activity", *Bull. Math. Biophysics*, Vol. 5 pp.115-33.

McEvoy, J. G. 1987. "Causes and Laws, Power and Principles: The Metaphysical Foundations of Priestley's Concept of Phlogiston", in Anderson, R. and Lawrence, C. (eds.) 1987, *Science, Medicine and Dissent: Joseph Priestley (1733-1804),* London: Wellcome Trust.

McHenry, L. B. 1992. *Whitehead and Bradley: A Comparative Analysis*, Albany, NY: State University of New York.

McKenna, T., Moon, T., Davis R., & Warne L., 2006. "Science & Technology for Australian Network-Centric Warfare: Function, Form and Fit", *Australian Defence Force Journal,* No.170, pp.62-76.

Medler, D. A. 1998. "A brief history of connectionism". *Neural Computing Surveys*, *1*(2), pp.18-72

Melle, U. 1998. "Responsibility and the Crisis of Technological Civilization: A Husserlian Meditation on Hans Jonas", *Human Studies*, 21:329-345.

Meneguzzi, F. R. and Luck, M., 2007. "Composing high-level plans for declarative agent programming", in *Proceedings of the Fifth International Workshop on Declarative Agent Languages and Technologies*, Hawaii, USA.

Meneguzzi, F. R., Zorzo, A. F., Da Costa Mora, M., and Luck, M., 2007. "Incorporating Planning into BDI Systems"**,** *Scalable Computing: Practice and Experience*, **8**(1), 15-28.

Mennell, S. 1992. *Norbert Elias: An Introduction*. Oxford: Blackwell.

Mensch, J. R. 1998. *Intersubjectivity and Transcendental Idealism,* Albany, N.Y.: State University of New York Press.

Mensch, J. R. 1997. "Freedom and Selfhood", *Husserl Studies*, 14: 41-59.

Mensch, J. R. 1998. "Instincts – A Husserlian Account", *Husserl Studies*, 14:219-237.

Meschini, D., Markku, L., and Piilonen, J. 2004. "Geometry, pregeometry and beyond", arXiv e-print (arXiv:gr-qc/0411053), pp. 1-43.

Moore, J. 1985. "Evangelicals and Evolution: Henry Drummond, Herbert Spencer, and the Naturalization of the Spiritual World", *Scottish Journal of Theology*, 38:383-417.

Moran, D. 2000. *Introduction to Phenomenology,* London: Routledge.

Moretti, S. 2002. "Computer Simulation in Sociology: What Contribution?" *Social Science Computer Review,* Vol. 20, No. 1, pp. 43-57.

Morie, J.F., Williams, J. Dozois A. and Luigi. D-P 2005. "The Fidelity of Feel: Emotional Affordance in Virtual Environments", in *Proceedings of the 11th International Conference on Human-Computer Interaction,* Las Vegas, NV.

Moss, S., Sawyer, R. K., Conte, R., and Edmonds, B. 2001. Sociology and social theory in agent based social simulation: A symposium. *Computational and Mathematical Organization Theory*, 7(3), 183-205.

Müller, J.P. 1997. "A cooperation model for autonomous agents" in Müller, J.P. Wooldridge, M.J. and Jennings, N.R. (eds.), 1997 *Intelligent Agents III - Agent Theories, Architectures, and Languages*, (*LNCS, Vol. 1193*), Berlin: Springer-Verlag, pp.245-260.

Müller, J. P. and Pischel, M. 1994. "An architecture for dynamically interacting agents" *International Journal of Intelligent and Cooperative Information systems*, 3(1): 25-46.

Muller, J. P., Pishel, M. & Thiel, M. 1995. "Modelling Reactive Behaviour in Vertically Layered Agent Architectures", In Wooldridge, M. & Jennings, N. (eds.) 1995 *Intelligent Agents*, Heidelberg: Springer Verlag, 261-276.

Müller, H. J. Malsch, T. and Schulz-Schaeffer, I. 1998. "Socionics: Introduction and Potential", *Journal of Artificial Societies and Social Simulation* vol. 1, no. 3. Available at: http://www.soc.surrey.ac.uk/JASSS/1/3/5.html pp. 1-6.

Murray, J. H. 2005. "The last word on Ludology v Narratology in Game Studies" (delivered as preface to keynote talk at *DiGRA 2005*, Vancouver, Canada, June 17, 2005). Available at: http://www.lcc.gatech.edu/~murray/digra05/lastword.pdf

Nareyek, A. 2004. "AI in Computer Games", *Queue,* (Feb. 2004), pp. 58-65.

Newell, A. and Simon, H. A. 1976. "Computer science as empirical enquiry". *Communications of the ACM*, 19:113-126.

Newell, A. 1980. “Physical Symbol Systems”, *Cognitive Science*, Vol. 4, pp.135-183.

Newell, A. 1982. "The Knowledge Level", *Artificial Intelligence*, Vol. 18, pp. 87-127.

Nickles, M. Rovatsos, M. & Weiß. G. 2004. "Empirical-rational semantics of agent communication". *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems* (AAMAS, pp. 94-101). Available at: http://www7.in.tum.de/~weissg/weissg-pub.html

Nickles, M. Rovatsos, M. Brauer, W. & Weiß. G. 2004. “Communication Systems: A Unified Model of Socially Intelligent Systems”. In Klaus Fischer, Michael Florian, Thomas Malsch (editors): Socionics: Its Contributions to the Scalability of Complex Social Systems. Springer LNAI, 2004. http://www7.in.tum.de/~weissg/weissg-pub.html

Nickles M. & Weiß G. 2005. "Multiagent systems without agents: mirror-holons for the compilation and enactment of communication structures", in K. Fischer, M. Florian, & T. Masch (Eds.), *Socionics*. Lecture Notes in Computer Science, Vol. 3413. Springer-Verlag. (pp. 289-313).

Nidditch, P. H. 1962. *The Development of Mathematical Logic,* London: Routledge and Kegan Paul.

Nielsen, K., Appel J., and Demazeau, Y., 2006. "Applying AI to Cooperating Agricultural Robots". AIAI 2006: 262-270. Available at: http://www.mycasa.dk/files/CCI53.pdf.

