University of Southampton

A Phenomenological-Connectionist Theory of Computational Agency

Volume I of II

by

Daniel W. Joyce

A thesis submitted for the degree of
Doctor of Philosophy

in the
Faculty of Engineering and Applied Science
Department of Electronics and Computer Science

April 2001

University of Southampton

ABSTRACT

Faculty of Engineering and Applied Science

Department of Electronics and Computer Science

Doctor of Philosophy

A Phenomenological-Connectionist Theory of Computational Agency

Volume I of II

by Daniel W. Joyce

This thesis presents a theory of computational agency. Computational agent theory differs from 'classical' artificial intelligence by committing to the view that a computational artifact is situated, and that its rationality is limited by the constraints of this 'situatedness'. Contemporary literature is surveyed and models of situated computational agency placed in their philosophical contexts. From this, a critical reconstruction of the notion of an agent is given from a phenomenological perspective. It is proposed that everyday *routines of activity* underpins agency and computational implementations of this substrate can take the form of connectionist networks.

The proposal is tested in technical practice on two domains; agents for multimedia content-based navigation/retrieval, and a simulated environment which explores the key properties of the proposed phenomenological agent theory. Recent proposals for goal-directed behaviour in connectionist systems (largely from the cognitive and behavioural neurosciences) are critically evaluated, and integrated into an agent architecture. This results in an architecture utilising suitably controlled reinforcement learning. The architecture implemented is then evaluated against the agent theory, and examples of 'routine behaviour' analysed in stationary and non-stationary environments.

Semiotic analyses are then proposed as an alternative theory of representation, as they are compatible with, and simultaneously possess explanatory power at a level beneath, the usual sentential/propositional level.

The thesis contributes a phenomenological theory of agency and gives examples of its influence on technical practice. Outlines of connectionist architectures are presented, implemented, and evaluated with respect to the agent theory proposed.

Acknowledgements

I would like to extend thanks and gratitude to Dr. Paul Lewis, whose supervision gave both the necessary freedom to explore ideas and invaluable direction to enable the completion of this thesis. Also, debts of gratitude and thanks are owed to Dr. Stuart Gilson and Jim Beacom (1974–1998) for years of valued friendship, advice and support. Thanks also to Dr. Rob Tansley and Dr. Mark Dobie, for help and support with software integration in the MAVIS2 project.

Contents

List of Figures

List of Tables

Part I

Literature Review and Critique

Chapter 1

Introduction

1.1 Project History

This thesis represents a culmination of work in a number of different directions originally inspired and funded by the EPSRC project MAVIS2 (Multimedia Architecture for Video Image and Sound). The original project (MAVIS1) aimed to provide content-based retrieval and navigation in multimedia information spaces. Part of the proposal for MAVIS2 was to include intelligent software agents which would assist in retrieval and navigation in such multimedia systems. Principally, they would represent autonomous entities which acted to combine low level media information with higher level knowledge utilising pattern classification techniques such as those implemented by artificial neural networks. The knowledge level would be provided by a restricted conceptual network (Sowa, 1984) of symbolic information provided during authoring of a multimedia application. Broadly speaking, the knowledge level can be considered an taxonomy and the lower level media information as signifiers or exemplars representative of the taxonomic categories.

Under a superficial examination, the agents can be seen as classification engines embedded in some autonomous software unit as, for example, those given in (Bigus and Bigus, 1998). However, this ignores an important qualitative difference which arises if an agent utilises artificial neural network, or as Smolensky called them, sub-symbolic techniques. Much agent theoretic research is concerned with questions about what an agent *should be*, and the analysis is ostensibly at the epistemic level. An agent is said to 'possess an ontology' relevant to its domain of discourse. This is in some sense represented or contained by an agent's internal state. In the case of traditional purely symbolic approaches, there is systematic interpretability of the agent's internal state : in fact, the design of the mechanisms for initialising and maintaining such state guarantee a direct correspondence between the environment and a unique denotative symbol in the agent's internal state.

Issues of self-organisation, learning and representation *with respect to* the agent's perception of the environment are all issues which need a coherent theoretical basis. Presently, agent theory largely concerns itself with questions such as whether game theory, market economic models, Belief Desire and Intention (BDI) models and recently, automata theoretic models (Wooldridge, 1999), should form the basis of agent architectures and implementations. These agent theories have an ontological commitment to traditions in AI, cognitive science and philosophy which might well be examined in a 'pre-ontological' fashion to find what assumptions or commitments there are. In this thesis, the very nature of agency is taken to be phenomenological and this commands different approaches. A broadly automata-theoretic approach was taken based on a modified situated automata theory (Rosenschein and Kaelbling, 1995).

The question for the MAVIS2 agents (again, at the usual agent theoretic level of analysis) is what would it mean if a software agent implemented its internal state via a connectionist, sub-symbolic system such as an artificial neural network? The contrasting ideology being the use of explicit knowledge structures to build such software agents. Immediately, the denotation of the agent's internal state (i.e. epistemology) is not as obvious and, depending on the history of agent-environment interactions, the internal state of two similar agents might be different.

At the level of grounded implementation, there is a distinction between connectionist and knowledge-level internal state. If we follow the literature, there are clear philosophical divides about the mechanisms agents use to implement their observable behaviour. The classic distinction between these two mechanisms was explicated by Smolensky (1988) who observed that the key difference is the *level of representational token*. In the symbolic agent, the denotation of the internal symbol (the token) is in direct correspondence with the real world. Smolensky claims that the level of representation is coincident with the level of computation. This kind of Platonic correspondence supports the traditional view of AI as being internalised, central world models with invariant features which are 'mentally' processed. In the sub-symbolic agent this may not be true: the transient activation of an artificial neuron may not explicitly denote one feature of the world. This kind of distinction draws a continuum of localised *versus* distributed computation. Some have argued this distinction is artificial in a philosophical context for example, see (Ramsey, 1997) and (Fodor and Pylyshyn, 1988). Despite these debates, there is certainly an essential difference between the implementation of behaviour on a symbolic as opposed to a sub-symbolic substrate. Notably, the latter uses continuous mathematics to produce behaviour, relying on (usually) stochastic interpretations of behaviour, probable outcomes and processing in an "holistic" way (that is, each element of the input is simultaneously used to compute the outcome with little procedural reasoning or deliberation). The former relies on representation and (usually) the discrete formalism of logic as both procedural and declarative representation of planning and knowledge respectively. Given both the technical promises and philosophical dichotomies, this question has been given some weight in the consideration of reactive agents and their implementation with connectionist substrates.

If computational agency is more than a programming methodology, a metaphor for distributed systems or pessimistically *only* a metaphor, then an agent theory is required which is general enough to span both physical and virtual agents in a plausible way, without explicit ontological commitments to deliberative planning, theorem proving, internal representation or sub-symbolism. This 'unifying theory of agency' was alluded to in (Huhns and Singh, 1998) and might be compatible with (or possibly identical to) complex adaptive systems theory cf. (Holland, 1975; Holland, 1995). Other disciplines have begun using agent theory to explore problems, such as (Liebrand, Nowak and Hegselmann, 1998) which describes modelling of social processes from a variety of computational disciplines including distributed AI, agents and cellular automata. This has caused certain controversies, such as the validity of evolutionary algorithms (as mechanisms explaining ideological conflict in populations) and their covert technological assumptions.

In exploring agent theory from the computational and phenomenological perspectives (see Chapter 5) it is hoped this thesis offers some contribution in the direction of an 'agent science'.

1.2 Thesis Position and Methodology

Much of the computational agents paradigm is concerned with the description of systems *as if they were* entities possessing intentional properties. Hence, there has been considerable recasting of many existing systems using the agent metaphor. The classic degenerate cases being Shoham's light switch agent (Wooldridge and Jennings, 1995; Watt, 1996) and McCarthy's thermostat agent (McCarthy, 1978).

One principle common to most intelligent software agent enterprises is the commitment to the symbolic AI programme; both implicitly (in the use of discourse to describe the agents) and explicitly (in technical practice, e.g. the use of declarative, knowledge-based languages for implementing agents). While rarely made explicit, landmark papers (Wooldridge and Jennings, 1995; Jennings, Sycara and Wooldridge, 1998; Shoham, 1993; Rao and Georgeff, 1991; Rao and Georgeff, 1995) constantly employ characterisations of the agent and its environment in the symbolic, first order and modal logics. However, more recently, the concept of agent and environment has been given a clearer taxonomy by Wooldridge (1999) reflecting a trend to consider the interactions of software agents and environments in a fashion familiar to behaviour-based robotics and ethologically motivated agents research. Even still, the characterisation is given in terms of modal and predicate logic. Ironically, other research programmes more concerned with adaptive behaviour (e.g. reinforcement learning) have used classical uncertainty to characterise this relationship. Environments are described as stochastic systems with probabilistic state dynamics e.g. (Sutton and Barto, 1998). Again, this emphasises the disparity between the software and robotic agents community.

If we include the broadly computational disciplines of robotics and physical agents, there has been a trend oppositional to traditional AI such that the 'new-AI' ideology is to move away from Platonic models of the agent as a deliberative actor in its environment (Clancey, 1997). The much cited work of Brooks (1986), Agre and Chapman (1987) and Rosenschein and Kaelbling (1995) on so-called reactive agency has emphasised the commitment to theories of agency which do not rely solely on the traditional deliberation, planning and world-modelling ideologies of traditional AI. Again, these arguments have both a philosophical and technical grounding. Brooks ostensibly defends his perspective first on technical points, later on more philosophical issues. Agre and Chapman approached the problem from a philosophical perspective and then developed the corresponding technology. The notion of critically examining the implicit philosophy of a particular technical approach is given extensive treatment in the AI context by Agre who states : "Technology at present is covert philosophy; the point is to make it openly philosophical." – (Agre, 1997), pp. 240. This principle of critical reconstruction reveals some difficulties in the software agents paradigm which are explored in this thesis.

Brooks' proposal and manifesto questioned the core methodology of AI. However, Etzioni (1993) defended software agents research in the light of Brooks' manifesto (Brooks, 1991a; Brooks, 1991b) whose monist proposal for revising the goals of AI was generally to; a) orient all research towards situated action in real environments, and b) remove abstraction and model building from technical practice. Etzioni disagrees, claiming Brooks had ignored certain fundamental abstractions in his own work (that of a digital solid state substrate for implementation) and that further still, a class of virtual agents called "softbots" could in fact be described in agent/environment terms which (whilst abstract in terms of traditional robotic notions of taxis and locomotion) were equally as "situated"[1] and direct as any physical robot. It is this category of "softbots" (which here are referred to as reactive software agents) which are the focus of this thesis.

Despite there being very little consensus on the very notion of agency, what can be ascertained is that work on agents usually divides in two; a philosophical approach (the agent-theoretic arguments) which then underpins a technical discipline (the implementation of agents under such a theory). Genesereth and Ketchpel (1994) provide a useful division of the broad domain of computational software agents research. They describe the efforts as :

- Agent theories: how agents are conceptualised, what properties agents should possess and what are the suitable models for analysing agents (e.g. game theory, economic models, complex adaptive systems etc.)

- Agent architectures: moving from specifications of the character of agency to architectures and implementations that support the principles from agent theories.

- Agent languages: tools to allow the implementation of agents given an agent-theoretic

[1] An agent which is described as *situated* or as having the property of *situatedness* is one where it is in continuous sensori-motor interaction with the environment

stance and architecture. What are the appropriate primitives and paradigms (e.g. procedural or declarative) for these construction tasks ?

This thesis is concerned with the notion of adaptive behaviour in reactive software (or virtual) agents. Specifically, it considers the issues of providing what might be termed *learning* in a vernacular sense. Superficially, it might stated that this work is concerned with agents which acquire rules governing their behaviour through adaptation and situated activity as opposed to being explicitly programmed, as in the work of (Yang, Pai, Honavar and Miller, 1998). Following from (Agre, 1997) and the critique of (Agre, 1995), this thesis is situated somewhere between agent theory and architectures, in that arguments about the foundations of agents and appropriate architectures are considered. Agent theory (i.e. what is characteristic of 'agency') is taken from a phenomenological standpoint, focusing on routinised activity without explicit reasoning (cogitation or deliberation). In terms of machine learning, it lies between reinforcement learning, associative neural networks and learning automata. Some necessary assumptions have been made; notably that connectionism (specifically, the class of techniques falling under the slogan artificial neural networks) provides a valid and potentially fruitful source of theories and techniques for discussing adaptive behaviour and that there is some validity (both philosophically and technically) to *some* of the claims of the reactive, non-deliberative manifestos.

Artificial neural networks have long been a source of controversy, both in terms of their equivalence to traditional AI and, in addition, their roles as cognitive models. Some researchers focus on the precise neuronal dynamics which mimic phenomena observed in biological neurons i.e. computational neuroscience. Others study them as abstractions of loosely coupled cognitive systems where the phenomena of interest are the behaviour and dynamics of the network. In any case, a consistent presence is the notion of adaptive or learning behaviour.

Clearly, there is some benefit in agents which possess the associative learning capacity of connectionist networks, but simultaneously, such approaches defeat the verifiable semantics thesis strongly sought by a majority of the software agents community; specifically, those concerned with agent-based software engineering e.g. see (Jennings, Sycara and Wooldridge, 1998). It can be seen that superficially a technique relying on stochastic or non-deterministic methods is not a reliable, robust model suitable for deployment in applications requiring verifiability. However, they possess behavioural characteristics with undeniable value in engineering systems which are situated in environments cf. those divorced from environments such as the classic expert systems (Jennings, Sycara and Wooldridge, 1998). Notably, such agents are capable of learning and adapting learned behaviour to new environmental contingencies.

Despite focusing on so-called sub-symbolic techniques, this thesis should not be read as being oppositional to other agent theories because some application domain specific functionality is necessarily available only as knowledge level rules. As an example, consider information retrieval agents. An agent embedded in a text-only environment will be asked to respond to certain queries or requests explicitly stated in terms relevant to the application. Beneath this specific functionality, a more generalised architecture might support features such as learning and acquisition of retrieval models (for example, probabilistic classification of document relevance with respect to a given query). This hypothesis suggests a vertically layered architecture, where a domain specific top layer integrates with lower levels which provide the basic functionality of that class of agents.

This is a typical view of software agents. The principle difference, then, between software agents research and the position of this work is that the agent-theoretic and technical substrate is not classical symbolic logic. Moreover, it will be argued that in adopting this stance, notions of the agent and environment need reconsideration and we find in the literature a wealth of different, and potentially valid, ideas for questions which are not easily conceived of in terms of symbolic logic cf. (Dreyfus, 1992).

1.3 Goals and Objectives

The goals and objectives of this thesis are :

- to determine the implications for a corresponding agent theory and architecture if learning and adaptation are taken as core features of an intelligent software agent
- to investigate the scope for migrating principles across the boundary of implementation and establish if the principles of software agents and physical agents (with their sometimes divergent philosophies) are distinct by virtue of one being virtual and the other physical
- to develop an alternative (but not exclusive) theory of computational agency not bound to strictly sentential-symbolic models and which take the notion of situatedness to be the foundation of the theory and subsequent models
- to identify appropriate learning mechanisms which might form the basis for such an architecture – including their connections to multi-level cognitive modelling.
- to explore how traditional artificial neural network techniques can be considered as agent control mechanisms where the supervisor becomes the environment and not an explicit error signal
- to explore how the traditional supervised learning model of *training then performance* can be broken in order that training and performance are concurrent (a necessary feature of an agent engaging in interactions with its environment).
- to explore applications to demonstrate the viability of such software agents e.g. their use in MAVIS2 and simulations

Although these goals may seem numerous and diverse, a survey of literature and techniques quickly reveals that they are closely related. For the purposes of evaluating the thesis, they have been stated explicitly and separately.

1.4 Structure

The thesis is divided into two volumes – this volume is the substantive work, and the second volume includes the two technical reports (Joyce, 2001b; Joyce, 2001a) which include detailed data from experiments and other exploratory work on reinforcement learning models.

This volume is structured as follows:

- *Part I* – presents an overview of current and seminal literature in agent theory and shows the relationship to agents that use connectionism to control and generate behaviour. Chapter 2 focuses primarily on software agents, extracting core themes which serve to differentiate 'reactive' and 'deliberative' agent theories. It emphasises the emerging trend to view agent theory as variations on the themes of bounded rationality and embodiment, rather than the 'strong' and 'weak' definitions of (Wooldridge and Jennings, 1995) which focused on intentionality. It concludes by highlighting those factors most pertinent to the goals and objectives stated above. Chapter 3 elaborates on connectionist models of behaviour, mostly from the robotic agents community. The discussion focuses on properties of agency identified in the literature review, notably, internal state and the acquisition of experience from interactions with the environment. It concludes by showing how properties of a connectionist system can be mapped onto the characteristics of agents extracted from Chapter 2, focusing on internal state, production and reproduction of behaviour.

- *Part II* – contributes a connectionist model of perception and action. Chapter 4 demonstrates how a generalised perception faculty can be constructed and how reinforcement learning can be integrated with connectionism in two ways; as a model of Hebbian learning in a local, topographic neural network, and using function approximation in a distributed, multi-layer perceptron network. Chapter 5 then introduces a new phenomenological theory of agency, extending the work of (Agre, 1997) to be compatible with the neural network techniques discussed in the preceding chapters. The chapter emphasises how Heideggerian phenomenology, and his theory of 'breakdowns', can be extended to account for routinised activity in agents employing connectionist models. From this, semiotic explanations are sought which provide a complementary theory of 'representation'. The chapter concludes by showing how semiotic and phenomenological explanation adequately account for the peculiarities of connectionist 'representation', internal state and the production of routine behaviour.

- *Part III* – introduces the two substantive experimental environments in which the agent theory presented evolved. Chapter 6 describes the application of such an agent theory to an adaptive pattern classification task. The agents are generalised inductive concept learners, bridging the gap between low level feature-based information from multimedia data and knowledge-level information provided by a taxonomic knowledge representation (the Multimedia Thesaurus). This work primarily emphasises the hermeneutic aspects of the agent theory, leaving explorations of embodiment and routine behaviour for Chapter 7, which describes a parsimonious simulated environment which facilitates exploration of routine activity in stationary and non-stationary scenarios.

- *Part IV* – discusses the construction of an agent architecture which attempts to explore the fuller version of the agent theory of Part III. In particular, Chapter 8 describes steps toward integrating the concepts of Chapters 3, 4 and 5 to construct an autonomous agent which can perform in the simulated environment of Chapter 7. Chapter 9 evaluates the architecture of Chapter 8 with respect to the phenomenological agent theory. Throughout, the integration of the hermeneutic, embodied and phenomenological components of the agent theory are emphasised. This naturally leads to a discussion of model building, and a constrained connectionist approach is taken (which is to say, the phenomenological framework constrains the model 'top-down' while motivation for the connectionist implementations of perception and action are constrained by cognitive and neuropsychological principles 'bottom-up'). This represents an attempt toward the 'naturalisation of phenomenology' enterprise – i.e. placing phenomenal experience in a scientific context, which in this thesis, is through connectionism and agent theory. The thesis is concluded in Chapter 10, where further work on the agent theory and its implementation strategies are described.

Chapter 2

Contemporary Agents: Theory and Practice

For over five decades, research has attempted to synthesise intelligent behaviour in machinery. During this period, many disciplines have contributed to this field and various schools have emerged. One recent addition that has flourished has been variously labelled 'agents', 'intelligent agents' and 'multi-agent systems'.

This field has borrowed from and been inspired by AI, cognitive science, and philosophy. This is the familiar view for computer science so we will begin here with an exposition of this area. From this, we will examine the broader context to which the computer science paradigm of agents belongs and reconsider its position with respect to the disciplines that inspired the field.

After summarising current ideology on "computational" agents, the areas which have received little attention (such as a model of learning and reactive agency) are detailed. By broadly polarising the agent community into deliberative and reactive, it is possible to explore a continuum of agent characteristics. It is proposed that *most* agents research (particularly, disembodied, virtual intelligent software agents) is heavily situated in traditional assumptions from the AI community (symbolic AI or *deliberative agents*) and that in diverging from AI, we have avenues of research activity in the (so called) "new-AI" (Dorffner, 1997) school that can contribute to the difficult problems of agent learning, adaptation and coping with uncertainty in their environments. A caveat is also required: techniques from new-AI may not necessarily substitute symbolic theories; it is probable that a hybrid of high level cognitive ability in agents can be provided by symbolic theories of reasoning accompanied by (but not operationally separable from) the lower level mechanisms. This debate will be futher investigated in later sections of this chapter.

2.1 Agents from a Computer Science Perspective

The computer science perspective on agents is not well established but there exists a common body of theory that is cited as the foundation of most agent applications. The field of computer science agents is often cited as being divided into two categories (Wooldridge and Jennings, 1995)

- *weak agency*: characterised by properties of *autonomy, social ability, reactivity* and *pro-activeness*. This is often interpreted as a natural extension to the object oriented methodology.
- *strong agency*: characterised by belonging to the AI community, defined in terms of mentalistic notions such as belief, desire and intention (Shoham, 1993) and possessing meta-knowledge and the ability to reason over it.

Other workers have defined similar categorisations; notably (Nwana and Ndumu, 1996), and the stronger definition of agents *not* simply as a metaphor but a goal of research in intelligent systems (Watt, 1996). A more thorough discussion of the ideology of agents in computer science can be found in (Dale, 1997; Genesereth and Ketchpel, 1994; Wooldridge, 1999; Jennings, Sycara and Wooldridge, 1998).

Agents can be seen as a natural extension to the object oriented paradigm. An agent can be viewed as an autonomous software entity that we are able to defer certain, possibly tedious or computationally complex, tasks to. Not only is the entity an encapsulation of state and behaviour, but it possesses the ability to select appropriate behaviour – for example, to decide on a suitable response to a certain request. In this capacity, an agent can be employed in many roles which are empowering computer technology but not necessarily characteristic of what the AI community view as intelligent agents.

2.2 Agent Theory and Technical Practice

In most agents literature, a distinction is introduced which can largely be attributed to work in adaptive robotics and a divergence from the assumptions of traditional AI. The most frequently cited work is (Brooks, 1986) and (Agre and Chapman, 1987) which advocates the *reactive* paradigm in contrast to the deliberative. The choice of the word 'reactive' is unfortunate, because it rarely describes the actual properties of behaviour-based agents entirely, and has connotations that suggest the agent *only* reacts to the current environment situation with no regard for the history of previous interaction. From initial successes, Brooks went on to define a manifesto (Brooks, 1991a; Brooks, 1991b) for producing intelligent systems that rejected wholesale the traditional approaches to AI.

Broadly, we can make the following distinctions (for agents):

- deliberative agency: centralised, symbolic representations of environment/world, planning and heuristics.
- reactive agency: no explicit or centralised symbolic representation of environment/world, no explicit planning or heuristics, varying levels of abstraction to define the granularity of behaviours.

Perhaps most pertinently, as (Agre, 1995) describes, work in engineering such systems carries tacit philosophical commitments such as those outlined above. Agre's attempt at critically reconstructing AI by reflecting on the practices and their implicit philosophy revealed that little serious attention has been paid to literature on embodiment, routinised situated activity and its implications for AI's brand of representationalism; usually formal symbol systems with denotational properties. In Chapter 5, a particular perspective on routine activity will be presented.

The manifestos for both deliberative and reactive paradigms are lengthy and detailed, but we should highlight the objections and assumptions of each.

2.2.1 Situatedness and Agents

The demarcating boundary between agents and classical AI seems to be the notion of situatedness. In as far as this is concerned, the current orthodoxy of computational agents theory has been (independently) characterised as an engineering of "transactional" systems. (Clancey, 1997) describes transactional systems as those which engage in situated sensori-motor loops. In terms of communication, Argyle and Kendon's social-psychological model (Argyle and Kendon, 1967) (outlined in Figure 2.1) gives us an appreciation which might justifiably shape computational agents. A transactional model binds the agent to the environment and the agent's state is then tied to the environment. However, it is the bounding of rationality (and its non-committal status on intentionality and representation) which justifies the leverage of Argyle and Kendon's model. If agents are to be genuinely synthetically social artefacts, then a model such as that shown in Figure 2.1 orients all action and communication in a necessarily bounded rationality.

This model is close (although not nearly as computational in origin) to Wooldridge's recent work (Wooldridge, 1999) which, in an attempt to generalise what an agent is (agent theory) at an appropriate level, sides with automata-like systems.

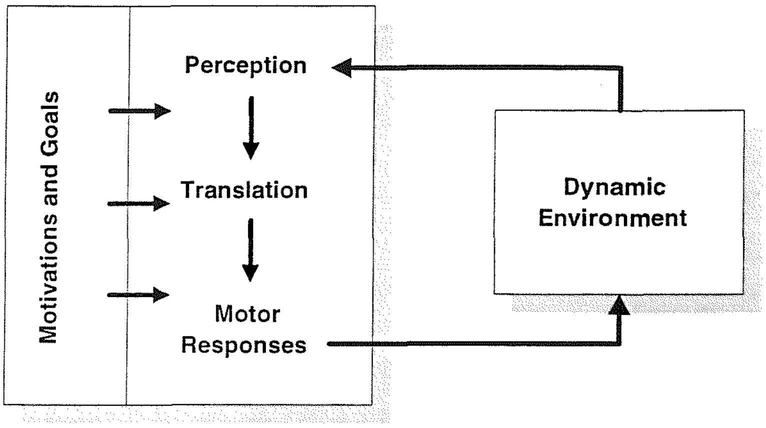

Figure 2.1: Towards a Transactional Model for Reactive Agents

---

2.2.2 The Character of Reactive Agency

Examples of work in agent theories and architectures[1] that include reactive components, or are wholly reactive, include:

- Brooks (1986) subsumption architecture (mobile robotics)
- Action-selection mechanisms (Maes, 1989)
- Agre and Chapman (1987) Pengi
- InteRRaP (Müller, Pishel and Thiel, 1995) – activity in the vertical layers is mutually exclusive
- TouringMachines (Ferguson, 1991) – activity is parallel in the vertically layered architecture cf. (Verschure and Voegtlin, 1998)

However, applications of these architectures and theories are limited. With the exception of Pengi, the other reactive architectures have been largely concerned with embodied, physical mobile agents. Many of the architectures have been simulated or developed without embodiment, excepting subsumption and its accompanying manifesto. The features of reactive agency that distance it from deliberative agency are:

- No explicit symbolic representation of the world
- Solipsistic sub-systems and behaviour modules (especially, in subsumption architectures)

[1]Research which has been conducted by the mainstream agents community under the label 'reactive' is given here – this does not necessarily reflect other work which might also reasonably be included under the reactive paradigm.

- Behaviour-oriented control systems (for example, the object avoidance behaviour is a "primitive"control module in Brooks' subsumption implementation) designed to meet specific intelligent task requirements (cf. 'top-down' functional decomposition for deliberative AI)
- Emergent behaviour from cooperating or competing sub-systems
- Simple agent behaviour can combine to produce societal level behaviours which are deemed intelligent

The proposal for intelligent behaviour without explicit representation, e.g. see the collection of papers spanning 1986-1991 given in (Brooks, 1999), has gained widespread interest. Indeed, one possible implementation substrate for implicit, or distributed, representation (at a finer granularity than symbols that refer to world objects) is the connectionist, or sub-symbolic, paradigm.

2.2.3 Critique of the Reactive/Behaviour-Based Methodology

Critics of reactive agents claim that:

- Reactive agents have only demonstrated simple goal directed behaviour
- Emergent dynamics are hard to analyze (cf. the computational irreducibility of cellular automata, where the observed patterning of behaviour can only be produced by explicitly iterating the equations governing system dynamics)
- It is difficult to see how individual components can be designed to guarantee the desired global behaviour (Maes, 1991)

Naturally these issues raise problems, as the deliberative community strive to provide formalisms for designing and implementing agents, such as agent oriented software engineering (D'Inverno, Fisher, Lomuscio, Luck, De Rijke, Ryan and Wooldridge, 1997). With specific reference to subsumption, we should note that the granularity of the control modules and *just how* solipsistic they are will determine the information carrying capacity of connections between them. Brooks' proposal is that they send varying signals to other modules. For example, the "feel force" module (Brooks, 1986) provides a force signal to the "runaway" module. This signalling line is essentially an information or content bearing entity: over time the "variable" represented by the connection between modules indicates the force encountered by a sensor. Subsumption, therefore, does not necessarily imply total rejection of a representational intentionality. Perhaps more specifically, Brooks' signalling lines between modules represent *his* interpretation of a similar biological agent's sensor to motor inter-operability. Brooks' later recanted his broad "intelligence without representation" dictum and replaced it with "without explicit, symbolic centralised representation". The clear dividing line is that reactive architectures and theories do not rely upon formal abstractions of the world, encapsulated in model-theoretic conceptualisations such as those examples given in (Guarino and Giaretta, 1995).

In addition, there are many debates about what constitutes a representation and what constitutes reasoning in the abstract formal logic sense. Hybrid reactive/deliberative systems demonstrate that pluralism has its advantages: we can inherit the fundamentals of reactive agency, but diverge from Brooks' manifesto for the sake of engineering pragmaticism. In the example given in the introduction, we might require agents that have an explicit symbolic capacity simply in order to provide necessary anthropomorphism so that the agent communicates to the user effectively.

We also have the additional burden of considering the polemics of, and the culture that the reactive paradigm inhabits. Brooks, Agre and Chapman are seen as the most vocal exponents of this paradigm but Brooks attempted to espouse his reactive paradigm as an all-encompassing approach to intelligent systems research and in doing so, also rejected *any* of the classical AI approaches. However, this leaves Brooks' thesis open to attack. (Etzioni, 1993) reminds us that Brooks argues for mimicking lower animal intelligence (such as that of insects) because evolution produced such organisms first. There is, however, contrary evidence: Brooks fails to realize that by accepting digital computers as a substrate (be they situated in robots that inhabit real worlds or otherwise) he has effectively bypassed the evolutionary epochs that results in a stable biochemistry for life.

This is not an attack at the engineering stance Brooks adopted to build robots, but on his extrapolation of this stance to a general philosophy of artificial intelligence. It also introduces us to debates surrounding computational theories of intelligence generally. Subsumption is, therefore, an effective engineering tool. It does not substantiate a reactive monism. Note that Minsky's *Society of Mind* (Minsky, 1986) is also similar to a reactive architecture: dedicated agents, in this case, perform relatively unintelligent tasks, but the collective behaves intelligently. The significant criticism of this proposal is that it is modular, which again is a criticism of the proposal as a general theory of intelligence and mind, not of its engineering utility. The rigid distribution of intelligent functionality over a collection of dedicated subsystems in this fashion denies the empirical fact that biological intelligence does not display the same separation of cognitive systems, perceptual systems and reasoning systems. Minsky's and Brooks' theories are similar apart from Brooks' insistence that embodied intelligence is the only viable route to achieving an artificial intelligence.

The reactive paradigm needs revision if it is to apply anywhere other than in embodied agents. There is very little detailed work on adaptive behaviour and learning in reactive agents that has been expressed in terms of software agents. This is largely due to most work in these areas being focused on sensori-motor ability and the provision of simple naturally occurring behaviours. In software agents it may be true that *some* sensori-motor ability is necessary (superficially, we can think of mobile software agents as possessing motor capacity) but a better focus might be on the issues that the deliberative paradigm has uncovered, such as language, communicative acts, co-operation and action selection mechanisms. There are certainly some challenges here.

2.2.4 The Character of Deliberative Agency

Deliberative paradigms inherit from classical GOFAI[2]. The emphasis is on:

- Correct representation which reduces the domain to problem solving
- Abstraction and decomposition: selecting the correct primitives for representing and acting on those representations by functional decomposition of the observed (or desired) competence
- Platonic split of agent and world (forms/world distinction: a form is some mental content internal to the agent, whereas the world is the external, distal objects this content refers to)
- Environment as divorced from reasoning (or cognitive ability)
- Mathematically formal systems of logic, behaviour, causality and knowledge
- Functional decomposition of system (mechanisms built from an analysis of how to perform the task, in contrast to behaviour oriented models of reactive agency)

Therefore, to correctly build *any* agent that can succeed in performing intelligent behaviour, it is necessary to build the *correct* representations of the world. That is to say, tokens or symbols that represent (or denote) the actual objects in the world which then define the domain of intelligent behaviour as manipulations on these symbols and produce or employ the correct axioms and rules of inference for manipulating symbols and their aggregations. This is shown in Figure 2.2.

The success of this philosophy is evident in knowledge engineering and expert systems where codified knowledge is manipulated via propositional and predicate logics. Exemplars of this approach include:

- Newell and Simon's (1963) article "GPS: A Program That Simulates Human Thought" described in (Charniak and McDermott, 1985), the physical symbol system hypothesis (Newell and Simon, 1976) and theorem proving by pattern-directed and resolution-based methods (Bratko, 1990) chapter 20
- The application of theorem proving and problem solving to mobile robotic agents (Fikes and Nilsson, 1971)

[2]Good Old Fashioned AI – a somewhat perjorative term that describes the deliberative, symbolic logic-driven approach of planning systems such as STRIPS

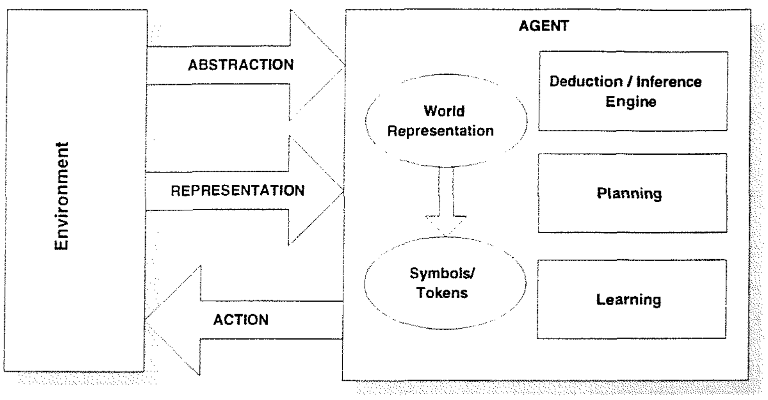

Figure 2.2: Schematic of a Deliberative Agent

- Winograd's SHRDLU and the natural language understanding effort (Winograd, 1973)
- Vision projects for mobile robotics, e.g. (Moravec, 1990) and model/knowledge based vision enterprises (cf. (Cliff and Noble, 1997) for a position on alternatives and the representation debate
- Albus'architecture for a general model of intelligent behaviour (Albus, 1991)

This tradition is distanced from the position taken by most computational agents workers (Wooldridge and Jennings, 1995; Wooldridge and Jennings, 1996; Genesereth and Ketchpel, 1994). The key notions are situatedness and the nature of rationality, which describes the coupled relationship of the agent with the environment and the constraints on decision-making.

2.3 Rationality and Agents

More recently, focus has shifted to characterising agents by first assuming they are situated entities and then focusing on the kinds of reasoning embodied in them. It is possible to identify the following broad categories:

- Perfect rationality
- Calculative rationality
- Purely reactive
- Belief, desire and intention (BDI) models and practical reasoning
- Bounded Rationality

2.3.1 Perfect Rationality

(Russell and Subramanian, 1995) defined *perfect* rationality as characteristic of the notion of an "agent" in economics and philosophy. The agent enjoys the property of being able to maximise the expected utility of any action taken at any time. This places the weakest of constraints on computational agency, where such an action is selected by an algorithm subject to the formal results of algorithmic complexity. No "real" agents can possibly exhibit this property since all deliberation and cogitation (and subsequent action selection) requires computation. The idea that an agent can decide in an infinitesimal time span over its complete world knowledge is clearly only a philosophical notion, unrealisable in natural and artificial agents.

2.3.2 Calculative Rationality

From the ideal of perfect rationality, *calculative rationality* defines an agent which receives some kind of perceptual input which cues it into deciding on its consequent action. After plan selection or building, execution of the plan results in relevant actions being taken. These actions *would* have been rational (e.g. would maximise the utility of the action in terms of goal satisfaction) if they had been taken at the time the percept was incident on the agent. Naturally, this definition embodies the idea that for a given percept p there exists a perfectly rational action a. In addition, there is some enumeration of every (p, a) which forms a policy which takes a small constant amount of time to search. Effectively under calculative rationality an action is in principle the best action for a percept, but is merely a theoretical label since it ignores the pragmatics of actually computing such a result.

2.3.3 Purely Reactive

Purely reactive agents needs little further explanation. The agent makes *some* decision and we make no claim on its rationality. Suffice to say that the most common manifestation of purest reactivity is in behaviour-based agents (Arkin, 1998). In this case, the rationality of action is defined by the micro-rationality of the behaviour or task modules which contribute to the macro behaviour of the agent.

2.3.4 BDI and Practical Reason

The notion of BDI architectures (and accompanying theory) is based on the notion of practical reasoning. Much has been imported from Dennett's work on intentional stance (Dennett, 1978; Dennett, 1987) about the necessary and sufficient internal mechanics for a computational agent to possess beliefs, desires and intentions. These proposals hinge on a form of calculative rationality since they usually employ modal logics (Rao and Georgeff, 1991; Rao and Georgeff, 1995; Shoham, 1993). They also hinge on something about which Dennett is deliberately non-committal, the *nature* of BDIs. To clarify, Dennett does not adhere to the notion that because we can reify a belief by disclosing it in propositional form that this is an internalised calculus of action for that agent. However, his powerful intentional systems theory permits hypothesis forming on *any* agent and its discussion using the linguistic device of the propositional attitude.

As a calculus of action, the operators, axioms and theorems of a given modal logic (e.g. see (Shoham, 1993) for a description) are explicit sentential-symbolic denotational systems. These enable systematic interpretability of an agent's internal state. A strong commitment to an internal calculus of action (Brooks, 1997) represents a philosophical commitment to computational agents as possessors of sentential-symbolic internal state, establishing a strong theoretical link (that is, exactly denotational) with the 'language of thought' proposition and its implications for agency (which are, on the whole, that propositional mentalism underpins cognition and therefore action).

Shoham built a simple modal system from which the following belief modality (derived from a theory of practical reasoning (Bratman, 1987) instantiated in logic) is most pertinent:

$$\mathrm{BEL}_a^t \Phi \tag{2.1}$$

For an agent a to believe something at time t (or apparently believe in the service of the designers or users being able to understand and engineer the agent) is for an agent to possess a token (e.g. a proposition Φ) in a "belief box". This is also illustrated by Fodor's exposition "Since according to RTM [representational theory of mind], concepts are symbols, they are presumed to satisfy a type/token relation; to say that two people share a concept (i.e. that they have literally the same concept) is thus to say that they have tokens of literally the same concept type." – (Fodor, 1998) pp. 28.

Shoham's commitment to realizing this framework excludes the full theorem-proving of his chosen modal logic – noting the problem of belief revision and assimilation is computationally intractable for propositional logics and undecidable for first order logics[3] (Shoham, 1993) pp. 77. Agents internal (mentalistic) state is simplified in such a way that assimilation and revision of facts is linear in time. This is achieved by disallowing any logical operators except negation.

In keeping with the sentential-symbolic tradition, Shoham (as (Rosenschein and Kaelbling, 1995) do with situated automata) defines the operational agent (e.g. machine-level, compiled working program) as separate from the method used to reason about the agent. This is the implementation independence of computational functionalism (Block, 1995) – the body, or implementation substrate of the agent, is independent of the agent's specification (of mental state in functionalism). This philosophical doctrine is the super-theory of Newell and Simon's physical symbol system hypothesis (Newell and Simon, 1976), which states that human behaviour (or intelligent action) will reveal a symbolic system, realized in the physical but fundamen-

[3]An intractable problem is one for which no polynomial-time algorithm can be found (Harel, 1992) and an undecidable theorem is one for which no effective procedure (in the Turing sense) can be found that proves its truthhood in a formal theory (Mendelson, 1987)

tally independent from its realisation. Newell and Simon close their seminal article with a conclusion that states that the rise in computational power will enable more elaborate implementations of formal symbol systems, and that these symbolic structures are the "raw material of thought". Newell and Simon's argument hinges on the philosophy of semantics derived from syntax, formal denotation and representation and finally, the virtual machine realised on the physical machine.

2.3.5 Bounded Rationality

Given results from computational complexity theory, we can be sure that certain in-principle algorithms for planning are intractable i.e. they cannot provide a result in a reasonable time. For example, first-principles planning with conditionals (e.g. re-planning when assumptions are not holding) have been shown to be intractable (Chapman, 1987).

At the other extreme are reactive agents which, we might say, can guarantee *a* result in a finite time. In between, are bounded rationality agents where, much like the heuristic search proposal of (Newell and Simon, 1976), there is a finite limit on the computation resources given over to any decision. Of course, the mathematical elegance of solutions can be compromised by this sufficing principle.

2.3.6 Optimality and Constraints

To conclude this section, the notion of optimality and its relationship to rationality need commentary. In the above descriptions, a definition of rationality was implicit. Namely, that an agent's purposeful behaviour that appears to meet some conjectured goals is rational. If one of the agent's goals is to survive, or not to run out of a certain resource, then this explicit goal will constrain the actions of the agent, making it behave in some rational way (after all, if it worked against the goal, it would not be considered rational). There is another implicit notion in rationality, especially in computational agency. If an agent is observably rational (i.e. works to achieve its specific goals) and these goals are in some way constraints on its behaviour (for example, survive or keep a certain resource level in equilibrium) then the agent will be behaving in an *optimal* fashion for a certain environment. Much work on reinforcement learning (Sutton and Barto, 1998) is built from the assumption that goal directed behaviour constrains an agent. Implicitly the combination of the goal, the resulting parameters that constrain behaviour and its observable success in achieving that goal defines its optimality.

For example, in instrumentally conditioning agents to behave in a certain way, they are asked to maximise the reward function:

$$R_t = \sum_{k=0}^{T} \gamma^k r_{t+k+1} \tag{2.2}$$

where R_t is the total reward measure (presented to the agent as a temporal sequence of real values in the range $[0,1]$) gained over a period $t \dots T$ with geometric discounting of future rewards r_{t+k+1} by a factor γ. Purposive behaviour is defined as maximizing the return.

Similar arguments pervade adaptationist thinking in evolutionary biology, where natural selection is taken to be a teleological optimizer of organisms by operating on populations of different genetic material. This optimizing conception of evolution has been challenged (see (Dennett, 1987; Dennett, 1996; Landauer and Bellman, 2000) for discussion. Instead, Dennett echoes the pluralist approach: that sufficiency in design is the best condition. Dennett gives the example of a dismasted boat, which ungracefully sails into port using a cobbled together rigging. As a best design for a boat it is weak, but given the goal (to safely reach port *given* the available materials and the dismasted, broken rigging) and the subsequent constraints, it is a *satisfactory* design.

Shoham's implementation of modal logic follows a similar sufficing principle; by restricting allowed operators and compositions of predicates. Effectively, for computational agents, the idea that an agent is sufficiently designed (adapted) to its environment so that goals and their constraints are satisfied is strong and plausible enough. It seems that mechanisms which cause behaviour might be best seen as necessary and sufficient for the task. Further ties to optimality might prove counter-beneficial. In later chapters, the ways in which adaptive mechanisms defined by optimally rational constraints will be explored and shown to be less than *generally* adaptive for long-term agent/environment situated activity.

2.4 Representation and Deliberation

In this section, we attempt to clarify the view of representation in deliberative paradigms. A simple example serves to illustrate the implicit philosophy and commitments. (Guarino and Giaretta, 1995) give an example of a simple blocks micro-world, consisting of a table and blocks which can be arranged in a variety of fashions.