Nobo, J. L. 1979. “Transition in Whitehead: A Creative Process Distinct from Concrescence”, *International Philosophical Quarterly,* 19: 265-283.

Nobo, J. L. 1986. *Whitehead’s Metaphysics of Extension and Solidarity*, Albany: State University of New York Press.

Nobo, J. L. 1998. “The Approach to Whitehead: Traditional? Genetic? Or Systematic?” *Process Studies* 27/1-2: 48-63.

Noë, A. 2004. *Action in perception*, Cambridge, MA: MIT Press

Norman, D. A., 1988. *The Psychology of Everyday Things*. New York: Basic Books.

Norman, D. A., 1999, "Affordances, Conventions and Design". *Interactions* 6(3):38-43, May 1999, ACM Press. http://www.jnd.org/dn.mss/affordance_conv.html

Nowak, C. 2001. “A note on process ontology for agents”, Workshop on Ontologies in Agent Systems, 5th International Conference on Autonomous Agents, Montreal. Available at:

http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=E886D656467E35C2668FDBD90865CF76?doi=10.1.1.21.5320&rep=rep1&type=pdf

Nowak, C. 2003. "On ontologies for high-level information fusion". In *Proceedings of the Sixth International Conference on Information Fusion*. IEEE Cairns, Australia.

Nowak, C. & Lambert, D. 2005. "The Semantic Challenge for Situation Assessments", *Eighth International Conference on Information Fusion* (FUSION), IEEE.

Noy N. and McGuinness DL. 2001. *Ontology Development 101: A Guide to Creating Your First Ontology.* Stanford University, Technical report SMI-2001-0880.

Nunes H. A. 2001. *Investigation of motivation in agents using the simulation of 5-aside football,* MA thesis, Computer Science, University of Hull, November 2001.

Oakley, J. 1992. *Morality and the Emotions*, London: Routledge.

Ostrow, J. 1987. "Habit and inhabitance: an analysis of experience in the classroom", *Human Studies*, 10: 213-224.

Ostrow, J. 1990. *Social Sensitivity: A Study of Habit and Experience,* Albany: SUNY Press.

Oxenham, M. G., Percival, J., Price, R., and Lambert, D., 2003. "Information Fusion And Extraction Priorities For Australia's Information Capability", in *Proceedings of the Nato Conference on Military Data and Information Fusion,* Prague, Czech Republic, October 20-22, 2003.

Paci, E. 1964. "Whitehead e Husserl", *Aut Aut*, Vol. 84, pp. 7-18. English translation (by Bianchi, L. M. 1998), available at:
http://www.yorku.ca/lbianchi/paci/whitehead_and_husserl.html

Pagels, H. 1988. *The Dreams of Reason: The Computer and the Rise of the Sciences of Coomplexity,* New York: Simon & Schuster.

Parsons, T. 1937. *The Structure of Social Action* (Vol. 1 & 2), New York: Free Press.

Parsons, T. 1951. *The Social System*. New York: Free Press.

Parsons, T. 1954. *Essays in Sociological Theory*, New York: Free Press.

Parsons, T. 1959. "An approach to Psychological Theory in terms of the Theory of Action", in Koch, S. (ed.) 1959 *Psychology: A Study of a Science,* Vol. 3. New York: McGraw-Hill, pp. *612-709.*

Parsons, T. 1974. "Comment on 'Current folklore in the criticisms of Parsonian Action Theory' [Turner & Beeghely, 1974]", *Sociological Inquiry*, 44/1:55-58.

Parsons, T. 1967. *Sociological Theory and Modern Society*, New York: Free Press.

Parsons, T. 1977. *The Evolution of Societies,* Englewood Cliffs, N.J.: Prentice-Hall.

Parsons, T. 1978. *Action Theory and the Human Condition*, New York: Free Press.

Parsons, T. 1979. "Religious and Economic Symbolism in the Western World," *Sociological Inquiry,* Vol. 49, pp.1-48.

Parsons, T. 1981. "Letter from Talcott Parsons", in *Sociological Inquiry*, vol.51, no.1, pp.35-6.

Parsons, T. and Bales, R. F. 1955. *Family, Socialization and Interaction Process*. New York: Free Press.

Pater, R., Hsieh M., & Silverman, B., 2007. "CTDB 3D Viewer Editor". Available at: http://www.seas.upenn.edu/~cse400/CSE401_2007/PaterHsieh/PaterHsieh.pdf.

Pawlaszczyk, D.; Dietrich, A. J.; Timm, I.; Otto, S.; Kirn, St. 2004. "Ontologies Supporting Cooperations in Mass Customization – A Pragmatic Approach", in: Piotrowski, M. (ed.): *Proceedings of the International Conference on Mass Customization and Personalization – Theory and Practice in Central Europe*, UITM, Rzeszów. Available at: http://www.tu-ilmenau.de/fakww/fileadmin/template/fakww/Institute_und_Fachgebiete/Wirtschaftsinformatik/Wirtschaftsinformatik_1/Bilder/Mitarbeiter/Dirk/ICMC_MC-Ontology_FINAL_040412.pdf.

Peel, J. D. Y. 1971. *Herbert Spencer: the Evolution of a Sociologist*, London: Heinemann.

Perugini, D. Lambert, D. Sterling, L. and Pearce, P. 2002. "Agents for Military Logistic Planning," paper presented at a workshop *"Agent Technology in Logistics," 15th European Conference on Artificial Intelligence (ECAI-2202),* Lyon, France, 2002.

Perugini, D.; Lambert, D.; Sterling, L.; Pearce, A. 2003. "Distributed information fusion agents", *Proceedings of the Sixth International Conference of Information Fusion, 2003* Volume 1, Page(s):86 – 93.

Perugini, D.; Lambert, D.; Sterling, L. & Pearce, A.2003. "A distributed agent approach to global transportation scheduling", *IAT 2003. IEEE/WIC International Conference on Intelligent Agent Technology,13-16 Oct. 2003*, pp.18 – 24.

Perugini, D.; Lambert, D.; Sterling, L.& Pearce, A. 2004. "Agent-Based Global Transportation Scheduling in Military Logistics", AAMAS'04, July 19-23, 2004, New York, USA. ACM.