An agent interacting and situated in this world would, according to the deliberative paradigm, require an explicit internal model of this world. Guarino and Giaretta show how this might be captured using the notion of *conceptualisations* which capture the essential properties in a formal model. For example:

$$\langle \{a,b,c,d,e\}, \{on, above, clear, table\} \rangle \tag{2.3}$$

This describes the world as a set of extensional denotations for the blocks (the domain of discourse) and a set of relations. The strict and important property that each block *in* the agent's world is *denoted* by a unique token in the model (e.g. a) is the property of systematic interpretability i.e. see (Harnad, 1992) and cf. (Agre, 1997). Therefore, if we specify:

$$\langle \{a,b,c,d,e,table\}, \{on, above\} \rangle \tag{2.4}$$

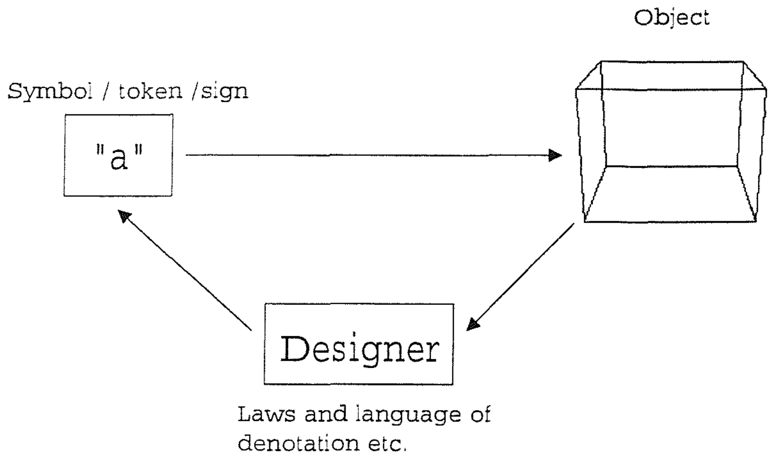

Figure 2.3: Symbol-Designer-Object Relationship

---

we are specifying a different conceptualisation of the world. In both cases, note that the choice of tokens (e.g. the table and blocks) and relations is arbitrary: the agent's designer will decide and model the world *for* the agent. Also, the meta-linguistic device of calling the table either an object or relation is what gives the model its interpretability. If we supplanted *on* or *above* with code (e.g. a binary representation of these collections of characters) and further we do not retain this code, then the tokens loose their intuitive meaning. This process of "codification" inscribes the designer's ideas into the "shape" of the tokens, signs or symbols.

Figures 2.3 shows the semiotic relationship between the symbol *a* and the block object, emphasizing the involvement of the modeller or designer. It is the modeller who possesses some formalised knowledge or representational schema which enables the production (semiosis) of a symbol or token *a* which according to the representational schema known to the modeller, will strictly individuate the object in the world.

Agre (1997) summarises the GOFAI tradition of formal mathematical modelling accompanied by tacit (and in his view, unquestioned) philosophy:

> Having formulated a mathematical semantics for one's new language, one may proceed to realize it inside a computer, probably using conventional data structures and pointers, and presumably building computational facilities that perform deduction within the new language.

More emphatically, Brooks describes this as partial "physics envy" (Brooks, 1999) pp.3 and the error of abstraction, leading to his well-known slogan "the world is its best representation" – a debunking of what was later refined to be *centralised* traditional world modelling as outlined above.

Fodor's influential Language of Thought (Fodor, 1979) has dominated discussion in cognitive science since Fodor's pronouncement that it was the only available contender for a cognitive theory of intentionality (see (Fodor, 1998) for a recent summary). Arguments which favour cognitive and interpretive systematicity (Fodor and Pylyshyn, 1988) as a nomological necessity seem to be often misconstrued as arguments *in favour* of an internal calculus of action and cognition. Beatrice de Gelder summarises the position by dividing Fodorian representational theory of mind and language of thought from sentential cognitivism:

> For example, on Fodor's view, to believe that p [where p is a propositional attitude], is to stand in a relation to a representation conceived of as a syntactically structured object, expressed in the language of thought. Sentential cognitivism promises to deliver the formula which will allow us to cash in our intuitive notion of a mental representation for that of an internal sentence in the language of thought. – (Gelder, 1996) pp. 152

For example, Fodor's thesis is that to have a notion of *something* requires possession of a representation of that *something*. A logico-linguistic formulation of this idea is the propositional and predicate calculi. Fundamentally, we can make an agents' epistemology explicit in symbolic computation and representation. In addition, we can capture even more essential properties of cognitive activity if we further the representational thesis into the domain of intentionality and doxastic state: beliefs, desires and intentions.

Computer science, then, seems to favour a Fodorian view which is in fact more akin to sentential cognitivism. In traditional AI, an agent is said to contain a representation p. By de Gelder's argument, this is sentential cognitivism and Fodor's position is more sophisticated. Fodor uses an English-like meta-language in a propositional framework, but does not argue for a realisation of that same language in the mind of the agent. Fodor seems to use this as a device for explicating what must be true of intentionality in principle.

Finally, in summarising the reactive approach and emphasising its distance from GOFAI Brooks (1997) states:

> Gone are representations of beliefs, desires and intentions. The very notion of a speech act no longer makes sense. And gone were the grand difficulties that AI artificially introduced by mistakenly trying to reduce an essentially external descriptive language to an internal calculus of reason at run-time.

2.5 Subjective vs. Objective Description

The description of the blocks micro-world given above is described as a conceptualisation; one of many different formal models of an ontology, or "what there is" in some objective reality.

Note that a conceptualisation arises only through the intervention and semiosis of the designer who attempts to inscribe their model of the world in the artefact they create.

Heidegger's phenomenological reconstruction of metaphysical ontology (i.e. to bring accounts of *what there is* into a context that acknowledges its dependence on the thing conducting the enquiry) states clearly that:

> All ontology, no matter how rich and tightly knit a system of categories it has at its disposal, remains fundamentally blind and perverts its most proper intent if it has not previously clarified the meaning of Being sufficiently and grasped this clarification as its fundamental task – Heidegger, Introduction to *Being and Time*; translation in (Krell, 1996) pp. 53.

Heidegger's claim is that all ontological study must fundamentally acknowledge the subjective and fundamental role of the agent at the centre of the analysis. Effectively, we can see Heidegger's writing on ontology as a rejection of an objective, total account of "the agent and environment". Further, this clarifies what a conceptualisation really is: it is one possible formal model devised by the agent's designer. It is a conceptualisation and ,therefore, a partial "capturing" of an ontology only in as much as the symbols stand in a relationship *via the designer*. This revisits systematic interpretability and, likewise, meshes with Newell and Simon's physical symbol system hypothesis; of which one constraint on symbols is that they denote "things" in the designer's conception of the world.

In keeping with the semiotic analysis given above, any symbol system is then a conceptualisation (not really an ontology) and one of *many possible* conceptualisations capable of being generated by an agent's designer. There must be at least three things in the interpretation of a conceptualisation: (i) the symbol, (ii) the designer (we might say the owner of the denotative law) (iii) an object.

Why is it necessary to speak of an agent's own "internal" or phenomenological conceptualisation? Superficially it is a peculiar philosophical analysis but, as will be argued in depth later, the self-organisation of an agent's internal mechanisms requires such an interpretation if we are to avoid infinite regress on the nature of representation. Additionally, the conceptualisations (and objective ontology) approach might fundamentally miss the important factors which are key to an agent successfully perceiving, acting and communicating in complex domains. This proposal about the failure of GOFAI to exhibit working systems in everyday domains because it ignores the nature of everyday experience and activity is at the root of critiques of AI, including (Dreyfus, 1992), (Winograd and Flores, 1986) and (Agre, 1995; Agre and Chapman, 1987).

Although Brooks was fundamentally driven by engineering concerns, his methodology implicitly incorporates the notion that "the representation is in the mind of the observer" and discounts the need to build representations to facilitate deliberative reasoning over them. In his view, all disembodied knowledge representation projects will ultimately find themselves in a regress to the point of embodiment – the only fundamental level of "knowledge" representation being that necessary and functionally sufficient to enable situated activity (Brooks, 1991a) –

reprinted with commentary in (Brooks, 1999).

The most pertinent feature of Brooks' anti-knowledge representation proposal is that he directly cites Lenat's attempt to build an ontology (read: conceptualisation). The CYC project Lenat (1995) is truly an attempt at capturing a classic ontology (e.g. an account of what there is) in a conceptualisation, attempting to describe the world objectively without reference to phenomenal aspects of agent perception and sensing of the world. Brooks' argument in part reiterates the work of Dreyfus, summarised in (Dreyfus, 1992). It is worthy of note that these criticisms of traditional knowledge representation-focused AI might be applied (with due caution) to the emerging trends of computational agents as conceived in mainstream computer science, supplicating a defense of agents as something fundamentally new in AI. In terms of the proposal that agents are *not* traditional AI, it is interesting to note the engineering focus on objective ontology, see for example the overview of (Huhns and Singh, 1997).

2.6 Propositional Attitudes and Belief in Purposeful Behaviour

In preceding sections, the characteristics of deliberative and reactive paradigms for computational agents have been outlined. In the case of deliberative systems of any sort, they are in some way constrained by world models, ontologies and knowledge representation, accompanied by some scheme for reasoning over such models. In section 2.3 the "kinds" of rationality which are often embedded in such agents were discussed. We now take up the question of how Dennett's intentional stance has been introduced into discussion of rational agents.

The definition of intentional systems is due to (Dennett, 1978; Dennett, 1987) where he introduces the term in an essay on philosophy of mind. In effect, Dennett proposes that we ascribe mental characteristics to agents which are justified when they help explain the behaviour of the agents.

In computational terms, Wooldridge and Jennings offer John McCarthy's justification for this philosophy, who states in (McCarthy, 1978) that the ascription of mentalistic notions to machines is justified when:

> "... the ascription helps us understand the structure of the machine, its past or future behaviour, or how to repair or improve it"

Dennett proposes that when we fail to understand a system (specifically, human or animal behaviour) from either the *design* or *physical* stance, we (as humans with minds) resort to the intentional stance: that is, we ascribe the system properties we are familiar with that enable us to assimilate the system's behaviour into our current understanding[4]. The intentional stance provides us with an abstraction on a system's behaviour which is congruent with the

[4]The anthropomorphism of complex inanimate objects (such as a user stating "The computer won't do X") demonstrates our everyday use of folk-psychological ascriptions to enable our understanding without resorting to a deeper knowledge of the mechanism (which, of course, we may not have access to).

most familiar folk psychology humans use to describe other humans' behaviour. In addition, we ascribe similar folk psychological properties to devices or systems we do not understand explicitly, but whose behaviour we are exposed to.

A simple illustration reveals why this has utility as a tool for understanding systems with apparently impenetrable behaviour. With reference to:

- design stance: if we have a design description (completely) for a von Neumann digital computer (a device or system) then we could predict some behaviour (say, failure of the system) by following the precise sequence of activity in the micro-circuitry. This is not a feasible proposition, despite being an *in principle* complete description. We might argue that this is the wrong level of abstraction and that despite its reductionist appeal, is impractical for all but the most trivial of systems.

- physical stance: if we have no description of the design, but can see the responses of the system in terms of physical state (that is, environmental conditions and their observable effect on the system) – we can subsequently predict its behaviour. This is an extensional view of a system. However, experience reveals that rarely (in the computing example given above) do we have access or understand fully the effects of system state.

- intentional stance: the tactic employed by an agent when the design and physical stance are unable to empower us with an understanding or predictive success in describing the system's behaviour.

From this, we should ask when it is *appropriate* to employ the intentional stance. In effect, intentional stance is an *in practice* argument which does not deny the *in principle* causal mechanism that might be responsible for the behaviour exhibited. As the anti-reductionist movement have stated before, e.g. see (Fodor, 1980), there are legitimately studies of systems at different levels of explanation cf. (Marr, 1982; Arbib, 1995b; Arbib, 1989). Stuart Watt's psychological agency (Watt, 1996) utilises the intentional stance and adds a further level (the psychological level) that provides agents with properties that usefully exploit the notions inherent in the strong definition of Wooldridge and Jennings. In conclusion, he argues that his approach is:

> "... fundamentally quite radical. Instead of building progressively more complex agents, I propose building progressively more human agents ... we must remember that we are human, and begin to design more and more human agents."– (Watt, 1996) pp.96.

Watt argues that an agent might have been constructed with the intentional stance in mind and he assumes the intentional stance in providing another layer (the psychological) which exploits properties of human behaviour. He does, however, argue for agents that are ostensibly interacting with humans. Dennett's use of the intentional stance is interpretational, not necessarily for the *construction* of a system. Watt (1996) proposed that artefacts (or agents) should possess behavioural characteristics like extant agents which use (even if the evidence is anecdotal) intentional stance. Dennett's intentional stance is an *interpretational* theory of mind. The inherent philosophy of the GOFAI programme is "representational" e.g. (Fodor, 1998) and hence "realist" in the sense that to hold a belief you must possess a representation of the belief. To interpret an object and ascribe it agency, belief, desires and the range of propositional attitudes is something that serves our interests and is instrumental to us understanding behaviour. Such ascription makes no solid claim on possession or even the existence (as entities) of such propositional attitudes. In summary, to adopt the intentional stance is, to a certain degree, a rebuttal that representation of the propositional attitudes is real. It is then, from a purely philosophical perspective, a mistake to argue for intentional stance and the representational thesis of GOFAI. Dennett is not espousing a *formal representational intentionality.*

This view has been widely contested; that intelligence (or perhaps more precisely, cognition and purposeful action) is *not just* computation as championed by and summarised in (Dreyfus, 1992), in the Chinese room argument of (Searle, 1980) and Harnad's notion of necessarily hybrid systems (symbol grounding) and "total Turing tests" (Harnad, 1990; Harnad, 1992).

In part, then, this shows that even if the defining argument for situated agents (Jennings, Sycara and Wooldridge, 1998) is present, and perfect rationality gives way to practical reason, the underlying philosophical commitment is to the internal calculus of action and denotative mental state cf. Fodor's language of thought and representational theory of mind.

The dividing line between contemporary agent theory and GOFAI is not in the methods and assumptions inherited (i.e. symbolic representation, problem solving and purposeful action by theorem proving and so on) but that same ideology is bound by situatedness and constraints on rationality.

2.7 Computational Agency from the Reactive Perspective

At this juncture, it is sensible to ask what a reactive (or more properly, non-sentential symbolic) agent theory might look like. At least, what properties an agent theory would have to possess in order to be considered something more than a simple control system and less like a representation-based deliberative agent.

2.7.1 Appropriate Substrate

When speaking of computational agents from a deliberative or symbolic-AI approach, we have an array of mechanisms which form the foundations of the agents implemented. For example, predicate logic (for knowledge) and modal logic (for belief) all carry theorem proving and resource-bounded mechanisms for rational action (Russell and Subramanian, 1995; Shoham, 1993).

Reactive, non-sentential dependent agents are a relatively recent addition to the field of agents (e.g. Brooks' papers from 1986 onwards are taken to be the resurgence of new-AI). (Arkin, 1998) Chapter 4 provides an interesting collection of architectural devices used in reactive, behaviour-based systems. In this thesis, connectionism is evaluated with respect to providing a substrate for a theory of agency.

2.7.2 Principled Representation

If a reactive agent does not utilise the arsenal of symbolic methods for representing beliefs, knowledge and so on, then what does it use? Are there any principled tools that help explain this apparent gulf?

It will be suggested that an appropriate unification can be found in a fuller theory of intentionality (Dennett, 1987) and sub-symbolism (Smolensky, 1988) when situated activity is considered as the driver forcing representation to become relevant. Essentially, we take a functionalist perspective on representation and engage recent work on deictic and indexical representations on sub-symbolic substrates. These vital contributions shift focus away from processing (say) symbolic structures to the interaction between the agent, environment and the genus of representations which are compatible. Key to this enterprise is the notion of deictic representations see (Agre, 1997) and interactionism. The former is the idea of the instantaneous relationship between something *inside* the agent and the environment; similar to intentionality, but less strong for example, deictic representations have little say on the original or derived semantics (the word *deictic* meaning to point out, individuate or show).

The notion of interactionism helps move away from the idea that the agent is the creator, maintainer and effector of structure in action. For example, in the GOFAI tradition, the world is modelled internally. This internalised world model is reasoned over and the actions are decided upon and effected. The world then changes and the agent must redress the disparity between *its* internal world (of cogitation and deliberation) and the external world. Interactionism relocates both the representational burden (e.g. things which factor into purposeful action are *not just* maintained as internal representations) and focuses the environment as being reciprocal in supporting purposeful action. From this, the notion of routine activity arises. A routine activity is one which repeats itself (to an observer) and is functionally dependent on the environment as well as the agent. A more philosophical treatment is given later in Chapter 5.

2.7.3 Goal-Directed Behaviour

It is somewhat difficult to locate goal directedness in the behaviour-based paradigm, let alone within a connectionist system. For example, the augmented finite state machines in Brooks' subsumption architecture each contribute to the overall observable goal directed or purposeful behaviour. For this reason, it is necessary to tease apart the structures that support this observable behaviour in connectionist agents. Since much of the work documented here uses reinforcement learning principles, the nature of goal direction is approached from the bottom up (e.g. networks which exhibit reinforcement-like learning effects) and then meshed with the machine learning technique of reinforcement learning, which is finally implemented on a connectionist substrate.

2.8 Towards a Unifying Theory of Agency

To conclude this chapter, a justification of this thesis' approach is necessary. Essentially, we are concerned with agent theory – and specifically, the epistemic and behavioural questions about situated activity in an agent. The approach taken will be to explore sub-symbolic or connectionist approaches. Such an approach is necessarily inspired by bio-behavioural models at multiple levels of explanation.

In their edited volume, Huhns and Singh (1998) present an introductory discussion about *what* computational agency is about and what fundamentally drives the research programme. They conclude that the study of agency is concerned with (largely) computational artefacts which possess properties usually attributed to other extant agents (e.g. animals and humans). Under this umbrella definition there are a wealth of disciplines which contribute, including: robotics, ethology, philosophy of mind, cognitive science and related disciplines. The study of computational agency is the unifying umbrella under which all these contributions fall. Casting such a wide net imports the possibility of grossly over-generalising. For example, is a robotic agent subject to the same design constraints as a small mammal, a human, or an agent in a market-economy? If such a 'great unifying theory' (GUT) of agents exists, then it might reasonably be defined, as other workers have, as any complex adaptive system cf. (Holland, 1975; Holland, 1995).

Assuming that such a mature GUT exists, then it must account for all facets of agent-like phenomena. In such a theory, a robot and a disembodied virtual agent must be equally accounted for. The perspective taken in this thesis is that the narrowly conceived models of computational agency which currently exist are not wholly adequate to scale-up in any way to a broader theory. Likewise, if a software agent is truly embedded in its (artificial) environment, then we must surely consider the following case: the theorem-proving "internal calculi of actions", e.g. (Fikes and Nilsson, 1971), approach has proved at least brittle in robots situated (and embedded) in the physical world. Therefore, as a general theory of agency, this may be a weak perspective. If the difference between an expert system and a disembodied virtual/software agent is the fact that it is embedded, situated and interacting with its environment, then we might predict that an internal calculi of actions might not be applicable to truly situated software agents either.

Accordingly, the goal of a GUT for agency must include crossover between the physical and virtual agent and both must be treated as situated agents in their respective physical/virtual environments. Also, it must account for the kinds of interactions which shape behaviour beyond the internalising of world models. While little will be said in conclusion to this question, it is hoped that this thesis offers a perspective on computational agency.

2.8.1 Conclusion and Justification

The implications of Wilson and Churchland's theses are congruent with revisionist thinking on agent-environment models. The fully integrative perspective of an agent and its environment (in fact, so integrative that the discourse structure "agent" and "environment" can be said to be artificial) is more congruent with the *structural coupling* theory of Maturana and Varela (1980a) than the conventional organisation of behaviour models of (Simon, 1985).

In terms of research culture, Newell and Simon's work in AI encapsulates the classical GOFAI assumptions in theory and practice (Newell and Simon, 1972). Simon's earlier work on organisational behaviour and structure is rehearsed in Chapter Seven of (Simon, 1985). One of his characterisations of the environment is *hierarchical nearly-decomposable systems* (e.g. as in government, business and commerce). The argument to support this observed phenomena of environment structure is that it is governed by the limited rationality of its agents. For example, an agent can only perform a limited amount of work. Workload, responsibility and job descriptions therefore define a hierarchy which divide the overall throughput of an organisation into responsible subsystems with the atomic unit of analysis being the agent. In this sense, we have a heterogeneous hierarchical organisation and corresponding description or ontology of agent and environment. The short-run dynamics of the inter-agent interactions are independent, but the global, long-run behaviour depends on these interactions. Essentially, this represents a functional decomposition of a collective, multi-agent system where the recursive decomposition of organisational task and subtasks halts when the atomic unit of the agent is encountered.

It is appropriate that this kind of environment/agent distinction works well for multi-agent systems which simulate or synthesise work-flow management systems. In terms of heritage, the environment-agent definition of Simon's earlier work (Simon, 1947) has been carried over into the (totally congruent) functional decomposition of competence approach of GOFAI. Computational agent theory has also inherited this environment distinction, unsurprisingly, given the emphasis on abstraction, decomposition and modularity in computer science.

The question posed here is whether or not this is the only way to characterise the agent and the environment. This thesis is in agreement with Wilson's bottom-up approach. Descriptions of the agent and environment in Simon's terms might be premature, since they are concerned with an atomic unit of calculating procedural bounded rationality and the intellectual lineage and practices which beset this approach.

In addition, Simon's characterisation is predicated heavily on the functional extension of the organisation (multi-agent system) and how an agent fits into this observed organisation. The environment/agent distinction (indeed, the very boundary between the two) is defined in service of the organisation's behavioural "output". We might say that only when the system has been subjected to analysis and decomposition is the agent/environment boundary defined, and then it is predicated on the mode of analysis and subsequent decomposition.

To summarise, characterising agents by their contribution to nearly decomposable systems is akin to looking at the agent and environment through a filter where the atomic unit of analysis is the agent. In this work, the agent and the environment will be built bottom up, aiming for a holistic interpretation and model of the agent/environment interaction.

Fundamentally, ecological theories of perception such as Gibson's *affordances* (Gibson, 1979) and Maturana and Varela's *structural coupling*, remind us of the closely woven fabric of agent design and the environment, such that the design enables survival. Gibson's anti-constructivist view proposed that the environment contains enough information to support visual perception cf. radical reactive agency and its dependence on information carried in the immediate sensory state (Jennings, Sycara and Wooldridge, 1998). Perhaps as each natural organism occupies an ecological and evolutionary niche, then similarly a truly situated artificial agent needs to similarly be engineered to exploit its environment. A generalised attempt to capture the nature of *all* environments and agents by one principle (for example, if we chose Simon's nearly-decomposable approach) would result in an agent which was not so much situated in its actual environment, but one which generalises the environment and the agent's behaviour to mesh the two at a more abstract level.

What motivates the inclusion of mechanisms constrained by cognitive and biological principles? Wilson's celebrated "animat approach" (Wilson, 1991) contends that a fundamental driver for any synthesised behaviour in an artefact should be similar to the driver in nature – survival. He suggests that fully integrating simple but holistic models of "animal robots" (or *animats*) in their environments enables the research community to progress "bottom-up" with synthesising behaviour. (Dennett, 1987) pp. 257 defends a similar hypothesis; that AI might progress by ignoring human micro-competence and exploring holistic behaviour in lower-order animals. For example, (Meyer, 1997) reviews some candidate mechanisms inspired by biology that exploit environmental features to serve the agent's design goals (e.g. to survive in a hostile environment). In concluding his comparative analysis, he concludes that more fundamental understanding of low level mechanisms of biological agent behaviour mechanisms need to be incorporated to fully imbue the engineered animats with the necessary adaptive machinery to ensure long-term performance in environments.

Wilson's approach is potentially one of many possible alternatives to the competence-based, functional decomposition of traditional (and symbolic) AI. He begins by characterising environments and then proposes vertical levels of complexity. For each level, a potentially different agent architecture might be needed to cope with the environment's complexity. The design strategy for an animat is to then exploit the characteristics of the environment in the agent's design (almost as neo-Darwinian selection is proposed to do).

How can we justify research effort in this direction? A philosophical answer has been offered in cognitive science and philosophy of mind. Paul Churchland offered some direction in (Churchland, 1979) with his proposal of scientific realism. Later work by Patricia Churchland (Churchland, 1986) embraces this principle and seeks to identify how mind/brain might be seen under a unified theory by mutual co-illumination of philosophy of mind and mature neuroscience.

A superficial interpretation of the Churchlands approach would be that they simply identify mind and brain by means of physical identity. This was attempted by (Smart, 1959) and forms the identity thesis; all mental states *are* merely physical states. Thinking a certain thought is identical to the agent's brain (or part of the brain) being in a certain physico-chemical state. Difficulties in demonstrating this hypothesis resulted in the reductionist theory that mental state might reduce to (and therefore be replaced by) a physico-chemical explanation at a lower level. This is also hard to justify. The process of theory reduction proceeds by trying to define a suitably broad physical state that co-occurs *every time* with a given mental state (see (Kim, 1998) for a fuller discussion). It transpires that examples abound of identifiable mental states which occur with such broad categories of physical state that the potential to discretise the physical states into some kind of finite set is impractical[5]. This so-called multiple realisability of mental states (that one mental state might co-occur, or be realised, by many different physical states) has been widely held to defeat the reductionist approach.

Such reductive philosophy on mind and body has undergone revision. For example, Kim's attempt to preserve reductionism in the light of heterogeneous disjunctions of physical states (Block, 1997), which holds that a valid definition of a physical state could be a disjunction of realisers (physical states).

Patricia Churchland's tack is different. She claims that pragmatic co-evolution of parallel theories at multiple levels of explanation is the most fruitful and progressive approach (Churchland, 1986) pp. 284. In effect, continuous evolution of psychological and neuroscientific theories (taken to be high and low level explanation respectively) that are joined by explanations at appropriate junctures, i.e. when both theories offer a mutual point of contact.

The reason for reviewing the Churchlands' perspective is that the unifying theme is the removal of folk-psychology (and intentional stance) from the problem and its research agenda. To establish a mature theory of behaviour, action, intention and cognition generally requires that we abandon folk-psychological explanation because:

1. the Churchlands' propose that folk-psychology is purely derivative; it is a pragmatic but not analytical mechanism for explaining and predicting human behaviour and cognitive functioning (Dennett, 1987; Dennett, 1978).

2. the Churchlands' argue that mind, purposeful behaviour and agency are not unique to

[5]Such individual mental states are questionable, since reporting a certain state requires the agent *itself* to do the reporting, opening the way to arguments about the introspective process involved and the validity of the resulting report on the mental state

humans – animals are also agents – but folk-psychology is uniquely human and cannot be generalised over all animals

3. therefore, the only way to proceed is to find and explain mechanisms and phenomena which are similar in both animals and humans (Churchland, 1986; Dautenhahn, 1997) – they conclude that neuroscience is the only common denominator in natural exemplars of intelligent, purposeful behaviour. All animals have nervous systems that adapt and learn in similar ways, and so a constructive approach is to study these mechanisms and the behavioural phenomena which emerge.

Patricia Churchland's claim can be found in the way similar folk-explanatory theories stand in relation to the "real" theories or sciences. According to Churchland, folk-psychological notions (which for precision we might limit to propositional attitudes such as "*x believes* that *p*") stand as everyday, common-sense explanatory mechanisms because of the complexity of understanding the fundamental causal mechanisms beneath, or responsible for generating, the observed behaviour. This is compatible with the intentional stance (Dennett, 1987).

Similarly, (Hayes, 1992) describes an approach to common-sense reasoning about physics in AI. His "naive physics manifesto" describes some basic *concept clusters* which must be present and realised in an algorithm that enables an agent to act in its world. For example, he defines a notion of "support" which describes how an object might be physically supported by another object if it is situated on top. Providing an agent with this common-sense intuition (readily observable in everyday practice and experience) engenders the agent with a naive understanding of some very complex physical science which explains precisely why this supporting relationship holds. Similarly, infants acquire such experience, and arguably then possess knowledge of support, but know nothing of the true Newtonian mechanics which more completely (and reducibly) describe the relationship. What Hayes has done is to carve the world into necessary but sufficient tacit principles (the concept clusters) and has thus created a folk-physics. However, any agent possessing Hayes' naive knowledge would invariably be polluted by Hayes' own embodiment, sensory experiences and these would not necessarily mesh with similar systems for an artificial agent. We therefore find an instance of non-transferable, agent-specific (in Hayes' case, human specific) ontological commitments and common-sense folk-explanatory mechanisms. No one can say for sure that a rodent uses similar notions of naive physics any more than a rodent uses folk-psychology to explain the behaviour of its kin. Hayes' manifesto is to physical science what Churchland claims folk-psychology is to cognitive science. Correspondingly, despite the pragmatic appeal of the concept cluster "supports" and the doxastic state "believes", the Churchlands have argued that they are scientifically moot for the study of both physics and cognitive science. We note that Hayes' proposal was not intended as a model for studying physics, but in fact representing physics to an agent so it might utilise this in its reasoning about world.

From this, we can simply establish the link to this work as follows: Much of agent theory is concerned with representing (in the GOFAI sense) doxastic and epistemic state – for example, using Kripke's modal logic (Kripke, 1963) to express possible-world semantics and identifying agent's internal state with such calculi of intention. An agent might formally possess the belief that it is safe to open a door. Doxastic state is defined firstly as a propositional attitude, then expressed and represented as agent internal state using a modal logic. Therefore, despite the qualifying notion of situatedness, agent theory requires an account of situatedness from first principles cf. the work of (Agre, 1997; Agre, 1988) and interactionism. However, connectionism has received much explicit attention as a proposal for implementing such a theory.

Chapter 3

An Agent Perspective on Connectionist Control

One goal for this thesis is to explore the virtual/embodied boundary and explore the application of the emerging definitions of an agent to both sides of the division. This was argued as necessary for a true unified theory of computational agency. There is also an intuitive appeal to an agent which can learn to act appropriately given very little or no description of its world. Connectionist models provide a structural and dynamic base for many kinds of learning; from simple associative learning through to classical and operant conditioning.

This chapter gives examples of agents (mainly from the robotic domain) which use neural networks explicitly to implement behaviour. The techniques used in experimental work are then described from the perspective of agent perception.

The chapter concludes by summarising the key observations which are carried forward as key design criteria in the subsequent chapters which develop the agent model.

3.1 Introduction

Most robotic agents directly employing connectionist internal mechanisms do so using derivatives of well-known network architectures. Usually, they are used to learn stimulus-response behaviour when no model for the behaviour exists. The resulting internal mechanism is taken to be a kind of "holistic" controller. The focus on robotic agents is necessary simply because most adaptive behaviour work (of the kind that emphasis behaviour learning from experience in the environment) has been conducted in either software simulations of robots, or on robot hardware.

In disembodied, virtual or software agents discourse, discrete actions are more common than real-valued actuators signals. In the next chapter, we will show how the pattern-recognition basis of the models discussed, such as in the work of Tyrell (1993), relates to a fully parallel model of action selection akin to the Pandemonium model of (Selfridge, 1959) see (Remez, 1987) and (Arbib, 1989) pp. 172. (Humphrys, 1997) indicates that Pandemonium models (pattern-recognition driven) and action selection models have not been explored with respect to functional equivalence. The treatment of an integrated connectionist/agent model as essentially classification via sensori-motor co-ordination has been given by (Scheier and Pfeifer, 1995). The approach here is similar, but given that perception and action are implemented by the agent, the systematic relationships between perception and action are teased apart in a way more akin to an agent theoretic approach (e.g. as interacting sub-systems).

Similarly to Wooldridge (1999) and Rosenschein and Kaelbling (1995), an automata-theoretic model is adopted. However, Rosenschein, Kaelbling and Wooldridge do not treat adaptive behaviour explicitly. Their corresponding theories of representation are symbolic (model theoretic or possible worlds). The strong condition that internal state (as symbols) must correlate with observable behaviour is challenged in Chapter 5. Correlational definitions of internal state content are ascriptional cf. (McCarthy, 1978; Dennett, 1978; Dennett, 1987). However, this strong epistemic assumption can be weakened by drawing on the formulation of behaviour-functional models (Arkin, 1998) and using the concept of heterogeneous disjunctions of internal state as the agent component of correlation cf. (Kim, 1998). These arguments and the recovery of situated automata theory for connectionist-based agents is given in (Joyce, 2001b). For convenience, this report is reproduced in Volume II of this thesis. Generally, the automata model will be used as an abstract agent architecture throughout this thesis.

3.2 Connectionist Networks and Robotic Agents

There is a vast literature on the use of neural network models such as the multi-layer feed-forward networks (MLFNNs) based on multi-layer perceptrons (MLPs) (Rumelhart and McClelland, 1986). Most involve some kind of connection of sensors to actuators in either a direct or indirect configuration. Some models have used MLPs with one layer of hidden nodes acting as intermediate processing units between sensing and actuation, whereas some, notably (Braitenberg, 1984), use direct connections from sensors to actuators without such intermediate processing.

Assuming some number of sensors, the output of the sensors is the input of the artificial neural network (ANN). This then "processes" the sensory input, usually yielding real values on the outputs which are then taken to be the "drive" strength of the actuators. This is summarised in Figure 3.1.

One class of models utilises the ubiquitous MLP as a generic model of supervised learning. In such a model, the agent must be provided with both sensory input (stimuli) and a

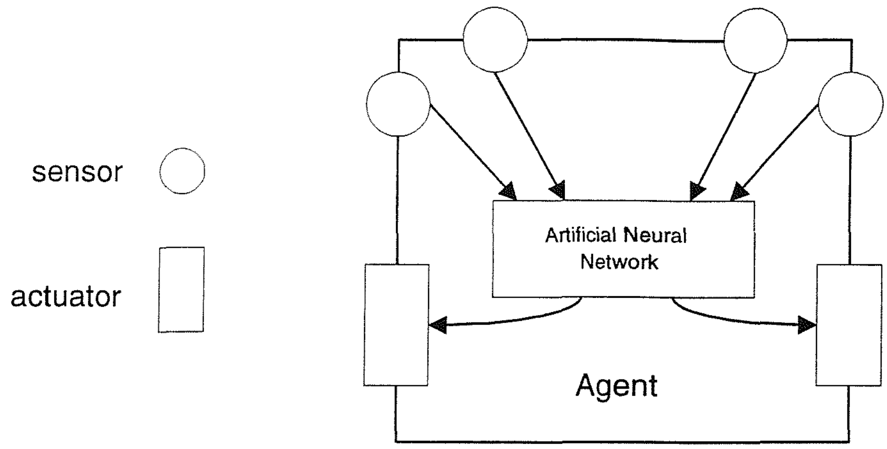

Figure 3.1: Schematic Illustrating the Role of ANNs in Robotic Agent Behaviour

corresponding "correct response" indication. The processing of sensory information through to actuation is then simply a series of numerical computations loosely inspired by "brain-like" models.

3.2.1 Implications for Agent Behaviour

If sensory information is represented as vectors, then these can be taken as realisations of the random vector $\mathbf{x}$ and multiple realisations form an input space. The output passed to an actuator using a single layer network is determined by a summation of the inputs multiplied by the weights connecting sensors to actuators:

$$a_j = \left(\sum_{i=1}^{M} w_{ij} S_i(a_i) \right) + w_{0j} \tag{3.1}$$

where a_j and a_i are, respectively, the activation of the output (actuator) node j and input (sensor) node $i \in \{1 \dots M\}$, w_{ij} is the weight leading from i to j, S_i is the signal function (mapping the potentially unbounded activation a_j in the range $[0,1]$) and w_{0j} is the *bias* node which is the metaphorical activation threshold above which the node "fires".

Such networks implement discriminant functions and form decision hyper-planes in the input space, where the bias or threshold w_{0j} is the offset of the hyper-plane from the origin of the feature space. So for a two actuator agent (i.e. one motor each for left and right wheels respectively) using a stimulus-response theory of behaviour for illustration, the possible behaviour is restricted by the hyper-plane $O_1(\mathbf{x}) = O_2(\mathbf{x})$. For a given stimulus, if $O_1 > O_2$ then the actuator for O_1 will activate, otherwise O_2 will respond.

It is easy to see how this restricts behaviour. If we assume that the stimuli fall into clusters in the input space, and the means of these clusters are in a XOR-type formation then we have a non-linearly separable problem. If each of these clusters represent stimuli for which different responses should be generated, then as Minsky and Papert demonstrated, there is not a single hyper-plane which separates them, thus disabling the ability of the agent to produce different behaviours for each case represented by the stimuli clusters. To summarise, the previous discussion describes the agent's perception and part of the action execution pathway. Output is determined as a function of the input activities and the weights. Hence the weights realise a set of possible mappings of input to outputs or behaviours cf. (Arkin, 1998).

3.3 Examples of Related Work

A number of workers (Nehmzow and McConigle, 1994; Braitenberg, 1984; Scheier, Pfeifer and Kunyioshi, 1998) have used the single-layer networks with McCulloch-Pitts type neuronal dynamics to control robotic agents under a suitable training regime. More generally, direct connections from sensors to actuators have been implemented using neural networks similar to those above. The structure of a single layer perceptron was used by (Nehmzow and McConigle, 1994) to implement stimulus generalisation which then fed a behaviour generator causing certain behaviours to be produced. The network architecture is used as a simple pattern associator, and the learning algorithm used is a supervised method based on the Widrow-Hoff method (Widrow and Hoff, 1960) of minimising the error on the nodes outputs with respect to the desired output.

3.3.1 Braitenberg Vehicles and "Parsimonious" Networks

Braitenberg's seminal work on what he called "synthetic psychology" (Braitenberg, 1984) demonstrated that simple neural networks can be made to exhibit complex behavioural patterns. Similar direct approaches include the so-called "five neuron trick" (Scutt and Damper, 1997) where the neuronal model is significantly more complex and intended to model classical conditioning at the level of neuronal spike activity. It is perhaps premature to include them in a comparative survey of techniques focusing on McCulloch-Pitts based neuronal models and their continuous signal function counterparts.

Similar work on Braitenberg vehicles, that use the familiar formulation of single-layer networks above, was conducted by (Salomon, 1997). The architecture is extremely simple, and can be expressed as a single layer network with two output nodes. However, Salomon claims that training of the weights is difficult because a training set cannot be determined and then used to train using "standard neural network training procedures" pp. 205. However, while the network's parsimony is elegant, it fails to capture certain architectural characteristics that other workers have, such as the use of novel training procedures incorporating reinforcement and the use of unconditioned stimuli which inform the agent of harmful or noxious environmental conditions. Salomon chooses to evolve the weights using a simulated evolution approach.

3.3.2 Neural Networks and Explicit Adaptive Behaviour

Scheier, Pfeifer and Kunyioshi (1998) use a simple single layer network to control a mobile robot, exploiting specific morphological features of the agent's embodiment to the advantage of the control system. They provide some evidence that the XOR problem is not relevant to situated robotic agency, citing experiments with primates failing to learn XOR-like relationships. A similar hypothesis has been explored by (Clark and Thornton, 1997) where they describe learning problems including XOR and the relationship with neural networks. One commentary on Clark and Thornton's article (Damper, 1997) proposes that XOR is not an issue because it is not a test for generalisation; that is, the network would have to "read minds" in order that it correctly predict (by generalising from other samples of the XOR problem space) the non-linearly separable cases. The fact that primate behaviour seems to be unable to cope with XOR-like problems, and Damper's suggestion that XOR isn't a generalisation problem anyway suggest that non-linear separability is not a major challenge to engineering intelligent agents. It could be concluded that in certain instances, multi-layer systems may present an over-engineered solution to many sensori-motor co-ordination tasks.

Similar work was conducted in a unified adaptive behaviour framework by (Verschure and Voegtlin, 1998) see also (Pfeifer and Verschure, 1997). They adopted principles of localised Hebbian learning, namely that adaptation should be a neuronally-localised phenomenon and layers of abstraction in the control architecture should be explicitly accounted for in the adaptive network architecture. Their DAC (distributed adaptive control) architecture has, at its core, an adaptive neural network partitioned into elements reflecting perception and action. Wrapped around the perception-action network is a layer which controls adaptive behaviour, and finally a higher-order layer co-ordinates the whole system in a vertical manner. Most significantly, they define a value system which indicates (to the adaptive layer) what is significant to the agent. In this thesis, both single-layer, self-organising unsupervised techniques *and* supervised multi-layer feed-forward networks will be examined.

Likewise (Doya, 1999b) proposed that different learning methods (e.g. unsupervised competitive / correlation, supervised and reinforcement) and architectures suit different functional components of purposeful action. (Nigrin, 1994a) (see (Nigrin, 1994b) for a summary) derived criteria which are necessary for an autonomous agent based on connectionist principles:

- The network must self-organise using unsupervised learning: an unsupervised system can be embedded in a framework for supervised learning, but the converse (according to Nigrin) is not possible. The ARTMAP networks of (Carpenter, Grossberg and Reynolds, 1991) illustrate this.

- The network must learn stable category codes: important existing categories should not be destroyed in the service of creating categories for novel stimuli

- A network should learn associations quickly if necessary (Nigrin states that "one should not have to touch a hot stove 500 times before learning one will be burnt"). This should be balanced with graceful long-term learning to allow generalisation.

- A network should use feedback and expectations to aid categorisation. For example, if previous attempts at successful action proved harmful to the agent, then some mechanism should reflect this on subsequent trials.
- The network should "unlearn" redundant categories

Although Nigrin lists 13 criteria, his work on SONNET as a general classification system for an autonomous agent leaves many of these untouched. Those presented above are the most relevant to this project.

Nigrin's insights are useful, but his claim for self-organisation as unsupervised learning ignores the relationship between situatedness and the agent's connectionist implementation. If a supervised learning technique is used (for example, back-propagation of errors with a MLP) to control an agent, then it is the source of training signal that defines its status as a supervised. Agency and embeddedness makes the notion of supervision more opaque than disembodied connectionism. The back-propagation method for changing weights uses gradient information to modify the final output of a layer of post-synaptic nodes. In principle, then, it is the source of the gradient information (i.e. the origin of the error relayed to the network) that makes the network supervised or unsupervised in the way Nigrin uses the term.

The notion of sensory classification and subsequent action selection is dealt with in (Scheier and Pfeifer, 1995). We are interested in this because it moves the treatment of neural networks closer to notions dealt with explicitly in agent theory. The hypothesis of Scheier and Pfeifer's work was that both action selection (picking the correct behaviour primitive such as "follow wall") and sensory classification are implicit in the neural network, sensors and the agent's interaction with the environment. No explicit internal mechanism or behaviour module needs to be constructed to cope with each observed, or in our case desired, behaviour. This may be compared with the ethological approach of (Tyrell, 1993) where observed behaviours had internal correlates, based on a hierarchy-of-behaviour approach inspired by (Tinbergen, 1951). The agents built by Scheier and Pfeifer were bottom-up mechanisms that co-ordinated behaviours (arranged in parallel as for the output nodes in a neural network) by interaction and learning. Action selection is implicit in the learned behaviour, and is, according to them, a false requirement which forces the building of internal behaviour modules directly correlated with desired observed behaviour.

While the approach of (Scheier and Pfeifer, 1995) is close to that used here, the argument against action selection is perhaps misguided. Action selection can be a by-product of the modelling or engineering approach adopted, for examples see (Joyce, 2001b) – also reproduced in Volume II of this thesis. However, it doesn't explicitly require internal behaviour modules to be correlated with observable behaviour; that is, simply the approach taken by one ethologist's perspective on producing a mechanism that explains observable behaviour.

Fundamentally, this thesis will integrate action selection with perception implicitly using neural networks, and this operant behaviour in the environment will cause other "layers" of control to influence the perceptual categories formed. In this work, classification is sensory motor coordination. However, because of the influence of the environment on higher-order control layers, *as well as* the adaptive networks, a number of influences from higher levels will affect lower level learning layers – a vertical distribution of control and influence. Most neural network based robotic agents are single-tiered and, therefore, vertical control is not an issue.

3.4 Mapping Connectionism to Abstract Agent Architectures

In this section, we attempt to provide interpretations of neural networks (such as those described above) as implementations of behaviour and interpret this in an agent-oriented fashion. Interpretations of behaviour congruent with both the neural network and intelligent agent communities will be investigated for a number of common features recurring in abstract definitions of agents.

This section concludes the motivation for what follows in Part II of the thesis; namely, that perception and action networks can justifiably be seperated and given concrete implementations in connectionist architectures and models.