Perugini, D.; Wark, S.; Lambert, D.; Sterling, L. & Pearce, A. 2004. "Agents in Logistics Planning – Experiences with the Coalition Agents Experiment Project", *Agents at work: Deployed applications of autonomous agents and multi-agent* *systems*, a workshop at the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2003).

Perugini, D. 2006, *Agents for Logistics: A Provisional Agreement Approach,* PhD Thesis, Uni Melb. Available at: http://eprints.infodiv.unimelb.edu.au/archive/00002517/01/DonPeruginiThesisFinal.pdf

Porter, D. H. 1975. "History as Process." *History and Theory* 14: 297-313.

Porter, D. H. 1979. "Explaining the Historical Process." *Process Studies* 9 (Fall-Winter): 73-93. Available at http://www.religion-online.org/showarticle.asp?title=2480

Powis, L.; Lambert, D.A.; Sarunic, P.W.; Jakobsson, A.; Stone, M. 1992. "Adaptive Radar Control of multifunction radars using artificialintelligence" *Radar 92. International Conferenc*, (12-13 Oct 1992). pp.426 – 429.

Price, L. 1954. *Dialogues of Alfred North Whitehead*, London: Max Reinhardt.

Primeaux, D. 2000. "Trust-based Learning of Data Characteristics by an Actual Entity." *Proceedings of the ACIS 1st International Conference on Software Engineering Applied to Networking & Parallel/Distributed Computing* (SNPD) May 2000, 239-244.

Primaux, D. & Saunders, B. 2006. "Finding Global and Local Maxima in 3D Space Using Swarm-like Behavior in a Colony of Prehending Entities", *2006 IEEE Swarm Intelligence Symposium.*

Putnam, H. 1973. "Meaning and Reference," *Journal of Philosophy* 70, pp.699-711.

Quine, W. V. O. 1960. *Word and Object*, Cambridge, Mass.: The MIT Press.

Raim, J. 2008. "Finite State Machines in Games", ppt presentation, Lehigh University, CSE Dept. Available at: www.cse.lehigh.edu/~munoz/CSE497/classes/FSM_In_Games.ppt

Rao, A. S. and Georgeff M. P., 1991. "Modeling Rational Agents within a BDI Architecture", *Proceedings of Knowledge Representation and Reasoning* (KR&R-91), San Mateo, CA.: Morgan Kaufmann.

Rasch, W. 2000. *Niklas Luhmann's Modernity: the paradoxes of differentiation*, Stanford, CA: Stanford University Press.

Raubal, M. 2001 "Ontology and Epistemology for agent-based way-finding and simulation", *Int. J. Geographical Information Science*, 15/7: pp.653-655.

Raubal, M., Miller, H., and Bridwell, S. 2004. "User-Centred Time Geography For Location-Based Services". *Geografiska Annaler B* 86(4), 245–265.

Raubal, M. and Moratz, R. 2008. “A Functional Model for Affordance-Based Agents”, in Rome, E., Hertzberg, J., Dorffner, G. (eds.) 2008 *Towards Affordance-Based Robot Control ,* Springer-Verlag, Berlin; pp.91-105.

Reid, D.J., Giffin, R.E., 2003. “A Woven Web of Guesses, Canto Three: Network Centric Warfare and the Virtuous Revolution”, *Proceedings of the 8th ICCRTS*, Washington DC, 17–19 June.

Reilly, W. S. N. 1996. *Believable Social and Emotional Agents*, Ph. D thesis, Pittsburgh, PA: Carnegie Mellon University.

Rescher, N. 1991. “The Promise of Process Philosophy”, in Rescher, N. 1991 *Baffling Phenomena and Other Studies in the Philosophy of Knowledge and Valuation,* Savage, Maryland: Rowman and Littlefield, pp.75-96.

Rescher, N.: 1996. *Process Metaphysics – An Introduction to Process Philosophy*, State University of New York Press, Albany, NY.

Rescher N. 2000. *Process Philosophy- A Survey of Basic Issues*, Pittsburgh: University of Pittsburgh Press.

Resnyansky, L. 2007. “Integration of social sciences in modelling: an interactionist approach to research practice”, *First International Conference on Computational Cultural Dynamics (ICCCD 2007), 27-28 August 2007.* Available at: http://www.umiacs.umd.edu/conferences/icccd2007/lucy-r.pdf

Resnyansky, L. 2007. “Integration of social sciences in terrorism modelling: issues, problems and recommendations”, *Technical Report, Defence Science and Technology Organization, Australia (DSTO-TR-1955).* Available at: http://hdl.handle.net/1947/4684

Riccardo Boero 2003. “Some Methodological Issues of Agent Based Models in Social Sciences.” Unpublished paper, available at: http://www.unisi.it/santachiara/aree/conf_phd_econ2003/conference_siena/papers/boero.pdf.

Richardson, D., Deckro, R. und Wiley, V., 2004. “Modeling and Analysis of Post-Conflict Reconstruction”, in: *Journal of Defense Modeling and Simulation*, Volume 1, Number 4, S. 201-213.

Richer, W. 2006 "NCW Risk Assessment – Towards a Compliance Policy and Process", *Proceedings of the 11 th. ICCRTS, Command and Control in the Networked Era*; Cambridge 2006. Available at http://www.dodccrp.org/events/11th_ICCRTS/html/papers/112.pdf

Ringel, C. M. 2008. "Extension in PR Part IV", in Weber, M. & Desmond, W. 2008 *Handbook of Whiteheadian Process Thought*, [Vol. 2], Frankfurt: Ontos Verlag, pp.131-148; 149-156. Available at: http://www.mathematik.uni-bielefeld.de/~ringel/opus/extension.pdf

Robertson, R. & Turner B. S. (eds.) 1991. *Talcott Parsons: Theorist of Modernity*, London: SAGE

Rockwell, W. T. 2001. "Experience and Sensation: Sellars and Dewey on the Non-cognitive aspects of Mental Life", *Education and Culture: the Journal of the John Dewey Society*, (Winter, 2001). http://users.california.com/~mcmf/sellarsdewey.html

Robert, A. 2000. *EMuds: Adaptation in Text-Based Virtual Worlds,* Ph. D thesis, Fribourg, Switzerland: University of Fribourg.

Robertson, D. S. 1998. *The New Renaissance: Computers and the Next Level of Civilization,* Oxford: Oxford University Press.