The features are given in (Jennings, Sycara and Wooldridge, 1998) (Wooldridge and Jennings, 1995) but the extraction of the most common features are summarised in (Wooldridge, 1999) as:

- state and "memory" : the notion that agents in some way store information about their environment, actions or past interactions
- situated : the agent is engaged in a direct perception-action loop with the environment cf. expert systems where the "agent" is a centralised entity which communicates only indirectly with its environment (Wooldridge, 1999) pp.11
- autonomy : taking action without human intervention according to design goals
- flexibility : defined as *social ability*, *responsiveness* and *pro-active*

The key to flexibility (not elaborated upon by Jennings *et al*) is the agent's capacity to learn. If an agent responds in a timely fashion (responsiveness) and pro-actively attempts to achieve its design goals, then a pre-programmed solution or solutions may not be appropriate. Adaptive behaviour is often omitted from agent definitions, but here it is treated as fundamental.

3.4.1 Long and Short-Term Memory as State

Different practitioners use the ubiquitous terminology of long and short term memory differently. In real-time neuronal models e.g. the work of (Schmajuk, 1997), the presence of a temporal pattern of activity impinging on a neuron (via its weight) is considered to be a short term memory trace. Similarly Carpenter and Grossberg (1987) have adopted the convention and taken the modification of the actual weights connecting neurons to be long-term memory.

If we attempt to map memory to an automata-like theory of agency cf. the abstract agent models of (Wooldridge, 1999), then we find language such as "state" to be more commonly occurring. However, such a simplistic translation of long or short term memory to automata state may be inappropriate. To establish this, we must examine the functional role of what is stored by the matrix of weights and how this relates to temporal sequences of activity in the agent.

3.4.1.1 Acquisition and Storage of Experience

Close inspection of the structure of a fully connected, fixed architecture neural network (such as the MLP) has shown how any processing and output is explicitly dependent on the contents of the weight matrix. Again, we abstract away from the *actual* method of *changing* the weights, but describe the interpretation of weights assuming them to be dynamic.

The activation and signalling behaviour of a multi-layer network can be recursively described as the output O_k dependent on the preceding layer of nodes O_j as follows. Let O_j be the output of a node j:

$$O_j(\mathbf{x}) = g_h\left(\mathbf{W}_j^T \cdot \mathbf{x}\right) \tag{3.2}$$

and similarly, for the next layer O_k is defined as:

$$O_k(\mathbf{o}_h) = g_o\left(\mathbf{W}_k^T \cdot \mathbf{o}_h\right) \tag{3.3}$$

where $\mathbf{x}$ is the input, $\mathbf{o}_h$ is the vector of signals from the preceding layer $\mathbf{o}_h = \langle O_1, O_2, \ldots, O_j \rangle$, g_o and g_h are the signal functions, $\mathbf{W}_j$ and $\mathbf{W}_k$ are the weight matrices.

By effectively partitioning the network into two discrete components, each can be given considered as a seperate implementation task – e.g. in the construction of appropriate connectionist models. Recall that if i indexes the input/pre-synaptic nodes, j the receiving (post-synaptic) nodes and w_{ij} is the efficacy of the synapse/weight connecting them. There are M pre-synaptic and P post-synaptic nodes. To begin, we can take either of these two networks and treat them individually by removing the signalling function from the definition and treating the activations in isolation. For the first layer of hidden neurons, we can state that the activation of each node will be:

$$a_j(\mathbf{x}) = \mathbf{W}_j^T \cdot \mathbf{x} \tag{3.4}$$

where $\mathbf{x}$ is the stimulus, input or percept incident on a field of neurons feeding activation to hidden nodes j. Similarly, we can decompose the output nodes, except that $\mathbf{x}$ becomes the vector of signals $\mathbf{o}_h$.

Hence, if we take some input $\mathbf{x}$ and the activation of the nodes as a vector $\mathbf{a}$ at some instant t, then we find the relationship is :

$$\mathbf{a}(t) = \mathbf{W}^T(t)\mathbf{x}(t) \tag{3.5}$$

where the full matrices are as follows:

$$\begin{bmatrix} a_1(t) \\ a_2(t) \\ \vdots \\ a_j(t) \\ \vdots \\ a_P(t) \end{bmatrix} = \begin{bmatrix} w_{11}(t) & w_{12}(t) & \dots & w_{1j}(t) & \dots & w_{1P}(t) \\ w_{21}(t) & w_{22}(t) & \dots & w_{2j}(t) & \dots & w_{2P}(t) \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ w_{i1}(t) & w_{i2}(t) & \dots & w_{ij}(t) & \dots & w_{iP}(t) \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ w_{M1}(t) & w_{M2}(t) & \dots & w_{Mj}(t) & \dots & w_{MP}(t) \end{bmatrix}^T \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_i(t) \\ \vdots \\ x_M(t) \end{bmatrix} \tag{3.6}$$

So, we can postulate a time series $t = \{1, 2, \dots, K\}$ such that there exists an individual matrix $\mathbf{W}(t)$ which represents the relationship between the stimulus $\mathbf{x}(t)$ and the *desired* activity on the post-synaptic nodes $\mathbf{a}(t)$ for example:

$$\begin{aligned} \mathbf{a}(1) &= \mathbf{W}^T(1)\mathbf{x}(1) \\ \mathbf{a}(2) &= \mathbf{W}^T(2)\mathbf{x}(2) \\ &\vdots \\ \mathbf{a}(K) &= \mathbf{W}^T(K)\mathbf{x}(K) \end{aligned}$$

and hence, each $\mathbf{W}(t)$ represents the weights (memory) of the association between the input and the output (e.g. each weight matrix implements an instance of a linear associative memory between $\mathbf{a}$ and $\mathbf{x}$). Note that each instance of the weight matrix is what is necessary to produce a known $\mathbf{a}$ given an input $\mathbf{x}$. Following (Haykin, 1999), the *total* experience acquired is described as the long term memory $\mathbf{M}$ of the agent:

$$\mathbf{M} = \sum_{t=1}^{K} \mathbf{W}(t) \tag{3.7}$$

The pattern of experiences (states) can be defined as a recurrence relation which emphasises the nature of state updating as a result of actions or responses (e.g. a temporal sequence of individual experiences or memories):

$$\mathbf{M}(t+1) = \mathbf{M}(t) + \mathbf{W}(t+1) \tag{3.8}$$

So, $\mathbf{M}(t)$ is the experience accrued up to time t and $\mathbf{W}(t+1)$ is the "new" experience. The effect on the total memory $\mathbf{M}$ of any further experiences is lessened as $t \rightarrow K$ because of the difference in absolute values of each element of the weight matrix $\mathbf{W}(t)$.

In this case, all the data is presented in advance to an associative memory e.g. as a set of all inputs $I = \langle \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K \rangle$ and all the *known* or desired outputs corresponding to each input are presented. However, an autonomous situated agent will not have access to the set of all possible stimuli and the corresponding best response to take given those stimuli in advance.

In the case where we *do not have* a desired activation vector *a priori*, a better description is that at a certain time, the weight matrix is update by some change $\Delta\mathbf{W}$ such that:

$$\mathbf{W}(t) = \mathbf{W}(t-1) + \Delta\mathbf{W} \tag{3.9}$$

where $\Delta\mathbf{W}$ can be computed by a variety of means see (Sutton, 1988). Then, the total experience is defined by:

$$\mathbf{M}(t) = \mathbf{W}(0) + \sum_{i=1}^{t} \Delta\mathbf{W}(i) \tag{3.10}$$

where $\mathbf{W}(0)$ is the null matrix.

In conclusion, the discussion above shows *how* state can be interpreted for an agent using such a neural network. The functional role state plays in producing behaviour is that of affecting activations, which in turn cause different nodes to signal or fire. Long term state is represented in the weights of the network. Short-term state is a more complex notion, but can be simply identified with the activation levels on a collection of neurons.

Note we have not discussed the objective world/internal state correlation problems which so profoundly affect AI and connectionism. The systematic interpretability of such stored state is a difficult issue. In this section, we have chosen to follow (Sharkey and Jackson, 1994) in that we have treated the weights as a different kind of functional state to the activations. The "activations as representations" perspective was described in Hinton, McClelland and Rumelhart's article; (Rumelhart and McClelland, 1986) chapter 3. They describe each active node (neuron) as representing a micro-feature of an item presented to the network, and each weight (idealised synapse) as representing "micro-inferences" between these features.

In this thesis, we will take an approach similar to the divide suggested by (Sharkey and Jackson, 1994) but incorporate it with the philosophical perspective of indexical-functional representation (Agre, 1997) see Chapter 9. In effect, this represents a migration of Sharkey and Jackson's division from higher-order cognitive functions (such as concept learning and language processing) to sensori-motor coordination.

3.4.2 Perception and Action

If the agent state $\mathbf{M}(t)$ reflects the sum of all previous experience $\mathbf{W}(0), \ldots, \mathbf{W}(t)$, then what role does this state take in affecting output? This question is answered concretely for neural networks, because the state plays a direct causal role in calculating the vector of activations for each layer in the network.

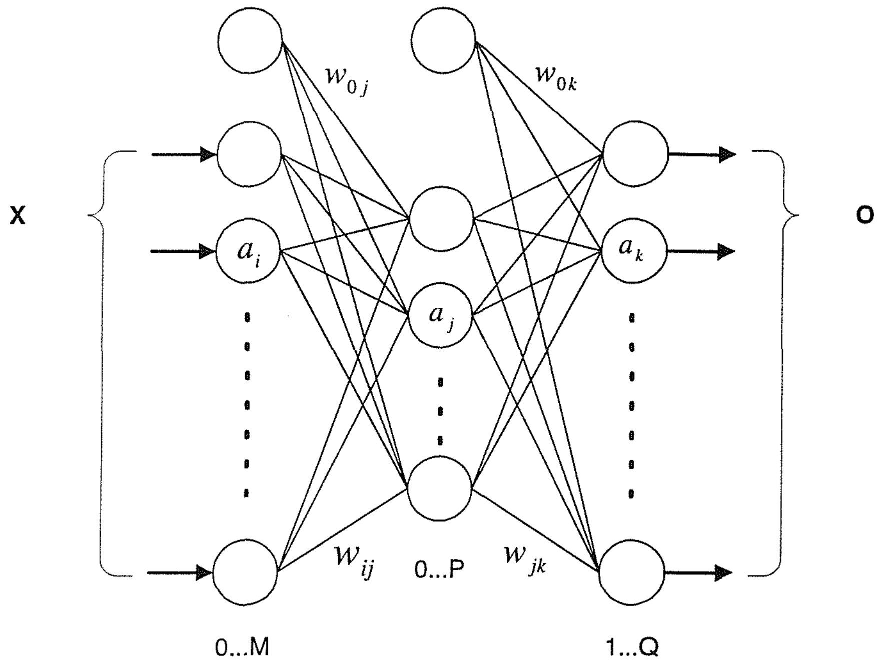

Figure 3.2: Multi-layer Network with One Hidden Layer

However, the interaction of perception and action with state is specifically tied to the kind of activation function used. Consider a two layer network such as that shown in Figure 3.2. Following (Sharkey, 1997) we can view the network as performing explicit mapping of raw sensory data onto a processed percept in the first layer. The question is 'how?'. Using the additive activation rules described here, each activation is the product of a vector $\mathbf{x}(t)$ (representing the stimulus at time t) and the weight state matrix $\mathbf{M}(t)$. This calculation causes a vector of activations on the post-synaptic neurons (e.g. those receiving input via the matrix $\mathbf{M}(t)$).

With reference to Figure 3.3, we call the activations of the post-synaptic neurons the *activation vector*, $\mathbf{a}$, then the vector of signals, $\mathbf{o}$, is a point in the bounded *signal space* $\Re^P$. Figure 3.3 shows the case for two neurons, which enables the view of the activation and signal spaces as a plane. Note that the weights are shown as elements $m_{ij}(t)$ of the *total* acquired memory at time t.

The computation is shown in Figure 3.4. The emphasis is that the agent's previous experience is coded in $\mathbf{M}(t)$ and this affects what is produced as the output signal vector (the percept).

The resulting signal vector is then passed to another network (i.e. layer of weights and neurons) which computes a mapping from percepts to responses. This view of the neural network in the context of memory, action and functional roles for intermediate states is elaborated upon later. A similar algorithm will map the vector $\mathbf{o}(t)$ to a vector of outputs node activations,

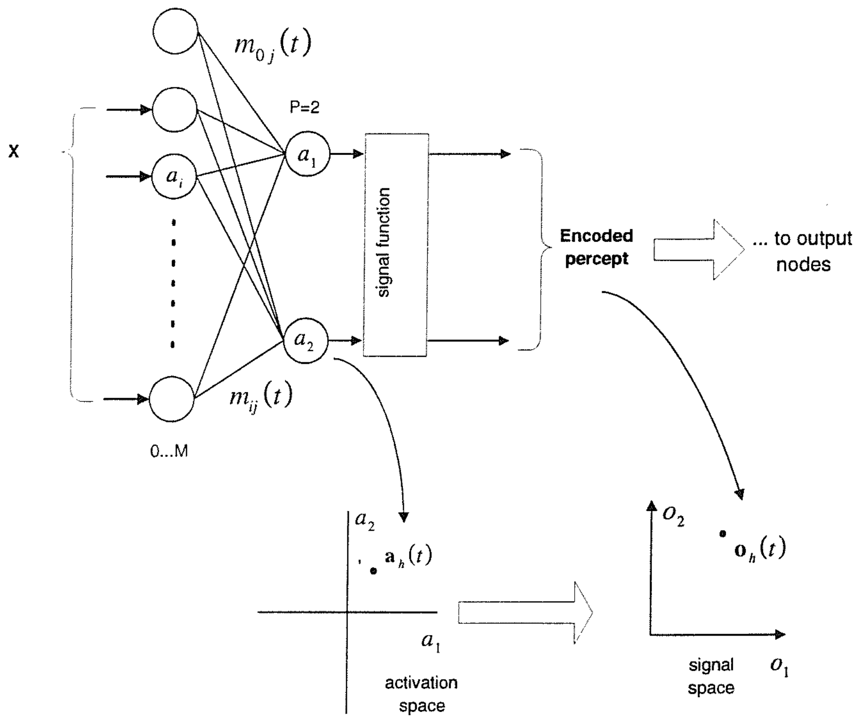

Figure 3.3: Encoding of Stimuli in One-layer of Weights

1. the stimulus is passed (by an identity function) "unprocessed" to the P post-synaptic nodes (activation spreading)
2. the activations of the P nodes are calculated (yielding $\mathbf{a}$)
3. the output/signal from each of the P nodes is computed (yielding $\mathbf{o}$)
4. the "percept" is then the signal vector $\mathbf{o}$

Figure 3.4: Computing Stimulus–Percept Mapping

representing a mapping from internal state to action.

The question of exactly *what* guides the creation of a signal vector (the internal state) is dependent on the network architecture, activation function and signal function. However, generally it can be stated the networks discussed here implement a product of two vectors; the weights impinging on a node and the signals transmitted along them.

In this case, each element of $\mathbf{a}(t)$ (the agent's internal state) caused by such an activation rule implements:

$$\mathbf{W}_j^T(t) \cdot \mathbf{X}(t) = \mid W_j^T(t) \mid\mid \mathbf{X}(t) \mid \cos\theta \tag{3.11}$$

which geometrically, implies that the agent correlates stored experience (memory) with new stimuli by the angle between the vectors in some space. Geometrically, then, the correlation of the input to stored experience is measured and the node j activated according to the similarity of its impinging weights and the stimulus. The activation of a node will be maximal when the angle between them (θ in equation 3.11) is zero. This is the basis for what Kohonen (1997) pp. 26 described as "classification of patterns by angles". To conclude, the basis of activation is stimulus similarity to stored internal encodings of past experience. This philosophically connects the neural network to the abstract agent architecture described in (Wooldridge, 1999). The adaptive resonance networks discussed earlier implement a functionally similar rule using the difference between the weight and the element of the input to the node. Note if all weights are equinormal, then the two different activations rules are functionally the same (Kosko, 1992) pp. 147 and competitive-type activation rules (based on differences between input and stored experience) reduce to the correlation measure.

3.4.3 Autonomy and Situatedness

As stated earlier, a signal vector resulting from the first layer of neurons is then passed to another sub-network (e.g. layer of weights and neurons) which computes a mapping from percepts to responses. This view of the neural network in the context of memory, action and functional internal state provides an important link to the notion of situatedness.

In terms of situatedness in cognition, perception and action, (Clancey, 1997) offers the following working hypothesis :

> "All human action is at least partially improvisation by direct *coupling* of perceiving, conceiving and moving – a coordination mechanism unmediated by *descriptions* of associations, laws or procedures. This mechanism compliments the inferential processes of deliberation and planning that form the backbone of theories of cognition based on manipulation of descriptions. Direct coupling of perceptual, conceptual and motor processes in the brain involves a kind of "self-organisation with a memory" that we have not yet replicated in a computer programs, or indeed in any machine" – pp.2 (emphasis in original)

It should be noted that Clancey's use of the word "descriptions" is in part technical, in that he uses the term to refer to what is called "deliberative" agency and symbolic AI. How does connectionism capture this? On its own, it is debatable that connectionism offers anything other than an implementation of computable function. However, if the agent-theoretic notion of situatedness *combines* with connectionism in an appropriate way (e.g. a valid training rule where weight modifications are a function of direct consequences and experience of the environment) then:

- there exists a form of *structural coupling* (Maturana and Varela, 1980a) whereby an agent engages in constant interactions with its environment, maintaining a suitable internal (self-organising) structure to enable it to "survive" in its domain given the (possibly changing) demands of that environment. We might build an analogy between structural coupling (or more precisely, the demand that structural coupling makes of the the agent in terms of adaptivity) and the self-organising of the weight matrices for a connectionist network.
- perception and action (loosely, stimulus-response behaviours) are effected directly (e.g. the input, hidden and output layers of the MLP) by super-positional representation which makes no recourse to explicit, descriptive world models.
- in accord with Clancey's definition connectionist networks can be the substrate of epistemology – but the denotative aspect of systematic interpretability is lost unless the theory of representation applied is not 'internal symbol to objective world' (see the discussion later in Chapter 9)

A connectionist system provides for a purely reactive system, but incorporates adaptive behaviour explicitly. It is reactive and implements *bounded rationality* because the time complexity of computing an action is dominated by the time taken to compute two matrix multiplications and the costs of calculating the signal function. It is difficult to state that the agent is purely reactive *and* rational (according to the optimality definitions of rationality), because the decision taken may not be optimal as the agent has not yet learned the best action given a certain contingency. So a basic implementation of the network of Figure 3.2 results in one activation for each hidden node P which requires a total of MP multiplications (i.e. the input vector multiplied by the weight matrix) and similarly each output node Q requires PQ multiplications. Let the signal function cost be constant at S_c then the total cost of computing an action C_A is of the approximate order of:

$$C_A = MP + PS_c + PQ + QS_c \tag{3.12}$$

For an local ART-based network, the cost is bounded similarly. It is difficult to state the overall cost at this stage because the network to map internal state to actions has not been elaborated upon. Suffice to say, it is certainly bounded by C_A, but the size of the network's P layer can grow (but hopefully to a limit). If weights are updated through learning on each iteration, then back-propagation adds a cost of $O(MP+PQ)$ since the derivatives are computed with respect to the weight matrices, so the overall time complexity is polynomial.

A final comment on the nature of situatedness is due. An MLP which is trained to perform and then has a fixed functional role in producing behaviour is not congruent with the notion of agent situatedness. This should not be confused with arguments about grounding sensori-motor ability in robotic capacities (this is a different philosophical arena for which connectionism arguably offers an explanation and technical practice). A trained MLP using back-propagation minimising the error measure between the actual output and the desired output is no longer adaptive, and therefore unable to reconfigure behaviour appropriate to the demands of the environment. A connectionist-based reactive agent which contains a continually adaptive system *is* different because it constantly modifies its behaviour based on feedback and situated action in the environment. Both the MLP and localised RBF/ART networks can be implemented in this way.

In terms of autonomy, we cannot proceed with invariant or generalised notions of the term. In order to move forward, it is suggested that a definition of autonomy is self-sufficiency (McFarland, 1996). An agent which can sustain continued interactions with the environment is considered autonomous in McFarland's interpretation (which treats internal state, particularly homeostatic state, as a dynamical state-space system). A useful characteristic, which may feed the idea of autonomy in connectionist systems, is to remove one conception of neural networks as passive receivers of information which "cue" the device into responding. In part, this is due to the pattern classification applications of the technology, but in Chapters 5, 7 and 9 we consider how a form of motivation can be added to connectionist systems, which reinforces the notion of motivated or goal-driven autonomous behaviour *and* emphasises interactionist theories of autonomy.

3.5 Conclusion

The notions of an agent recently emerging from the software and intelligent agents community has been biased by its grounding in traditional AI. As a result, consideration or systematic study of connectionism in intelligent agents (as they are now conceived at least in the computer science domain) has yet to become focused. As was suggested in the introduction, this could simply be because there are no migratable principles from the largely robotic agent research (where connectionism has received attention) and software agents. Positively, however, most work on such robotic agent principles begins as software simulation (purely for economic reasons).

In this chapter it has been shown that:

- notions of neural networks from robotic agents migrate across to generalised, abstract agent architectures

- neural networks (connectionist networks) have properties which qualify them as abstract agent architectures (e.g. state and stored experience)
- a neural network implementing an agent's internal mechanism constructs two kinds of stored memory (state):
  - a long term memory (state) of experience which have preceded the current action and which *directly* influence future actions
  - a short term memory (state) which represents the current mapping of stimulus to a "current" internal state with direct reference to the stored experience (long term state)
- response or action generation is an integral part of the mechanism, despite being viewed as a separate sub-network mapping internal state to action

The next chapter considers how learning and adaptation in connectionist systems can be implemented as a function of environment contingencies (emphasising the situated and embedded character of interactionist agent theory).

Part II

Towards A Phenomenological-Connectionist Theory of Agency

Chapter 4

Connectionist Models for Agent Perception and Action

This chapter describes work undertaken to implement a mechanism for agent action. It adopts the modified framework for behaviour-functional and situated-automata given in (Joyce, 2001b) – see Volume II of this thesis.

Using the motivations and division of perception and action common to agent theories and connectionist interpretations given in (Chapter 3), a series of suitable connectionist models are described. These models are later used in applications which require agents to perceive and act in their environments. These mechanisms are derived from an agent theoretic understanding of connectionism, and are presented as generally as possible. From such a generalised basis, they are modifed for each environment considered (e.g. see Part III). This in conjunction with Chapter 5 contribute a theory building exercise, which is then tested in Part III.