Rocher, G. 1974. *Talcott Parsons and American Sociology*, London: Nelson.

Rogers, B. L. 1991. *A new associative conditioning element for consumatory classical conditioning*. Unpublished PhD dissertation, Swinburne University of Technology, Melbourne.

Rogers, B. L. 1993a. "The associative conditioning element", in P. Leong & M. Jabri (eds.) *Proceedings of the Fourth Australian Conference on Neural Networks*, Sydney: Sydney University Electrical Engineering, pp.45-8.

Rogers, B. L. 1993b. "Towards a Computational Model of Whitehead's Metaphysics of Creativity", *Technical Report: SUT-CS-19/93*, published by Department of Computer Science, Swinburne University of Technology, pp.1-45.

Rome, E., Hertzberg, J., Dorffner, G. (eds.) 2008. *Towards Affordance-Based Robot Control ,* Springer-Verlag, Berlin.

Rovatti, P. 1968. "Whitehead e Husserl: una relazione", *Man and World*, Vol. 1, No. 4. pp.587-603.

Rovatsos, M. Weiß, G. & Wolf, M.. 2002. "An Approach to the Analysis and Design of Multiagent Systems based on Interaction Frames", in Gini, M., Ishida, T. and Lewis Johnson, W. (eds.), *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-02),* Bologna, Italy, July 15-19, ACM Press. Available at: http://homepages.inf.ed.ac.uk/mrovatso/papers/rovatsosetal-aamas2002.pdf

Rovatsos, M. Nickles, M. & Weiß, G. 2003. "Interaction is meaning: A new model for communication in open systems". *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems* (AAMAS, pp. 536-543). 2003. http://www7.in.tum.de/~weissg/weissg-pub.html

Rovatsos, M. 2004. *Computational Interaction Frames*. PhD thesis, Department of Informatics, Technical University of Munich.

Russell, B. 1921. *The Analysis of Mind*, London: George Allen and Unwin.

Russell, B. 1997. *The Collected Papers of Bertrand Russell, Volume 11: Last Philosophical Testament, 1943-68*. (Edited by John G. Slater with the assistance of Peter Kollner), London and New York: Routledge.

Russell, K. 2004 “Artifacts and affordance: the surface of meaning”, *Working Papers in Art and Design* **3.** Available at: http://sitem.herts.ac.uk/artdes_research/papers/wpades/vol3/krfull.html

Russell, S. J. and Norvig, P. 2003. *Artificial Intelligence: A Modern Approach (Second Edition),* New Jersey: Prentice Hall.

Russell, W. A. 1966. **“**A note on Lotze's teaching of psychology, 1842 – 1881”, *Journal of the History of the Behavioral Sciences,* Volume 2, Issue 1, pp: 74-75

Rutten, M. G. & Lambert, D. A. 2003. "Automated Adaptive Situation Assessment", in Reznik, L. and Kreinovich, V. (eds.) *Soft Computing in Measurement and Information Acquisition*, [*Studies in Fuzziness and Soft Computing*], Berlin and Heidelberg: Springer Verlag; pp.90-104.

Ryder, M. & Wilson, B.. 1996. “Affordances and Constraints of the Internet for Learning and Instruction”, Conference paper, *Association for Educational Communications Technology, Indianapolis, Feb. 14-18, 1996*. Available at: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/a1/22.pdf

Sallach, D. L. 2003. “Social Theory and Agent Architectures: Perspective Issues in Rapid-Discovery Social Science”, *Social Science Computer Review*, Vol. 21, No. 2, pp.179-195.

Salmon, W. C. 1984. *Scientific explanation and the causal structure of the world.* Princeton, NJ: Princeton University Press.

Sapaty, P., Morozov, A., Finkelstein, R., Sugisaka, M., & Lambert, D. 2008. “A new concept of fl exible organization for distributed robotized systems”, *Artificial Life and Robotics* (2008) 12:88–96. http://www.springerlink.com/content/k65v067l43601ug0/

Saunders, B. 2008. *Observational Intelligence. An Overview of Computational Actual Entities and Their Use as Agents of Artificial Intelligence*, VDM Verlag.

Sawyer, R. K. 2001. "Emergence in sociology: Contemporary philosophy of mind and some implications for sociological theory". *American Journal of Sociology*, 107(3), 551-585.

Sawyer, R. K. 2002a. “Unresolved Tensions in Sociocultural Theory: Analogies with Contemporary Sociological Debates”, *Culture and Psychology*, Vol. 8, No. 3, pp.283-305.

Sawyer, R. K. 2002b. "Durkheim’s dilemma: Toward a sociology of emergence". *Sociological Theory*, 20(2), 227-247.

Sawyer, R. K. 2002c. "Nonreductive individualism, Part I: Supervenience and wild disjunction". *Philosophy of the Social Sciences*, 32(4), 537-559.

Sawyer, R. K. 2003a. "Artificial societies: Multi agent systems and the micro-macro link in sociological theory". *Sociological Methods and Research*, 31(3), 325-363.

Sawyer, R. K. 2003b. "Nonreductive individualism, Part II: Social causation". *Philosophy of the Social Sciences*, 33(2), 203-224.

Sawyer, R. K. 2004a. "The mechanisms of emergence". *Philosophy of the Social Sciences*, 34(2), 260-282. Special issue: "Systems and mechanisms: A symposium on Mario Bunge's philosophy of social science." Guest Editor: Andreas Pickel.

Sawyer, R. K. 2004b. "Social explanation and computational simulation". *Philosophical Explorations*, 7(3), 219-231. Special issue: "The social explanation of action." Guest editors: Stefan Cuypers and Robrecht Vanderbeeken.

Sawyer, R. K. 2005. *Social Emergence: Societies as Complex Systems,* Cambridge: Cambridge University Press.

Sawyer, R. K. and Berson, S. 2004. "Study group discourse: How external representations affect collaborative conversation". *Linguistics and Education*, 15, 387-412.

von Scheve, C., Moldt, D. 2004. "Emotion: Theoretical Investigations and Implications for Artificial Social Aggregates", in Lindeman, G. Moldt D. and Paolucci P. (eds.) *Regulated Agent-Based Social Systems.* Berlin: Springer, pp. 189-209.