The breadth of material on the subject of reinforcement in learning, from lower level cognitive models through to machine learning abstractions, forced a revisiting of literature which, for brevity and convenience, is summarised in (Joyce, 2001a) and reproduced in Volume II of this thesis.

4.1 Adaptive Models of Perception

This section illustrates the ways single and multi-layer networks are used in this thesis to provide a generalised model of perception and the production of internal state. In Chapter 8, the same generalised model is expanded to model an holistic internal state which also includes homeostatic information about goal completion/satiation.

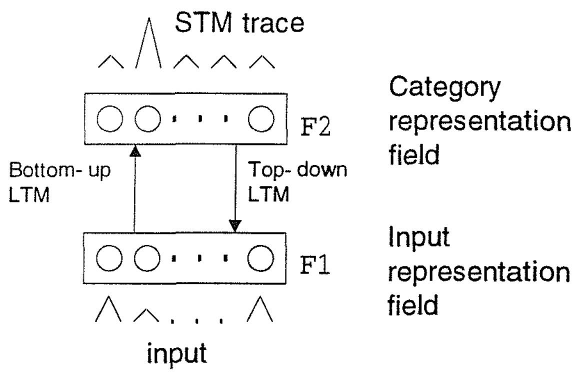

Figure 4.1: Outline of ART showing input / category fields

4.1.1 Perception and Modalities

In the neural reinforcement model of Schultz, Dayan and Montague (1997), they identified an architecture for an agent which locates the ventral tegmental area (common to the mid-brain of most animals, which is the region most implicated in evaluating emotional or affective conditioning in action) as a focus of descending encodings from different cortical modalities i.e. different classes of processed input to the agent. Similarly, self-organising connectionist models have been used by Sun and Peterson (1998) and any feed-forward network essentially implements a similar kind of self-organising mechanism by virtue of its signal function and weight modification (learning) algorithm.

Cortical modalities, as an abstraction, are useful here. They represent an efficient coding of the perception of the environment, and potentially, with somato-sensory state in an internal state coding. This thesis provides exploration of plastic self-organising networks, akin to (Carpenter and Grossberg, 1987) see also (Levine, 1991). The reason for considering adaptive resonance techniques (ART) was in their grounding in neural networks and perception/cognition. Various modifications of the original ART1 have been made; through computationally simpler models with real-valued inputs (Carpenter, Grossberg and Rosen, 1991) and finally culminating in analyses of ART modules as if they are self-organising systems with computational geometric properties e.g. (Williamson, 1996). ART networks can be implemented in a computationally simple fashion (by discretising them) and then they become functionally similar to radial-basis function networks. ART networks can be interpreted as building maps of the stimuli space (that is, they afford computational geometric interpretations). The key difference is the localised property of ART neurons; each neuron is said to "code for" certain stimuli or stimuli prototypes[1].

The basic ART strategy is shown in Figure 4.1. An input vector is incident on a field

[1]It is also interesting to note that these networks do not divide the input space using hyperplanes. Instead, they model the distribution of that space using kernels or basis functions

of nodes, each node representing a collection of neurons, which activate and send signals to F2 (where again, each node shown is modelling a collection of neurons that act as a receptive field). F2 nodes are connected together in a laterally inhibitory field, so each node inhibits the activation of each other node. This implements a winner-take-all (WTA) function over the category nodes. However, as the activations grow and compete on F2 nodes, top-down connections (weights or idealised synaptic paths) transmit growing activations which then act to suppress or further enhance F1 nodes. This coupled behaviour between neuronal fields F1 and F2 is modelled by a system of differential equations which (Carpenter and Grossberg, 1987) have shown to converge to a stable state with a winning category node in F2. Crucial to this selection of a category node is the definition of a scalar *vigilance* which determines the level to which input and category exemplar must match. If the match is too low, the network effectively shuts down the poorly matching node. The network continues to oscillate activations between F1 and F2 until either a winning node which matches *above* the vigilance parameter is found, or a new node is recruited to cope with the presumably novel input. Carpenter and Grossberg have described this as a neural realisation of a spatial search for matching patterns. Analogously, the search for an internal state given a mapping of the agent's percepts.

Note that Carpenter and Grossberg choose to interpret activations on F2 (such as a vector of real values representing points in a state-space) as short-term memory and the top-down and bottom-up weights connecting the fields as long-term memory. Similarly, the F2 nodes have been taken to be simpliciters of the visual cortex (Varela, Thompson and Rosch, 1991) pp. 96. This observation lends the use of such models to produce a generalised stimulus to percept mapping in this thesis.

4.1.2 Constructive Stimuli-Space Maps

Williamson (1996) explored a method of constructing ART modules by explicitly stating a generalisation function which is Gaussian in form and where the parameters of the function are learned weights. Williamson's specialisation of ART (called GART) is more robust with respect to noise because, unlike original ART, it forms smooth receptive fields over the stimuli space instead of discontinuous hypercubes.

The implementation used here is shown as a feed-forward network in Figure 4.2. Note that all weights and activations are bottom-up. The systems dynamics are as follows: Let the stimulus representation layer be S, category / percept layer P and the vigilance controlled layer be C. The nodes in S simply implement an identity signal function given some stimulus or input denoted by the vector ι. The first layer which performs any real computation is therefore, the P layer. Let nodes in S and P be indexed i and j respectively. The weight vector for a given P node is $\mathbf{W_j}$. As usual, the entire matrix of weights for connecting all S nodes to P nodes is the augmented vectors forming the matrix $\mathbf{W_{ij}}$.

The vector of signals from the S layer is simply the input vector $\mathbf{S_S} \equiv \iota$. Each P node com-

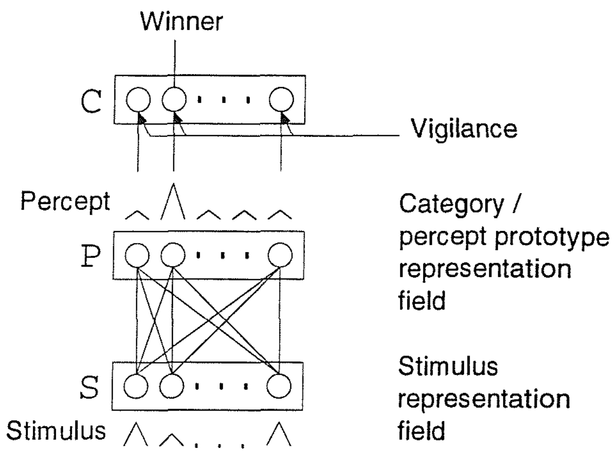

Figure 4.2: ART-like Network for Perception

putes its activation as the squared Euclidean distance between the input and stored (learned) weight vector:

$$a_j = \sum_i (\iota_i - w_{ij})^2 \tag{4.1}$$

so the nodes build activation as a quadratic function of the distance between the learned weight vector and the current stimulus ι. It is now possible to define the activity instantiated by the P layer as a point in activation state-space:

$$\alpha_P(t) = \langle a_1, a_2, \ldots, a_{|P|} \rangle \tag{4.2}$$

where $\mid P \mid$ is the number of nodes in the P layer. The computation resulting in the vector $\alpha_P(t)$ is the *percept* computed by the agent's virtual perceptual machinery. Note that this vector of activations is then used as the basis for selecting a single (winning) internal state.

Given the activation, a suitable signal function is e.g. (Moody and Darken, 1989; Duch and Jankowski, 1999):

$$S_j^P = \exp\left[\frac{-a_j}{2r}\right] \tag{4.3}$$

where r is the radius of the receptive field. This would normally correspond to the covariance matrix for a multivariate Gaussian representing node j's receptive field

4.1.3 Category Choice and Internal Automaton State

Utilising the automata-theoretic model of computational agency, there must at some time be an indicative internal state which configures the system for a decision, response or action generation. It must be (and under the scheme developed here is) a functional state; that is, causally efficacious in generating the response or action.

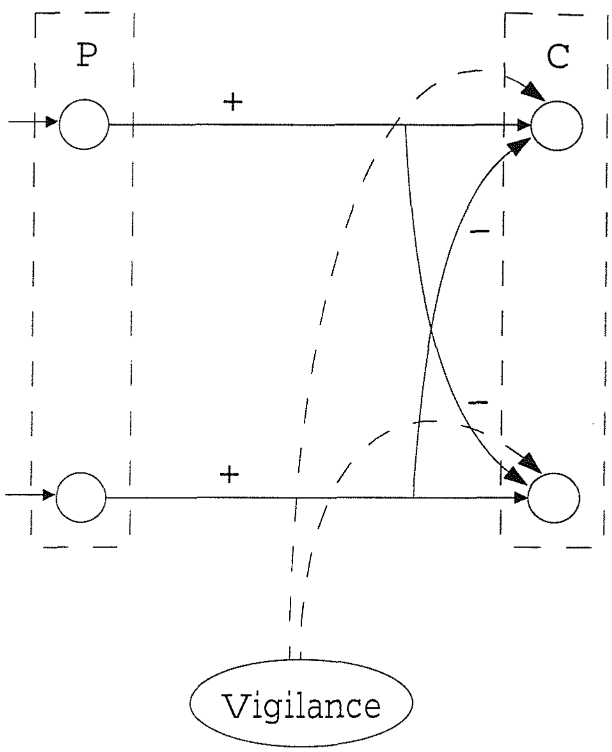

Figure 4.3: State Selection by Forward-Spreading Activation Inhibition (On Centre, Off Surround)

Such a mechanism is that implemented by the *C* layer. The key features are as follows:

- a winner-takes-all (WTA) mechanism must be implemented to select the most probable internal state of category; the cognitive efficiency of this (for the agent) is that the internal state selected configures it for a number of response or action alternative.
- each *C* node is partnered with a *P* node in a one-one mapping
- *C* node activations are proportional to *P* node signals
- the *C* layer configures the agent to choose and appropriate response governed by the current vigilance (e.g. attentional acuity)

It should be noted that for clarity and ease of exposition, the *C* and *P* nodes are treated differently, and the competition is implemented by *activation inhibition* using the Grossberg "on-centre, off-surround" mechanism. It is feasible to collapse the *C* and *P* nodes into one layer, and then implement the same functionality with intra-layer lateral, recurrent inhibition connections between nodes.

Figure 4.3 shows a basic configuration for two corresponding *P* and *C* nodes. Note that the connection from any *P* node j to *C* node $k = j$ is excitatory. There is also an "on-centre, off-surround" nexus of inhibitory fixed weights (Grossberg, 1976; Levine, 1991) so that any

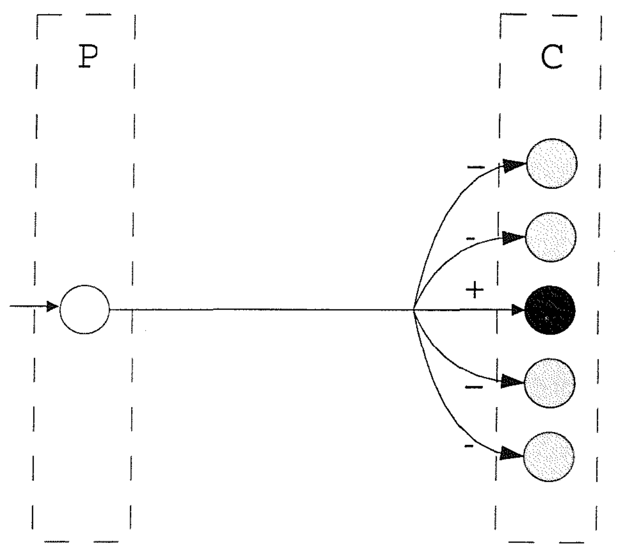

Figure 4.4: On Centre, Off Surround network for a single P node projecting to a number of C nodes

positive activation spreading from node j to k inhibits nodes $k \neq j$. In the full dynamical system underpinning this WTA mechanism, $P \rightarrow C$ non-adaptive weights would have corresponding $C \rightarrow P$ non-adaptive weights which as the C nodes begin to compete, the activation (and thus the signalling) of the other (losing) C nodes is flattened. Figure 4.4 elaborates an entire "on-centre, off-surround" mechanism for a single P node and the entire layer of C nodes.

The dynamics of this system are governed by the following equations:

$$a_k = \varepsilon_{k=j} - \zeta_{k \neq j} - \upsilon \tag{4.4}$$

$$\varepsilon_{k=j} = S^P_{k=j} \tag{4.5}$$

$$\zeta_{k \neq j} = \sum_{\forall j \neq k} S^P_j \tag{4.6}$$

where ε is the excitatory component (on-centre) and ζ the inhibitory component (off-centre). $\upsilon \in [0,1]$ is analogous to Grossberg's *vigilance* parameter, which is a scalar, non-specific level of attentional arousal. To explain the role of this parameter, recall that if P nodes effectively compute the similarity of the current input (stimulus) to a known situation coded in the weights $\mathbf{W_j}$ (where j is some P node). The corresponding signal S^P_j will be bounded in the range $[0,1]$ indicating the "match" between the current stimulus and the stored experience coded for by $\mathbf{W_j}$. This will excite the corresponding C node k (while simultaneously forcing the activation of neighbouring C nodes to 0). This activation is further reduced by ρ that acts as a dynamic threshold above which the match S^P_j must exceed in order to be counted as a *strong enough correlation* between stored experience and the current input. If $\upsilon = 0$, then any match which wins the competition will do, and conversely, if $\upsilon = 1$ then only the closest matching winning node is allowed. The inhibitive field can be derived from the shunting inhibition equations using an appropriate signal function (e.g. a sigmoid / exponential formulation) (Grossberg, 1976) see also (Schmajuk, 1997; Levine, 1991).

The WTA mechanism described is an activation-inhibiting system, where nodes compete *and then* signal (e.g. fire) if they are supra-threshold. The threshold for firing is that the activation must be greater than zero, thus:

$$S_k^C = \begin{cases} 1 & \text{if } a_k > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

So, it is expected that the network configuration ensures that once the vigilance parameter has been subtracted from the net activation (incorporating inhibition and excitation) only one node will have supra-threshold (e.g. > 0) activation. The binary signal function is then an indicator of the agent's internal state – effectively, a computation that selects the agents overall disposition to take certain actions.

4.1.4 Adaptation of the Stimulus Space

In the network presented here, the usual parameters for a receptive-field approach are present – the centres of fields (e.g. weights $\mathbf{W_j}$ for each P node j) and the radii (or covariance) for the fields (represented only as r in the signal function S_j^P).

The problem faced in implementing such a network was separating out the effects of the vigilance and the learning of centres. The decision was taken to simplify the problem so that effectively, the centres and radii are fixed constants. Adaptive context-sensitive behaviour is instead gained through a potentially infinite number of P nodes which can be recruited (as in ART) as and when new situations and contexts arise in the agent and the environment. The changing vigilance, υ, implements a dynamic radius for these otherwise static receptive fields. This is because υ implements an attentional mechanism which is identical to a uniform growth or reduction in the radius of all receptive fields. This is explained in Figure 4.5. The top diagram shows (for a one-dimensional stimulus space) the locations of the receptive fields. The thick black line is the stimulus and P nodes $j = 1$ and $j = 2$ are activated and signalling. The resulting activations on the category nodes (layer C) are shown below, with the result of the vigilance parameter reducing the activation of nodes $k = 1$ and $k = 2$. If a high vigilance is present, then the neither node signals, because the match in the P layer was insufficient. For the medium setting of vigilance, $k = 1$ wins.

In the former case, adaptation occurs; the network recruits, or grows, a new node centred on the stimulus which defines a new receptive field which accommodates the novel stimulus. If some C node is supra-threshold, then the weights coding the centres of the P nodes vary according to the competitive learning rule:

$$\Delta w_{ij} = -S_j^P w_{ij} + S_i^S S_j^P \tag{4.8}$$

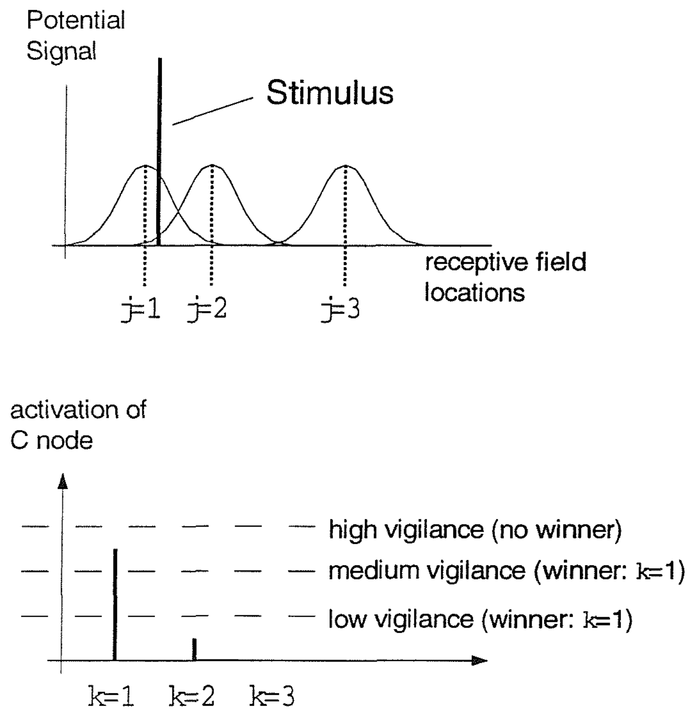

Figure 4.5: Explanation of the relationship between receptive fields, activations, vigilance and winning node

---

A consequence of not adapting weights is that it simplifies the agent's cognitive machinery. The negative consequence is that this machinery may grow without bound unless the agent enters into some routine activity where the number of required nodes is stable. Applying equation 4.8, the weight change would be $\Delta w_{ij} = S_i^S - w_{ij}$ because $S_j^P = 1$. The weight at time t can then be described as:

$$w_{ij}(t) = w_{ij}(t) + S_i^S - w_{ij}(t) \quad (4.9)$$
$$= S_i^S \quad (4.10)$$

so in one iteration of the deterministic competitive law, the novel input is encoded by adapting the weight to *become equal to* the new input. To summarise, when there is no supra-threshold node (after the vigilance has been subtracted from the C nodes) then a new node is recruited as described (e.g. the node codes for the novel input instantly). If a node *is* supra-threshold, then no adaptation of existing nodes occurs. This is similar to the mechanism of (Peterson and Sun, 1998).

4.1.5 Cognitive Basis of Perceptual Qualities

Following the work of Shepard (1987), Lee (1997) described a number of experiments with the ALCOVE neural network, which learns to produce 'psychological spaces'.

Generalising from stimuli to categories (and the learning of these categories) has been the subject of much study e.g. see (Harnad, 1987). For reasons of cognitive efficiency, an autonomous agent must assess the similarity of stimuli to those of previous experiences, but the cognitive computational structure is, as yet, poorly understood. The nature of this generalisation can be captured wholesale by a 'perception function' G cf. (Luce, 1959). If this is extended such that the computation of the next internal state is $\Gamma(d(\iota_a, \iota_b))$, where d is some similarity measure between two stimuli. For this reason, the number of possible generalisation functions is that of the number of computable functions. Lee went on to explore how this under-constraining might be approached with respect to available literature.

Shepard's work (Shepard, 1987) identified the Minkowski metric (with $r \in (0,2]$) as the basis of stimuli comparison (or distance) and then posited that the invariant generalisation gradient is an exponentially decaying function of this distance. This conveniently maps to the $r = 2$ metric used here along with the Gaussian basis function used to generalise within categories.

The computational model of reactive agency meshes with connectionism because:

- the basis of perception and internal state selection is based on category stimulus similarity and competition between categories
- (Luce, 1959) proposed that probabilistic behaviour should provide the basis of studies in psycho-physical decision tasks
- WTA, lateral inhibitory competition and internal state choice are based on generalisation, category learning and stimuli space mapping and are implemented in an unsupervised radial-basis function-like mechanism (e.g. which implements the prescription for one kind of distance metric and generalisation function)

The nature of unsupervised learning can now be addressed. Doya (1999b) proposed a gross categorisation of regions of mammalian brains that have computational analogs so that:

- perception (in its broadest cognitive sense) implicates the cerebral cortex and this largely relies on *unsupervised learning*
- learning sensori-motor tasks implicates the cerebellum as a mapping between perception and action using *supervised learning*
- motor 'output' to the peripheral nervous system is effected (again, coarsely) by the basal ganglia; a collection of regions which use *reinforcement learning*

Compatible with Doya's proposal, the perceptual mechanics of the local network use unsupervised learning. Later, the internal goal-directed aspects of the agent architecture (homeostatic state) are fed to the internal state computation. Doya (1999b) pp.966 notes that "unsupervised learning modules in the cerebral cortex work as the medium for representing the state of the external environment as well as the internal context. They also provide a common representational basis for the cerebellum and the basal ganglia, between which there are no direct anatomical connections." This suggests that some kind of 'contraction' of information and its exchange in a common 'currency' must occur between perception and action. This is effected by the WTA selection and its passing of a signal to the motor circuits (which are discussed later).

If the agent's discussed here were capable of locomotion, then the regions of the space activated by different stimuli would correspond to features of the environment such as walls and maze junctions, which constitute a map of the environment. Hippocampal regions (corresponding to receptive fields) of rats have been shown to fire in a spatially and orientation selective fashion (O'Keefe and Dostrovsky, 1971) and the asymmetric shape of these receptive fields shown to be dependent on experience (Mehta, Quirk and Wilson, 2000). In the model used here, the fields are all strictly symmetric and adaptation is limited. An extension to this work would be to incorporate these recent findings into a model that is able to cope with the skewed stimuli space of the agent.

In essence, the constraining factors on models tend to be both cognitive psychological (as in observed generalisation in discrimination/categorisation) tasks and biological/anatomical (for example, in the work of O'Keefe and Dostrovsky, Hubel and Wiesel (1965) and notably, the biologically-inspired levels which implement the computational level of (Marr, 1982)).

4.1.6 Extensions

There is no reason why a recursive structuring of P layers (e.g. multiple instantiations of the perception mechanism) cannot be used to implement a form of high-order relational learning of the kind described in (Thornton, 2000) and (Thornton, 1996). Thornton frames agent learning in the context of information theory (Shannon and Weaver, 1949), arguing that maximising prediction (analogously, choosing an action which will be successful and goal achieving) is minimising information.

In terms of generating an internal state, Figure 4.6 illustrates this principle with reference to a Mealy automaton model of a computational agent[2]. The hatched area shows the response generating part of the agent yet to be elaborated on. The function G is shown as being affected by the LTM (long-term memory) accumulated in the weight matrix $\mathbf{W}$ (which grows over

[2]A Mealy automaton has a function mapping input to internal state, and function mapping internal state to output cf. a Moore machine which has only a 'next internal state' function and the output is defined as a direct consequence of the state (rather than a function of internal state)

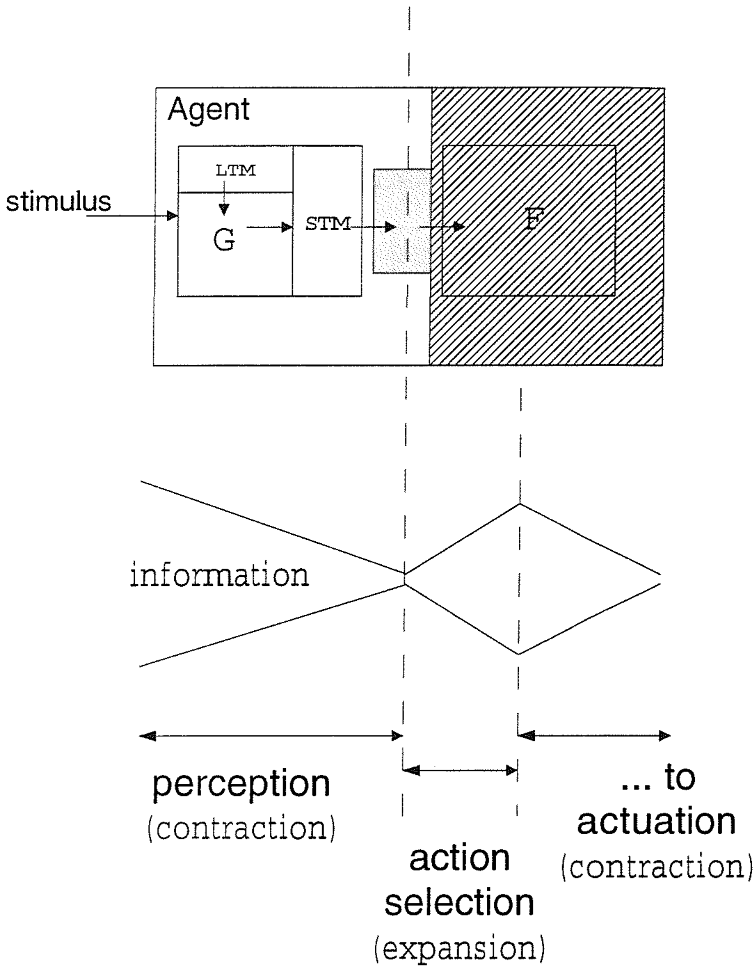

Figure 4.6: Diagram showing the qualitative mapping of information to features of the automaton model of computational agency; The function G represents perception and the holistic internal state of the agent, and F the action produced as a function of this state

---

time) and producing some STM trace (to use Grossberg's phrase) analogous to the percept α. The grey area in the diagram is the interface between perception and action, which is not as discretely divided as shown, but corresponds to the output of the C layer.

Thornton's approach is that by exploiting regularities in the stimulus space the learner (the agent) is able to justify assigning probabilities to outputs (actions) (Clark and Thornton, 1997). The argument progresses as follows; if an agent can find regularities *explicit* in the data (stimuli) then it can easily exploit these. A trivial example is that if a harmful consequence occurs when a 1 is present on a sensor, but when 0 occurs the result is always satisfactory, then the sensor and the learning network exploit this explicit regularity.

On the other hand, if this simple indicator is not present, then regularities implicit in the stimuli must be found over time (experience) which enables the agent to discover the relationships between features of the input and the consequences. This variety of learning is called *non-relational* – the former explicit case being *relational* – by Thornton and Clark. Thornton's direct proposal is that a recursive building of layer upon layer of a similarity-based learning mechanism enable further regularities of the non-relational kind to be discovered in the data. Thornton cites the XOR or parity problem as *perfectly non-relational* because all outputs depend on all values of the input so that similar, metrically adjacent inputs will have *non-adjacent* outputs see (Thornton, 2000) chapter 7.

In this work, the recursive building of P layers will not be attempted because it adds an order of complexity to the problem, but it is interesting to note that this is a potential way forward embracing the notions contained herein.

4.1.7 The MLP as a Bi-partite System

In the preceding discussion, an ART inspired system was described and simplified to a RBF network where only the centres of basis functions are learned see also (Peterson and Sun, 1998) where the same simplification was made. The input layer and the hidden nodes implement a kind of correlation-measure based perceptual mechanism. The activations of the hidden nodes are equivalent to the vector $\alpha_P(t)$. The connections from the hidden layer to the output layer implements a mapping from the internal state to actions.

The finite size of the MLP hidden layer make $\alpha_P(t)$ a more useful analog of internal state than the choice of winning C node in the ART-based localised network. However, the weights connecting the input layer to the hidden nodes still play a similar role; by adaptation, it shapes the internal states (e.g. the realisations of $\alpha_P(t)$) thus implementing the long-term associations between inputs.

As mentioned above, both implement a kind of quantisation of the input (stimuli) space. The MLP can be advantageously interpreted, as shown by (Harnad, Hanson and Lubin, 1994), to implement categorical perception effects which decrease within-category grouping of stimuli and increases between-category grouping (in Euclidean space). Harnad *et al*'s work illustrates that kernel formation (as for ART networks) is functionally present in MLPs also. The first layer weights are therefore implementing similar functionality to those of the ART weights, but their coding is quite different. The MLP codes by using an error measure back-propagated from the the output layer generalised from the Widrow-Hoff rule, whereas localised network weights code for the centres by iteratively converging the weight to be the mean of the stimuli category. The kinds of receptive field organisation exhibited by localised, distance-metric activation function networks also appear in MLP nets. Hidden units become 'tuned' to respond maximally to certain stimuli (representative of means of the category).

Similarly, the representational efficacy of the first layer weights has been described by Sharkey and Jackson (1994) as the context independent representation of the input space. While the activation vector represents the non-concatenative representation cf. (Smolensky, 1988; van

Gelder, 1990) of the complex of individual inputs, it is the functional role of the weights that determines the shape of this activation vector. The dynamics of the weight matrices in both local and distributed systems provides the causal (and thus, functional) force and utility. The role of the second layer of weights in an MLP determine the mapping (or 'carving') of the hidden unit space into regions which correspond to desired output. Qualitatively and functionally, the localised network and the MLP offer a similar perceptual mechanism, but they code this mechanism in different ways.

4.1.8 Examples of Related Work

Multi-layer networks have received a lot of attention as generalised learning techniques which can be applied to a variety of adaptive control problems. This section will focus primarily on examples of adaptive behaviour in agents.

4.1.8.1 Decentralised Perception

(Sharkey, 1997) presents experiments on using multi-layer neural networks to develop (by experiential learning) representations which are portable, that is, reusable if the morphology of the agent is different. Such a philosophy appears to run contrary to sub-symbolism, but in fact, Sharkey notes that "... the representations developed, for example, by multi-layer perceptrons during learning play a necessary role in the computation and are used whether or not they appear in the mind or the observer." pp. 346. Sharkey appeals to the necessity of *something* between perception and action which mediates the process. This is a special case of philosophical functionalism (Block, 1980a), where the representation is necessary and the content or semantics can be defined similarly with reference to the environment.

Sharkey's experiments use MLPs to control a robotic arm disembodied from its sensory faculties. Three networks were independently trained. A classification network was used to classify objects and spatial locations (a perception net), another to coordinate the arm's position in space with motor commands to the arm's actuators (a motor net) and then an "inter-net" which attempted to compute and communicate the invariants in the perceptual nets hidden layer to the motor net's hidden layer. The purpose was to explore the invariant representations which the inter-net formed to co-ordinate perception and action when they are separate physical entities. The importance of this work is the exploitation of representations captured by the inter-net and their functional role in exhibiting behaviour.

In terms of its relationship to agent theories, the decentralised representations employed by the components of Sharkey's model are more akin to the "society of mind" approach of (Minsky, 1986). This proposes that a "mind" is a collection of simpleton agents responsible, individually, for simple tasks. The collective behaviour constitutes a mind. Similarly, Selfridge's perceptual "demons" are an earlier instantiation of the parallel, distributed (and ultimately, connectionist) philosophy of agency (see (Remez, 1987) for discussion).

4.1.8.2 Neuro-evolution and Supervised Learning

An attempt at removing the supervisor from the training loop for agents using neural networks was attempted by (Nolfi and Parisi, 1993). They used two multilayer networks; a "teaching" network which supervises the other "learner" network. A simple analogy for the teaching network is a configuration of innate behaviours which are used to influence learning during the agent's lifetime. The original interest in this model was to explore the boundary between what is learned by the agent-environment interactions and what must be configured *a priori* by some other mechanism. In this example, the *a priori* behaviours are provided by evolution-environment interactions (the process of selection) and on a microscopic scale, interactions from the agent-environment influencing learning. A similar methodology, but where the architecture does not learn and a novel neuronal model is used, have also been implemented in the work (Channon and Damper, 1997) from (Cliff, Harvey and Husbands, 1992). Instead of the usual feed-forward arrangements, these systems use both inhibitory and excitatory parallel outputs, as well as time delayed activations, noise internal to the neuron and recurrent connections.

However, in this thesis, the actual mechanisms of adaptation are considered in an agent framework – simulated evolution coupled with connectionist models are impractical for wholesale application to software agents purely because of the massive redundancy/utility tradeoff that is necessary to establish a solid candidate solution for a given agent in an environment. Eventually, this "total autonomy" approach may be viable; where agents are being entirely designed via structural coupling with the environment and its Darwinian-like selection mechanisms.

For this thesis, the problem is approached from the perspective of lifetime learning only, using hand-configuration of any features which are environment specific (e.g. such as virtual sensors and unconditioned stimuli which have significant effects on behaviour that is learned). A similar relationship between learning and evolution in the philosophy of connectionism is explored by (Compiani, 1996) who provides connectionist analogues for the nativist vs. tabula rasa perspectives inherent in developmental psychology.

The reason for choosing this as an example here, is that it only employs simple neural network models for both the learning and evolutionary components (the teaching network). The division between *what* is to be learned and *how* it is learned is clearly defined by different mechanisms – namely, the lifetime learning network is entirely governed by the supervision of a network which is produced by a separate simulated evolution process and does not learn during the agent's lifetime.

Fundamentally, the simulated evolution produces a training set (expressed as a feed-forward neural network) which trains the agent's learning net. It is this learning and the structural and dynamic aspects of the neural network model that are of vital importance here. Simulated evolution is a simple way of producing the necessary training set for a particular environment without human intervention.

However, this intuitively notion of evolved innate behaviour which is adapted during life-time learning does not address the fundamental question of *appropriate* connectionist mechanisms for adaptive behaviour in general. For example, in (Nolfi and Parisi, 1993) the evolution of a separate network, a teaching network, shortcuts the real issue, which is how the agent learns from the environment. In their work, the agent learns because another network is evolved, which responds to the same environmental stimuli and delivers the phylogenetically determined correct response. We might say that in Nolfi and Parisi's work, we "consult" artificial selection in order to search for the correct response. This is then coded in the teaching network and subsequently life-time tested on the agent.

4.2 Synopsis of Literature for Learning Actions

In this section, three models suitable for connectionist and automata-like models were reviewed – the stochastic learning automaton (SLA) (Narendra and Thathachar, 1974), behaviour-functional (Arkin, 1998) and the more general reinforcement learning (RL) (Kaebling, Littman and Moore, 1996; Sutton and Barto, 1998). In addition, associative models where also examined (Chang and Gaudiano, 1998) which aim to produce reinforcement learning effects in simpler systems based on classical conditioning; see also (Damper, French and Scutt, 2000).

Most similar to the architectures explored here are:

1. a localist, topographic implementation (Moore, 1994) which similarly constructs state space maps using computational geometric principles, but via recursive decision-tree partitioning of the state space, whereas here, a 'flat' model is used; see also (Peterson and Sun, 1998).

2. an MLP (function approximation): (Tesauro, 1995; Tesauro, 1992; Sun and Peterson, 1998) and the complimentary reinforcement back-propagation technique of (Ackley and Littman, 1991a; Ackley and Littman, 1991b).

Other approaches to the implicit learning of action-selection, with some kind of internally generated training signal *and* sensori-motor interaction, include the 'instinct' rules of (Nehmzow, 1995). These approaches generate teaching vectors for a single-layer network trained using the delta rule. Likewise, the neuro-evolutionary approach of (Nolfi, Elman and Parisi, 1994; Nolfi and Parisi, 1993) which uses an entire evolved MLP to generate teaching signals. These do not use the reinforcement learning framework.

The deterministic finite-state automata approach to modelling behaviour (cf. situated automata theory) was described by (Luce, 1959) as an *algebraic* model of behaviour on account of its specification being driven by a model isomorphic to legal syntactic expression matching. Essentially, this represents a *formal grammar of action* for a reactive agent. Luce's alternative is the probabilistic specification of behaviours which is implicit in connectionist modelling. Specifically, Luce's model applies to psychophysical experimentation on choice behaviour in a variety of modalities and is the analogy (in cognitive science) to the probabilistic automata approach taken in this thesis.

In deciding between various approaches, two were chosen; a localist network implementation of Q-learning and a function approximation using the second layer of a MLP. Both have advantages and disadvantages for analysing the routine behaviour which, it is hoped, emerge from the agent/environment coupling.

There is, however, a distinction between two classes of connectionist system:

- the class of techniques where the reinforcement signals of classic reinforcement learning are examples of secondary reinforcing stimuli (Catania, 1992; Barto, Sutton and Anderson, 1983) which represent *a priori* associations between a, usually unconditioned, stimuli and its value to the agent. For example, if the agent is provided with a shock, which the designer decides is the result of poor action. This information is conveyed to the agent as a single negative real-valued reinforcement.

- the class of *associative networks* which *produce* reinforcement-like effects e.g. (Grossberg, 1972a; Grossberg, 1972b; Schmajuk, 1997; Chang and Gaudiano, 1998; Klopf, 1988) so that stimuli are used directly to influence conditioning mechanisms; the resulting behaviour resembling operant conditioning.

During development, an attempt was made to generalise action learning from the work undertaken on the MAVIS2 agents (Joyce and Lewis, 1999a; Joyce and Lewis, 1999b; Joyce, Lewis, Tansley, Dobie and Hall, 2000), which uses a purely associative network with no delayed reward. Integrating these different perspectives relies on generating the abstraction of the reinforcement learning signal from 'lower order' stimuli. In unpacking this alongside considerations of connectionist goal directed behaviour, see for example (Levine and Leven, 1991), it became clear that the mechanisms using associative learning of reinforced behaviour would have to be explored (particularly as the control architecture evolved).

4.3 Discussion

Many workers have used the function approximation ability of connectionist networks to estimate the values functions for states and actions see (Kaebling, Littman and Moore, 1996) for an overview. What is significant about connectionist techniques is that they can be an implementation of both the function V and the policy. The state value function for some policy π is:

$$V^{\pi}(x) = E_{\pi}[R_t \mid x_t = x] = E_{\pi}\left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \middle| x_t = x \right\} \tag{4.11}$$

where E_π is the expectation operator. The random variable R_t given the current state[3] can be visualised as follows. Assume that the agent only ever receives positive reward, such that $\forall t, r_t = 1$ then :

$$R_t \mid x_t = \gamma^1 \cdot 1 + \gamma^2 \cdot 1 + \gamma^3 \cdot 1 \ldots \tag{4.12}$$

so that for any state $x_t = x$, there is a realisation of R_t with an associated probability. However, an analytical solution for the distribution or density function for all values of R_t is not available, hence direct evaluation of V is not possible. Instead, RL methods can be used in combination with sampling of the state-space refine estimates of V.

Similarly, we need to define a function which yields the value of taking a certain action if the environment is in a certain state (e.g. the perceived stimulus for the agent). This is the *action-value* or Q function and is defined as follows :

$$Q^\pi(x,o) = E_\pi[R_t \mid x_t = x, o_t = o] = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \,\middle|\, x_t = x, o_t = o \right\} \tag{4.13}$$

where x and o are the state and action respectively. Hence the value function for a policy is related to the Q value by (Ballard, 1997):

$$V^\pi(x) = \max_o [Q^\pi(x,o)] \tag{4.14}$$

Updating of Q-values is as follows:

$$Q_t(x_i, a_j) = (1-\eta)Q_{t-1}(x_i, a_j) + \eta \left(r_t + \gamma \max_k Q_{t-1}(y, a_k) \right) \tag{4.15}$$

such that η is the learning rate, the γ is the future reward discount factor, x_i is the ith internal state, a_j the action taken when the agent was in state x_i and r_t the reward resulting from that action/state pair. Q_{t-1} is the previous value of the Q-value, and $\max_k Q_{t-1}(y, a_k)$ is the next predicted action (based on the old Q-value).

This can be shown to be a Hebbian-like learning rule by considering the levels of abstraction in modelling synaptic pathways. Figure 4.7 shows five abstractions, derived as follows:

1. From (Ballard, Hayhoe, Pook and Rao, 1997) and (Gerstner, 1999), neuronal spikes, i.e. models at the level of membrane biophysics, occur on the time-scale of order 1 msec.

2. From signal rate models where activation and spiking are time averaged using a temporal scale of around 10 msec (see (Ballard, Hayhoe, Pook and Rao, 1997) pp. 724). The classic model of McCulloch and Pitts (1943) is at this level of abstraction.

[3] the current state x is the agent's current internal state, *not* the environment state alone as is usual in RL *without* perceptual generalisation over inputs

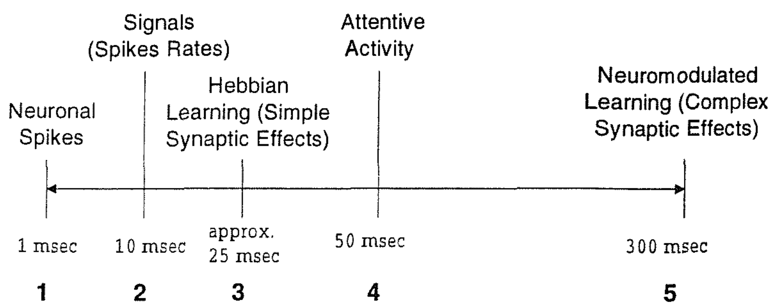

Figure 4.7: Multiple Levels of Abstraction in Modelling Learning (Not to Scale)

3. Studies of the induction of long-term potentiation (LTP) and depression (LTD) can be measured on the time scale of around 25 msec. For example, if post-synaptic signals occur around 20 msecs *before* correlating pre-synaptic signals, then LTD can be observed. Likewise, if post-synaptic signalling occurs within a 20 msec window *after* pre-synaptic signals, then long term potentiation is observed. These experimental results were reported by (Bi and Poo, 1998) in hippocampal cell cultures measuring excitatory post-synaptic potentials resulting after the above signalling regimes were used on the cells. See also the treatment of (Kosko, 1992). This abstraction is referred to by Shepherd (1994) as the *simple synaptic* model of adaptation corresponding to Hebb's postulate of learning (Hebb, 1949).

4. Signal correlation alone cannot account for delayed delivery of unconditioned stimuli e.g. those which report rewards. To account for delayed consequential stimuli, Klopf (1988, Barto, Sutton and Anderson (1983) introduced the notion of eligibility; acting as a measure of how much future stimuli, such as reward or punishment, affect a synaptic pathway. This therefore spans the range 'attentive activity' (cf. Ballard *et al*) to 'Neuro-modulated Learning'

5. Brown and Chatterji (1995) propose that a local Hebbian learning scheme can be viewed as the activity of a neuronal field or group whose behaviour (with respect to adapting weights) is affected by global neuromodulators which are representative of reinforcement and control the plasticity and changes to a group of synapses. Shepherd (1994) proposes this as a 'complex' model of synaptic dynamics, and concurs with Ballard (1997, Ballard, Hayhoe, Pook and Rao (1997) who argue that deictic codes (see Chapter 9) and reinforcement learning occur at the temporal scale of around 300 msecs. It is also the most convenient level of analysis for the drive and motivation of Hull's systematic theory of behaviour (Hull, 1943) and Thorndike's 'bonds'; see (Hilgard and Bower, 1966).

This enables the model of learning used in the localist network illustrated below. Let w_{ij}

be the idealised synaptic pathway between neuronal groups which model internal state i (the C nodes) and the motor pattern-generator (Ewert, 1995) for action j. Then, the Q-learning formula expressed as a Hebbian-like learning rule with complex synaptic effects is:

$$w_{ij} := (1-\eta)w_{ij} + \eta\left(r_t + \gamma \max_{j'} w_{i'j'}\right) \tag{4.16}$$

where i' and j' are the predicted next state and corresponding best action. Expanding and rearranging this equation:

$$w_{ij} := w_{ij} - \eta w_{ij} + \eta r_t + \eta\gamma \max_{j'} w_{i'j'} \tag{4.17}$$

and then re-writing using the differential form of the Hebbian learning rule (see (Kosko, 1992) for further discussion):

$$\begin{aligned} w_{ij}(t) &= w_{ij}(t-1) + \Delta w_{ij} && (4.18)\\ \Delta w_{ij} &= -\eta w_{ij} + \eta r_t + \eta\gamma \max_{j'} w_{i'j'} && (4.19) \end{aligned}$$

The first term is the passive decay of the weight, the second term ηr_t subsumes signal correlation and neuromodulatory effects (cf. Shepherd's simple *and* complex synaptic effects) and the final term includes neuromodulatory effects analogous to eligibility and predicted rewards/punishments (elaborated further in Chapter 8 when the proposal of (Doya, 1999a) is considered). For this reason, it is proposed that Q-learning is a localised Hebbian-based rule working at the functional abstraction of neuronal groups. This, then, forms the basis of the topographic implementation of action learning that follows.