Schillo, M. Fischer, K. and Klein, C. T. 2001. "The Micro-Macro Link in DAI and Sociology", in Moss, S. Fischer, K. and Klein, C. (eds.) 2001 *Multi-Agent Based Simulation: Second International Workshop on Multi-Agent based Simulation, Boston, MA, USA, July 2000. Lecture Notes in AI,* Berlin: Springer-Verlag, pp. 133-148. Available at: http://www.virtosphere.de/data/publications/articles/schillo+.lnai1979.pdf

Schilpp, P. A. (ed) 1951. *The Philosophy of A. N. Whitehead*, (2nd edn.) La Salle, Illinois: Open Court.

Schmidt, B. 2000. *The Modelling of Human Behaviour*, Ghent: SCS-Europe BVBA.

Schmidt, B. 2005a. "Human Factors in Complex Systems: The Modelling of Human Behaviour", *Proceedings 19th European Conference on Modelling and Simulation, ECMS Europe*.

Schmidt, B. 2005b. "Human Factors in Simulation Models"; http://www.scs-europe.net/services/ecms2005/pdf/plen-02.pdf.

Schmidt, B., & Schneider, B. 2004. “Agent-based modelling of human acting, deciding and behaviour - the reference model PECS”, *Proceedings 18th European Simulation Multiconference, SCS Europe, 2004*.

Schoen, P. 2007. "System of Systems Common Operating Environment (SOSCOE): Support to Net Centricity", ppt. presentation by Paul Schoen, Director SOSCOE FCS LSI , Boeing. Available at: http://www.dtic.mil/ndia/2007netcentric/Schoen_NDIAPublicReleaseversion0short.pdf.

Scholz, J., Hengst, B., Calbert, G., Antoniades, A., Smet, P., Marsh, L., Kwok, H., and Gossink, D., 2005. “Machine Learning for Adversarial Agent Microworlds,” in Zerger, A. and Argent, R.M. (eds), *MODSIM 2005 International Congress on Modelling and Simulation*, Modelling and Simulation Society of Australia and New Zealand, December, pp 2195–2201, ISBN: 0-9758400-2-9: www.mssanz.org.au/modsim05/proceedings/papers/scholz.pdf

Scholz, J. 2006. "Agreement Technologies"', Ppt slide presentation at 2006 seminar on Innovation-Focused ICT Development for Defence and National Security. http://www.ictoutlookforum.com.au/asset/298.pdf

Schutz, A. 1967. *The Phenomenology of the Social World*, [trans. by G.Walsh and F. Lehnert, of *Der sinnhafte Aufbau der sozialen Welt: Eine Einleitung in die verstehenden Soziologie* (1932)], Evanston: Northwestern University Press.

Searle, J. R. 1980. “Minds, brains, and programs,” *The Behavioral and Brain Sciences*, Vol.3, pp.417-424.

Senge, P. M. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organisation*, New York: Doubleday Currency.

Sengers, P. 1998. *Anti-Boxology: Agent Design in Cultural Context,* Ph. D thesis, Pittsburgh, PA: Carnegie Mellon University.

Sengers, P. 1999. “Practices for Machine Culture: A Case Study of Integrating Artificial Intelligence in Cultural Theory.” *Surfaces*, Volume VIII, pp.1-58.

Sherburne, D. W. 1981 [1966]. *A Key to Whitehead’s* ***Process and Reality***, Chicago: The University of Chicago Press.

Sherburne, D. W. 1971. “Whitehead without God”, in Brown, D., James, R.E., & Reeves, G. 1971 Process Philosophy and Christian Thought, Indianapolis: Bobbs-Merrill, pp.305-328.

Sherrard, P. 1987. *The Rape of Man and Nature: An Enquiry into the Origins and Consequences of Modern Science*, Ipswich, Suffolk: Golgonooza Press.

Silverman, BG, Johns, M, Cornwell, J, O’Brien, K, 2006a. “Human Behavior Models for Agents in Simulators and Games: Part I – Enabling Science with PMFserv,” *Presence: Teleoper. Virtual Environ.,* Vol. 15, No. 2, pp. 139-162.

Silverman, BG, Johns, M, Cornwell, J, O'Brien, K, 2006b. "Human Behavior Models for Agents in Simulators and Games: Part II – Gamebots for a Foreign Culture," *Presence: Teleoper. Virtual Environ.,* Vol. 15, No. 2, pp. 163-185.

Simons, P. M. 1987. *Parts. An Essay in Ontology,* Oxford: Clarendon Press.

Singer, P. W. 2009. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*, New York: The Penguin Press.

Sinisi, V. T. 1966. "Lesniewski's Analysis of Whitehead's Theory of Events", *Notre Dame Journal of Formal Logic*, 7: 323-327.

Smelser, N. J. & Smelser, W. T. (eds.) 1963. *Personality and the Social System,* New York: John Wiley and Sons.

Smith, B. 1995. "Common Sense", in Smith, B. and Smith, D. W. (eds.) 1995 *The Cambridge Companion to Husserl,* Cambridge: Cambridge University Press, pp.394-437.

Smith, B. 1998. "The Basic Tools of Formal Ontology", in Guarino, N. (ed.) *Formal Ontology in Information Systems*, Amsterdam: IOS Press, pp. 19-28. Cohn, A. & Varzi, A. C. 2003, "Mereotopological Connection", *Journal of Philosophical Logic,* Vol. 32, pp. 357-390.

Smith, B. 2003a. "The Ecological Approach to Information Processing", in Kristóf Nyíri (ed.), *Mobile Learning: Essays on Philosophy, Psychology and Education*, Vienna: Passagen Verlag, 2003, 17–24. Avialable at: http://ontology.buffalo.edu/smith/

Smith B. 2003b. "Ontology", in L. Floridi (ed.), *Blackwell Guide to the Philosophy of Computing and Information*, Oxford: Blackwell, 2003, 155–166. Avialable at: http://ontology.buffalo.edu/smith/

Smith, D.W. 2005. "Truth and Experience: Tarski vis-à-vis Husserl" in *Experience and Analysis: Papers of the 27th International Wittgenstein Symposium*, J. C. Marek and M. E. Reicher (eds), Austrian Ludwig Wittgenstein Society, Kirchberg am Wechsel, pp.111-126.

Smith, D.W. 2007. *Husserl*, NewYork, N.Y.: Routledge.

Smith, T. S. 1992. *Strong Interaction,* Chicago: The University of Chicago Press.

Smith, T. S., and S. Benard. 2002. ''Compassion and Social Integration: Computational Investigations of Moral Sentiments and Normative Behavior.'' Paper presented at the Annual Meeting of the American Sociological Association, Anaheim, Calif.

Smith, T. S. 2003. "Where sociability comes from: Neurosociological foundations for social interaction", in Camic, C. & Joas, H. (eds.) 2003 *The Dialogical Turn: New*

Roles for Sociology in the Postdisciplinary Age. Lanham, MD: Rowman & Littlefield, pp. 199-220.

Smith, T. S. 2005. "Dialogicality and social representation: the dynamics of mind" [review of Ivana Markova], *Contemporary Sociology*, 34/3:331-332.

Smith, T. S. & Stevens, G. T. 1997. "Comfort regulation as a morphogenetic principle: local dynamics of dominance, competition and attachment", *Advances in Group Processes*, Vol. 14, pp. 113-155.

Smith, T. S. & Stevens, G. T. 2002. "Hyperstructures and the Biology of Interpersonal Dependence: Rethinking Reciprocity and Altruism", *Sociological Theory*, 20/1: 106-130.

Sokolowski, R. 1988. "Moral Thinking" in *Edmund Husserl and the Phenomenological Tradition: Essays in Phenomenology* (edited by Sokolowski, R.), Washington, D.C.: The Catholic University of America Press., pp.235-248.

Sokolowski, R. 1989. "What is Moral Action?" *New Scholasticism,* Vol. 63., pp. 18-37.

Sowa, J. F. 2000. *Knowledge Representation: Logical, Philosophical and Computational Foundations,* Pacific Grove, California: Brooks/Cole.

Spiegelberg, H. 1965. *The Phenomenological Movement: A Historical Introduction,* Vol. 1 (2nd edn.), The Hague: Martinus Nijhoff.

Stamper, R.K., 2001. "Organisational semiotics: Informatics without the computer?", in K. Liu, R.J.Clarke, P. Bogh Andersen, & R.K. Stamper (eds.), *Information, organisation and technology: Studies in organisational semiotics,* Boston, MA: Kluwer Academic Publishers, (pp. 117-171).

Steinbock, A. 1989. "Whitehead's 'Theory' of Propositions", *Process Studies*, Vol.18, No.3, pp. 19-29.

Sterling, L. and Shapiro, E.1994. *The Art of Prolog: Advanced Programming Techniques*, 2nd ed., MIT Press.

Squire, K. D. 2004. *Replaying History: Learning World History through playing 'Civilization III'.* Phd. Thesis, Indiana University.

Squire, K.D. 2006. "Games as ideological worlds." Available at: http://www.academiccolab.org/initiatives/papers.html

Stell, J. G. 2000. "Boolean connection algebras: A new approach to the Region-Connection Calculus", *Artificial Intelligence* 122(1–2):111-136.

Stojanov, G., Trajkovski, G. and Kulkov, A. 2006. "Interactivism in artificial intelligence (AI) and intelligent robotics", *New Ideas in Psychology*, 24:163-185.

Sun, R. (ed.) 2006. *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Cambridge: Cambridge University Press.

Sun, R. 2006. "Prolegomena to Integrating Cognitive Modeling and Social Simulation", Sun, R. (ed.) 2006 *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*, Cambridge: Cambridge University Press, pp.3-28.

Suppes, P. 1960. "A comparison of the meaning and uses of models in mathematics and in the empirical sciences", *Synthese,* Vol. 12: pp. 287-301.

Sycara, K. P. 2008. "Modelling Cultural factors in collaboration and negotiation", *Research Proposal, The Robotics Institute, Carnegie Mellon University*. http://www.cs.cmu.edu/~softagents/papers/sycara_modeling_cultural_factors_collaboration.pdf

Sycara, K. P., Gelfand, M., and Abbe, A., (eds.) 2009. *Proceedings of W:24 Modeling Intercultural Collaboration and Negotiation (MICON), at International Joint Conference on Artificial Intelligence (IJCAI'09),* Pasadena CA, July 13 2009. Available at http://www.cs.cmu.edu/~softagents/papers/IJCAI-09-MICON-Proceedings-final.pdf

Szakolczai, A. 2003. "Religion and the Civilising Process", in *Irish Journal of Sociology* 12/1: 39-44.

Szakolczai, A. 2006. "World-Rejections and World Conquests: The Dynamics of War and Peace", in Tilo Schabert and Matthias Riedl (eds) *Die Menschen im Krieg, im Frieden mit der Natur* (Würzburg, Königshausen & Neumann, 2006), pp. 147-63.

Tarski, A. 1944. "The Semantic Conception of Truth and the Foundations of Semantics", *Philosophy and Phenomenological Research* Vol.4. pp. 341 -376.

Tarski, A., 1953. "A general method in proofs of undecidability", in Mostowski, A., Robinson, R. M., and Tarski, A. 1953. *Undecidable Theories. Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, pp.1-35.

Taylor, G. Frederiksen, R. Vane, R. R. Waltz, E. 2004. "Agent-based Simulation of Geo-Political Conflict", Proceedings of the National Conference on Artificial Intelligence (IAAI2004), No. 19, pages 884-891. Available at: http://www.soartech.com/pubs/IAAI04-GTaylor-AGILE.pdf.

Taylor, T. L. 2006. *Play Between Worlds: Exploring Online Game Culture*, Cambridge, Massachusetts: The MIT Press.

Thagard, P. 2005. *Mind: Introduction to Cognitive Science (Second Edition),* Cambridge, Massachusetts: The MIT Press.

Thagard, P. 2008. "How cognition meets emotion: Beliefs, desires, and feelings as neural activity", in G. Brun, U. Doguoglu & D. Kuenzle (Eds.), *Epistemology and Emotions*, Aldershot: Ashgate. pp. 167-184. Available at: http://cogsci.uwaterloo.ca/Articles/how-cognition.pdf.

Thompson, E. 2007. *Mind in Life: Biology, Phenomenology and the Sciences of the Mind,* Cambridge, M.A.: Harvard University Press.

Thuraisingham, B., Khan, L., Kantarcioglu, M., and Khadilkar, V., 2009. "Design of a temporal geosocial semantic web for military stabilization and reconstruction operations", in *Proceedings of the ACM SIGKDD Workshop on Cyber-Security and Intelligence Informatics* Paris, France, pp. 63-74.

Townsend, D. 1993. "The Aesthetics of Joseph Priestley", *The Journal of Aesthetics and Art Criticism*, 51:361-71.

Urban, C. 2000a. "PECS: A Reference model for the simulation of multi-agent systems", in Suleiman, R. Troitzsch K. G. and Gilbert, N. (eds.) *Tools and Techniques for Social Science Simulation*, Heidelberg: Physica-Verlag, pp. 83-114.

Urban, C. 2000b. "PECS – A Reference Model for Human-Like Agents", in: Thalmann, D. (ed.), *Deformable Avartars*, Boston: Kluwer.

Urban, C. & Schmidt, B. 2001. "PECS – Agent-Based Modelling of Human Behaviour", *AAAI Technical Report FS-01-02*. Available at: www.aaai.org

Vanzago, L. 2003. *Temporality and Subjectivity in Merleau-Ponty and Whitehead: A Comparison,* unpublished Doctoral Dissertation, Katholieke Universiteit, Leuven.

Valenza, R. 2008. "Vector mathematics: symbol versus form", in Weber, M. & Desmond, W. 2008 *Handbook of Whiteheadian Process Thought*, [Vol. 2], Frankfurt: Ontos Verlag, pp.87-96; 149-156.

Varzi, A. C. 1994. "On the Boundary Between Mereology and Topology", in Casati, R., Smith B. & White G., (eds.) *Philosophy and the Cognitive Sciences*, Vienna: Holder-Pichler-Tempsky, pp.423-442.

Velmans, M. 2008. "Reflexive Monism", *Journal of Consciousness Studies*, 15(2), 5-50. Available at: http://cogprints.org/6103/

Velmans, M. 2009. *Understanding Consciousness* (2nd edn.), London: Routledge.

Vernon, D. Metta, G. & Sandini G. 2007. "A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents", *IEEE Transactions on Evolutionary Computation, Special Issue on Autonomous Mental Development* (in press). Accessed 26/11/07 at http://www.robotcub.org/misc/review2/06_Vernon_Metta_Sandini.pdf

Vijayakumar, M. V. & Davis, D. N. 2008. Development of a "Society of Mind" Cognitive Architecture based on the Principles of Artificial Economics", Paper under review for International Journal.

Wacquant, L. J. D. 1992. "Review of James Ostrow, Social Sensitivity: A Study of Habit and Experience", in *Theory and Society*, 21:893-897.

Wallis, P. 2005. "Believable conversational agents: introducing the intention map". In Pelachaud, C. Andre, E. Kopp, S. and Ruttkay, Z. (eds.), *Creating Bonds with Humanoids (Proceedings of the Workshop at AAMAS'05)* Utrecht University, the Netherlands, July 2005.

Waltz, E. 2002. "Data Fusion Conceptualization, Specification and Information Operations", in *Proc. of DARPA Ontology and Visualization Workshop*, Buffalo NY, November 18-22, 2002. Accessed online October 2007 at Workshop website (ppt presentation, under 'publications: conferences') - http://www.infofusion.buffalo.edu/

Wark, S.; Broughton, M.; Nowina-Krowicki, M.; Zschorn, A.; Coleman, M. Taplin, P. & Estival, D. 2005. "The FOCAL point - multimodal dialogue with virtual geospatial displays", *SimTecT 2005*. http://www.siaa.asn.au/get/2411856058.pdf.

Wark, S. and Lambert, D., 2007. "Presenting the Story Behind the Data: Enhancing Situational Awareness using Multimedia Narrative", *Military Communications Conference*,. MILCOM 2007. IEEE Volume , Issue , 29-31 Oct. 2007 Page(s):1 – 7.

Wark, S. Lambert, D. Nowina-Krowicki, M. Zschorn A. and Pang, D., 2009. "Situational Awareness: Beyond Dots On Maps To Virtually Anywhere" *SimTecT 2009 Simulation Conference: Simulation - Concepts, Capability and Technology (SimTecT 2009),* Adelaide, Australia, 15 - 18 June, 2009.

Wark, S., Zschorn, A., Broughton, M. and Lambert, D. 2004. "FOCAL: A Collaborative. Multimodal Multimedia Display Environment." *Proc. SimTecT 2004*. Available at - http://hdl.handle.net/1947/2398

Wark, S., Zschorn, A., Perugini, D., Tate, A., Beautement, P., Bradshaw, J.M., & Suri, N. 2003. "Dynamic agent systems in the CoAX binni 2002 experiment", in: *Proceedings of the Sixth International Symposium on Information Fusion*, Cairns, Australia, 2003.

Wassermann, C. 1988. "The Relevance of *An Introduction to Mathematics* to Whitehead's Philosophy", *Process Studies*, Vol. 17, pp. 181-192.

Weber, M. & Desmond, W. 2008. *Handbook of Whiteheadian Process Thought*, [Vol. 1 & 2], Frankfurt: Ontos Verlag.

Weiss, G. (ed.) 1999. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence,* Cambridge, Massachusetts: The MIT Press.

Weiss, G. Nickles, M. Rovatsos, M. Fischer, F. A. 2007. "Specifying the intertwining of cooperation and autonomy in agent-based systems", *Journal of Network and Computer Applications*, 30/3 (August): 1196-1215. Available at: http://www7.in.tum.de/~weissg/weissg-pub.html

Wetzel, L. 2008. "Types and Tokens", *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*, Edward N. Zalta (ed.). Available at: http://plato.stanford.edu/archives/win2008/entries/types-tokens/

Weyns, D. Van Dyke Parunak, H. Michel, F. (Eds.) 2005. *Environments for Multi-Agent Systems: First International Workshop* E4MAS 2004*, Lecture Notes in Computer Science* Vol. 3374, Berlin / Heidelberg: Springer-Verlag.

Weyns, D. Parunak, H. Michel, F. Holvoet, T. and Ferber, J. 2005. "Environments for multi-agent systems, state-of-the-art and research challenges", in Weyns, D. and Parunak, H.V.D. and Michel, F., (eds.), *Environments for multi-agent systems*, *Lecture Notes in Computer Science*, vol 3374, pp. 1-48.

Whitehead, A. N. 1898. *A Treatise On Universal Algebra With Applications* Cambridge: Cambridge University Press.

Whitehead, A. N. 1906. "On Mathematical Concepts of the Material World", *Philosophical Transactions: Royal Society of London,* series A, Vol. 205, pp. 465-525, reprinted in *Alfred North Whitehead: An Anthology* (Eds. Northrop, F. S. C., & Gross, M. W.) Cambridge: Cambridge University Press, pp. 11-82.

Whitehead, A. N. 1911. "The Philosophy of Mathematics", in Armstrong, H. E., Farmer, J. and Freeman, W. G. (eds.), *Science Progress in the Twentieth Century: A Quarterly Journal of Scientific Thought and Work*, Vol. V, July 1910 – April 1911, London: John Murray, pp. 234 – 239.

Whitehead, A. N. 1910-13. (with Russell, B.), *Principia Mathematica*, 3 vols. Cambridge: Cambridge University Press.

Whitehead, A. N. 1911. *An Introduction to Mathematics,* London: Williams and Norgate.

Whitehead, A. N. 1916. "La Théorie Relationniste de l'Espace", *Revue de Metaphysique et de Morale,* Vol.23, pp.423-54.

Whitehead, A. N. 1917. "The anatomy of some scientific ideas", in Whitehead, A.N. 1929 *The Aims of Education and Other Essays*, New York: MacMillan, pp.180-231.

Whitehead, A. N. 1919. *An Enquiry Concerning the Principles of Natural Knowledge.* Cambridge: Cambridge University Press.

Whitehead, A. N. 1920. *The Concept of Nature*. Cambridge Uni. Press. 2004

Whitehead, A. N. 1922. *The Principle of Relativity with Applications to Physical Science*. Cambridge: Cambridge University Press.

Whitehead, A. N. 1926. *Science and the Modern World*, Cambridge: Cambridge University Press.

Whitehead, A.N. 1927. *Symbolism: Its Meaning and Effect,* New York: Macmillan.

Whitehead, A. N. 1978 [1929]. *Process and Reality: An Essay in Cosmology,* (Corrected edition, Eds. Griffin, D. R. and Sherburne, D. W.). New York: The Free Press.

Whitehead, A. N. 1942 [1933]. *Adventures of Ideas*, Harmondsworth, Middlesex: Penguin Books [N.B. the pagination in the Penguin (Pelican) 1942 edition differs from the MacMillan & Cambridge 1933 editions].

Whitehead, A. N. 1934. "Indication, classes, numbers, validation", in Whitehead (1948), pp. 227-240.

Whitehead, A. N. 1937. "Analysis of Meaning", in Whitehead (1948), pp. 93-99.

Whitehead, A. N. 1938. *Modes of Thought*, New York: The Free Press.

Whitehead, A. N. 1948. *Essays in Science and Philosophy*, London: Rider and Company.

Whitehead, A. N. 1951a [1941]. "Mathematics and the Good", in Schilpp, P. A. (ed.) 1951, *The Philosophy of A. N. Whitehead*, (2nd edn.) La Salle, Illinois: Open Court, pp.666- 681.

Whitehead, A. N. 1951b [1941]. "Immortality", in Schilpp, P. A. (ed.) 1951, *The Philosophy of A. N. Whitehead*, (2nd edn.) La Salle, Illinois: Open Court, pp. 682-700.

Willey, B. 1961. *The Seventeenth Century Background: Studies in the Thought of the Age in Relation to Poetry and Religion,* London: Routledge.

Willey, B. 1940. *The Eighteenth Century Background: Studies in the Idea of Nature in the Thought of the Period,* Harmondsworth: Penguin.

Wilson, F. 1985. "The Lockean Revolution in the Theory of Science", in S. Tweyman and G. Moyal, (eds.) 1985, *Early Modern Philosophy: Metaphysics, Epistemology and Politics,* Delmar, N.Y.: Caravan Books, pp.65-97.

Winch, P. 1958. *The idea of a social science and its relation to philosophy*, London: Routledge.

Wolfram, S. 2002. *A New Kind of Science,* Champaign, IL: Wolfram Media Inc.

Wolterstorff, N. 1996. *John Locke and the Ethics of Belief*, Cambridge: Cambridge University Press.

Woolridge, M. 2002. *Introduction to Multi-Agent Systems,* John Wiley and Sons.

Wrong, D.H. 1961. "The Oversocialized Conception of Man in Modern Sociology", *American Sociological Review*, Vol. 26, No. 2 (Apr., 1961), pp. 183-193

Wyman, M. A. 1956. "Whitehead's Philosophy of Science in the Light of Wordsworth's Poetry", *Philosophy of Science*, 23/4: 283-296.

Yang , K., Lo, A. & Steele, R., 2007. "Ontology-Based Multi-agent System for the Accommodation Industry", *Proceedings of The Thirteenth Australasian World Wide Web Conference 07.* Available at: http://ausweb.scu.edu.au/aw07/papers/refereed/yang/paper.html

Zacharias, G. L., Jean MacMillan, J. and Van Hemel, S. B., (eds.), 2008. *Behavioral Modeling and Simulation: From Individuals to Societies,* Washington: National Academies Press (pre-publication copy).

Zaino, J. 2008. "What's Next: Intelligence at the Interface" - Interview of Tom Gruber. Available at: http://www.semanticweb.com/article.php/3746011

Zambonelli, F. Parunak, H. V. D. 2003. "Signs of a Revolution in Computer Science and Software Engineering", in Petta, P. Tolksdorf, R. Zambonelli, F. (eds.) *Engineering Societies in the Agents World III*, Third International Workshop, ESAW 2002, Madrid, Spain, September 16-17, 2002, Revised Papers. Lecture Notes in Computer Science 2577 Springer, pp.13-28.

Zúñiga, G. 2001. "Ontology: Its Transformation from Philosophy to Information Systems," *Proceedings of the International Conference on Formal Ontology in Information Systems*, Ogunquit, Maine, USA , 2001, pp. 187-197.