4.4 Two Implementations of Reinforcement Learning

4.4.1 Topographical Implementation

Given that the perceptual machinery maps the stimuli space into a quantised set of categories, the connection to a local network essentially implements a lookup table. However, the justification for considering this a connectionist system is that (as illustrated above) a complex synaptic model can be extrapolated from the basic Hebbian form.

What is required, then, is a topographic associative map between internal state and the agent's response. Each internal state node will be connected to each possible action. Assuming an associative memory (with more complex learning), a topographic, localised network was implemented.

Recall that the C nodes code for a single category and internal state combination, and only one must remain active. This common currency of exchange is cognitively efficient for the agent, since the next computation (e.g. for action selection) involves only minimal choice. In

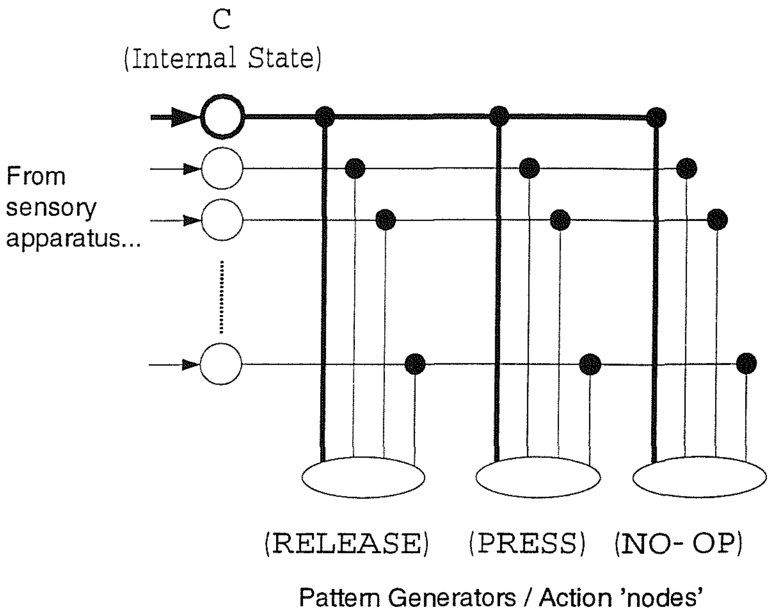

Figure 4.8: Associative Network for Action Learning

effect, the disposition to act enables a select pathway from the internal state to actuating machinery. Figure 4.8 shows this arrangement as an associative memory after (Barto and Sutton, 1981) for an agent that can choose between three actions – PRESS, RELEASE and NO-OP (signifying 'do nothing'). The diagram's appearance is mainly for expository convenience – the filled circles representing modifiable synapses. In the diagram, the thick line shows the result of the top *C* node being activated. Hence, the sum of the sensory input and internal variables has caused the mechanism to select a specific node which predisposes it to choose one of the three actions (represented as pattern generator nodes). The weights from *C* nodes to actions are a sensori-motor pathway, or schema, which is then subject to further regulatory functions to determine which action is taken and which weights are updated.

This network, combined with a Q-learning rule, enables the agent to learn more than stimulus response pairs. The network can, in principle, code for sequences of S-R contingencies. More recent studies of the rat hippocampus have shown that action-outcome pairing is also important in instrumental conditioning (Corbit and Balleine, 2000). That is to say, the action taken and its resulting outcome (e.g. the pressing of a lever and the delivery of a reward) are in some way encoded as well as the S-R association and that chains of S-R association being acquired and extinguished is not sufficient to explain instrumental behaviour. Their experiments showed that in a rat with a lesioned hippocampus, the *acquisition* of instrumental behaviour (e.g. behaviour leading to the attainment of a motivating goal) was the same for non-lesioned rats. So, they learned the action-outcome pair. However, if the probabilities of outcomes were changed so that the learned action-outcome pair was degraded in utility (e.g. if the rewards were delivered irrespective of the actions taken) then the lesioned rats failed to accommodate this change in the environment. They suggest that the action-outcome encoding function of the hippocampus is more closely related to context learning and the ability to recognise 'free' rewards so that when rewards are free, the causal connection between action-outcome is decayed because its instrumental value has likewise decayed. Evidence such as this suggests more sophisticated models of action are required to account for instrumental conditioning of the kind developed for the agents in this thesis. A full investigation of such theory would be suitable for potential future work.

In the connectionist paradigm, the behaviour functional notion of the response set $R(t)$ manifests itself as the gains or "action probabilities". The relationship between weights and action probabilities is given by the proportionality :

$$\rho_{ij} \in R(t) \propto \Pr[o_i \mid s_j] \tag{4.20}$$

which states that the probability of an action o_i given the internal state s_j is proportional to the presence, gain or strength of the element ρ in the response set. Recall that $R(t)$ is the collection of all possible actions and their viabilities, gains or strengths at time t.

To generate a behaviour, a method of *effecting the policy* is required. Recall that a response set at time t is a set of instantaneous options for action coupled with a real scalar. The activations caused by the weights transmitting and amplifying antecedent (state node) activity implements a particular set of response space values (o, ρ).

For example, the agent described in Chapter 7 has three actions available; PRESS, RELEASE and NO-OP. The nodes driving the action nodes can be any indicator of internal state – so far, this has been described as the activity of a category node given the generation of percept.

The response set in the behaviour-functional model (Arkin, 1998) is equivalent to the combination of the policy π and the mechanism for introducing non-determinism in the action selection. One common method is to use a Boltzmann distribution over the Q values to select an action which is usually the action which maximises V^{π} but occasionally, to select a less valuable action to encourage novel behaviour.

The implementation used here is to take the weights connecting C nodes to the actions as coding for Q values (see section4.3). At any one instance in time, the agent will select one internal state and this selectively causes specific activations on the action nodes according to the weights connecting the C node to the action. If the input to the action nodes is simple a weighted sum activation, then the node activates in direct proportion to the stored Q value coded by the weight.

4.4.2 Modelling Output/Action Selection

Recall that weights from C nodes coding for current internal state s_j map to action nodes and o_k let this weight matrix be $\mathbf{W}$ where w_{jk} denotes the weight from C node j to action node

k. Since only one C node is active, at any instant in time, a vector of relevant weights will be $\mathbf{W_k}$. Let the activation of the action (output) node k be a_k^O and the signalling of the C node j is $S_j^C = 1$ such that:

$$a_k^O = w_{jk} S_j^C \tag{4.21}$$

trivially, then, each action node is activated proportionally to the value of the state-action pair $Q(s_j, o_k)$ in the Q learning algorithm. The associative memory of Figure 4.8 therefore encodes the state-action contingencies in terms of previous reward. The probability of taking action k given internal state j should then be computed from:

$$\Pr[o_k|s_j] = \frac{\exp\left(\frac{a_k^O}{temp}\right)}{\sum_{k'} \exp\left(\frac{a_{k'}^O}{temp}\right)} \tag{4.22}$$

By increasing the Boltzmann temperature, the most likely action has its potential flattened to a level equal to that of the other actions. The neuronal activation dynamics in continuous time would enable an interpretation of the Boltzmann function as a collection of neurons (as in Figure 4.8) implementing a WTA circuit. (Kohonen, 1997) chapter 8 discussed hardware analog implementations of WTA networks, and (Kaski and Kohonen, 1994) describes models based on spiking behaviour rather than activation building which, while useful, are less congruent with the more familiar models which separate activation and signal dynamics as two interacting systems.

The circuit and discussion of (Yuille and Geiger, 1995; Yuille and Grzywacz, 1989) demonstrates a number of strategies for producing WTA effects from softmax, using statistical mechanics and constrained optimisation. In Yuille and Grzywacz's formulation, the WTA functionality is recovered from the more general softmax formula e.g. as the temperature in equation 4.22 tends to infinity. Their model is based on pre-synaptic inhibition between neurons, requiring that activation building be defined as primitive shunting inhibition with contrast enhancement provided by the exponential function (Yuille and Geiger, 1995):

$$\frac{da_k}{dt} = -a_k + I_k \exp\left(-B \sum_{k' \neq k} a_{k'}\right) \tag{4.23}$$

Note that the similarity with shunting inhibition is that the excitatory input I_k is multiplied by the inhibitory input. As with most continuous models, the first term is the passive decay. The effect of the other neurons' $k' \neq k$ activations are incorporated by an exponential function of their sum weighted by a negative constant B analogous to Boltzmann temperature. Most importantly, this equation also enables the variable B to be interpreted as analogous to a gain function. The temperature variable acts as a coefficient of inhibition and self-excitation. The shunting effect (division of activations) is amplified by the temperature.

Recall that it is usual to follow the RL policy π but occasionally, to encourage exploration, pick a random action. However, this requires a shift from the paradigm of a deterministic neuron spiking behaviour to probabilistic interpretation of firing rates. If each action node's activation is affected by some incoming signal (excitation) and intra-field (e.g. other action nodes) inhibition, then the dynamical system of equation 4.23 will yield a mean activation of each node over some Δt. During this interval, each neuron will sporadically fire if its activation is supra-threshold. At some moment $0 < t' < \Delta t$ during the interval, the neurons may or may not be spiking. This uncertainty is the exploit used to implement non-determinism in the model used here and thus, to implement non-deterministic (but controlled by temperature) WTA. A fuller discussion of the spike rate-codes and probabilistic measures can be found in (Gerstner, 1999; Maass, 1999).

If over the period Δt the mean activation is defined as:

$$\langle a_k^O \rangle_{\Delta t} = \frac{\exp\left(\frac{1}{\beta} w_{jk} S_j^C\right)}{\sum_{k' \neq k} \exp\left(\frac{1}{\beta} w_{jk'} S_j^C\right)} \tag{4.24}$$

then the winning node should *mostly* be picked based on the highest activation, depending on the temperature β. Recall that a discrete system assumes that this temporal summation occurs and that all firing is synchronous. However this does not reflect the theoretical premise stated that *sometimes* nodes with lower activations fire. This can be simulated by assuming that due to noise, at some time $0 < t' < \Delta t$, a sample of the activations is taken (assuming that they are still building over the interval Δt and that the final winning node's activation is not always larger over the period Δt) and the resulting highest activation fires. Effectively, the interval Δt is treated as a unit interval divided into regions which are proportional to the average activation over the period Δt. By then making the point at which signalling occurs a uniform random variable (analogous to the sample at time t') in that range, most often, the largest activated node will fire. Alternatively, other nodes with lower activations will fire. This is similar to the roulette wheel sampling used in genetic algorithms (Mitchell, 1996) Chapter 5.

The division of a unit interval is illustrated in Figure 4.9 (the top diagram shows example weights). Recall that the excitatory input to an action node is merely the weight, as the signal for the state node k is zero or one. The nodes then build activation over some interval and this is implemented as described above. The conceptual selection of the winning node (i.e. which action will fire) is illustrated in the bottom diagram, which shows (for low and high Boltzmann temperatures β respectively) the proportional allocation of interval range to probability of being the winning node by virtue of activation level. A uniform random variable simulates the moment t' where one node fires (in the diagram, the fictional value $t' = 0.56$ is given). Notice that if the temperature is low, the highest activated node wins, otherwise, the other nodes are selected.

The system described in Figure 4.9 was implemented and the temperature varied. The frequency of the most probable node (e.g. highest weight and therefore highest activation over

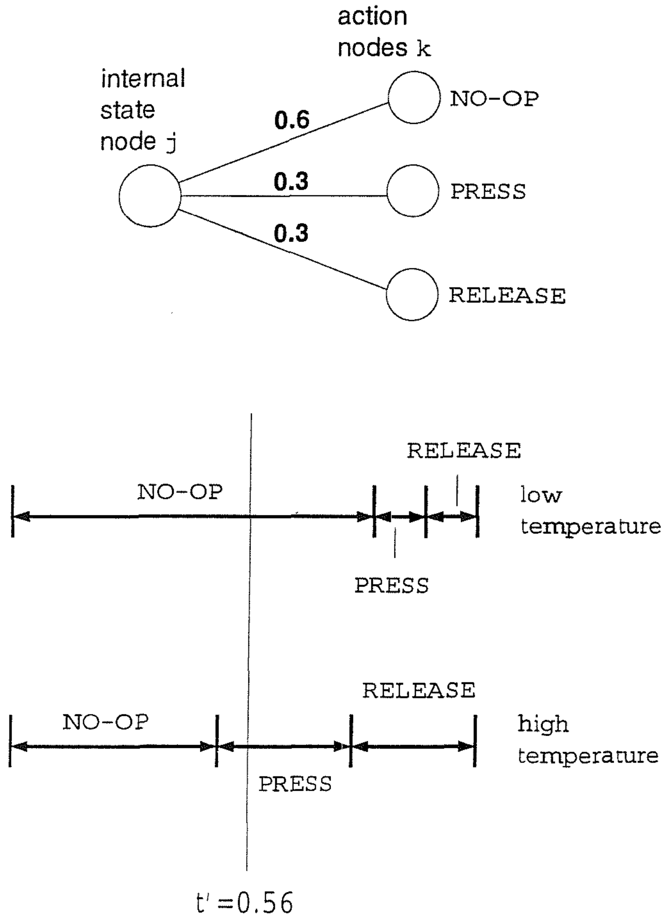

Figure 4.9: Top: Example weights from an internal state C node to the three action nodes; Bottom: Examples of proportional allocation depending on node activation, temperature and resulting probability of selection

the interval Δt) *versus* any other node (i.e. random action selection) being chosen was plotted over 10000 trials. The results are shown in Figure 4.10. Notice that as temperature increases, the winner selection method detailed above chooses more randomly, whereas if the temperature is low, the action corresponding to the learned (e.g. highest weighted action) always wins (e.g. the system collapses to deterministic WTA).

This interpretation is carried over to the MLP network, where the function approximator feeds activation to a WTA network as described above.

4.4.3 Function Approximator Implementation

Given the descriptions of MLP networks in Chapter 3, an MLP can be seen as:

1. a first layer of weights which quantises and maps the stimuli space using finite resources (the number of hidden units) according to the functional requirements of the desired

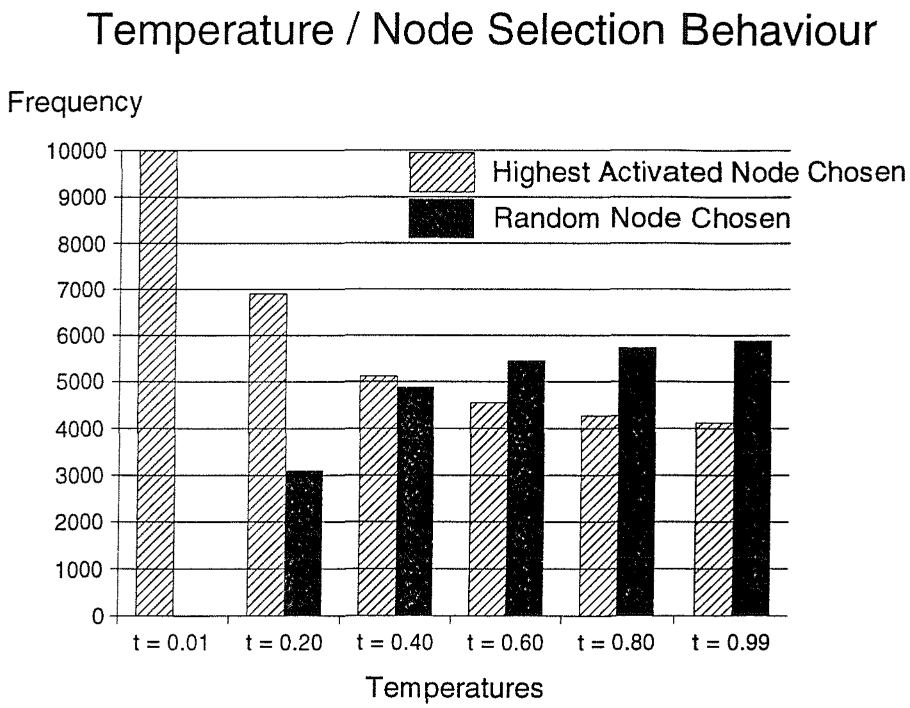

Figure 4.10: Frequency of highest activated node being chosen over other nodes with varying temperature

behaviour (that is, behaviour desirable in the environment)

2. second layer outputs the Q values for any given state (recall that the vector of hidden activations $\alpha_P(t)$ is not enumerable as the C nodes are in the topographic implementation above)

3. a training process that modifies both layers of weights; this implicitly means that both the shape of the stimuli map (i.e. receptive fields formed implicitly in the input layer weights) *and* the output layer of weights (coding for the Q values) is affected by environmental feedback (see below)

The usual error measure is sum-of-squares (SSE) between the output node signals (representing the actions) and the target vector (describing the desired signalling output of each node):

$$E = \frac{1}{2}\sum_k (S_k - t_k)^2 \tag{4.25}$$

where S_k is the signal from output (action) node k. This is interpreted as the estimate of the value $Q(x, a_k)$. Following from (Sun and Peterson, 1998), the network is taught to approximate the Q values according to the Bellman residual. At some time t, the agent responds with a vector of outputs and the stochastic Boltzmann action selection takes place. This results in one 'winning' node, denoted O_k corresponding to an action k (i.e. PRESS, RELEASE or NO-OP). Denoting the current state/action $Q(x,a_k)$ and the future most likely state (e.g. the most likely next state/action pair) $\max_{k'} Q(y,a'_k)$. The target vector is then:

$$t_k = \begin{cases} r_t + \gamma \max_{k'} Q(y,a'_k) - Q(x,a_k) & \text{if node } k \text{ wins} \\ 0 & \text{otherwise} \end{cases} \tag{4.26}$$

This value is given to the back-propagation formula which then recursively adjusts the weights in preceding layers of the MLP. The interesting feature of this rule is that all layers of the MLP are modified in the service of approximating the current Q function (representing the value of states and actions taken in those states). Both the perceptual 'layer' and the output layer of the MLP are thus affected by feedback from the environment in the form of r_t. The errors computed and used in the back-propagation formula are then dependent on the environment and not an external teacher. If the value of states changes over time, then the network must 're-learn' the Q function. The only obstacle to this is that that the learning rate is usually decayed by some schedule over time to enhance convergence of the output nodes in approximating the Q function.

Weight modifications in back-propagation algorithms are according to the gradient of the error with respect to weights. So, for output nodes where the signalling of node k is S_k, the error is defined as:

$$\delta_k = S_k - t_k \tag{4.27}$$

and the errors for hidden units j (those analogous to internal state) with signalling S_j and weights connecting hidden node j to output node k denoted w_{jk}:

$$\delta_j = S_j(1 - S_j) \sum_k w_{jk} \delta_k \tag{4.28}$$

which enables the derivatives to be formed for the input-hidden weights w_{ij} and the hidden-output weights w_{jk} respectively :

$$\frac{\partial E}{\partial w_{ij}} = \delta_j S_i \tag{4.29}$$

$$\frac{\partial E}{\partial w_{jk}} = \delta_k S_j \tag{4.30}$$

so that the weights are changed in the direction of the gradient of the error with respect to the weights:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \tag{4.31}$$

In general then, η controls the overall plasticity of a weight and δ_j the contribution of the Bellman residual. This latter component incorporates the reinforcement signal r_t and the temporal difference between states $\gamma \max_{k'} Q(y,a'_k) - Q(x,a_k)$ weighted by the future reward discount as for Q learning in a topographic representation.

4.5 Conclusion

The choice between function approximation and a topographic map is not arbitrary. Here, both the local topographic approach and a function approximator were tested. Arguments for a topographic implementation are concerned with the plausibility of back-propagation of errors and the structural features of the MLP network:

- the partial derivatives which must be made available to the back-propagation algorithm and the temporal differences algorithm are not likely to be available in a Hebbian-plausible network model
- there is some biological support for the topographic nature of areas of vertebrate animals responsible for action selection e.g. (Doya, 1999b) pp. 965 illustrates how the basal ganglia region might use reinforcement learning as the principal source of adaptation and that the descending morphology of the motor circuitry is organised in a highly topographic fashion. Berns and Sejnowski (1996) also propose a model which relies on highly modular and localised connectivity.

The regulation of activity for an autonomous agent will be complex, so it is desirable to have a transparent model of action selection where local and global regulatory effects (i.e. those affecting many idealised synaptic pathways and those affecting selective pathways) can be clearly demarcated. Essentially, this is an argument for interpretability of the mechanisms underwriting behaviour. Both mechanisms were implemented and tested in the simulation described in Chapter 7.

This chapter has introduced the connectionist mechanisms that were used as the basis for a connectionist-based agent architecture, assuming the division of mechanisms for perception and action in Chapter 3; one mechanism relies on topographic maps of internal state to actions where each weight codes for the estimate of the Q value, the other using the MLP as a function approximator. In addition, the chapter contributed a unified model of complex and simple synaptic effects, illustrating that the Q-learning rule is situated between signal correlated Hebbian rules and neuromodulated adaptation of idealised synapses.

Chapter 5

A New Phenomenological Agent Theory for Embedded Agents

This chapter contributes a theory of agency by proposing a means of realising phenomenology in a way compatible with semiotic theories of representation. There has been much discussion (especially in the cognitive sciences) about embodiment and the naturalisation of phenonemenology – that is, cognitive scientific explanations of the philosophy of subjective experience, for example (Roy, Petitot, Pachoud and Varela, 1999). Given that situatedness (loosely, the embodiment of an agent) is key to agent theory (i.e. to discriminate between the expert systems tradition and the newer proposal of an intelligent agent as a situated entity) , it is an interesting prospect that the naturalisation of phenomenology might converge with agent theory since their fundamental concerns are coincident yet their theories, methodologies and practices are different.

Rather than merely introducing and discussing these theories in the form of an esoteric description, this chapter will attempt an integrative formulation relative to the project of agency and connectionism. Semiotics provides a purely hermeneutic theory of representation and meaning. Heidegger's phenomenology provides a perspective on the grounding of deliberation in everyday routine activity (e.g. apparently mundane "regular" behaviour). The proposed theory of phenomenological agency attempts to close the distance between these two perspectives by offering an operational version of both traditions in a connectionist compatible way. What results is a proposed vertical layering of activity from routines to deliberation which indicates how the methods of AI and connectionism might usefully be deployed to implement these layers.

Throughout this chapter, discourse such as; 'input', 'output', 'stimulus' and 'response' will be used to maintain continuity with previous work. These terms have lost favour in the study of situated cognition because they imply a separation of agent from the world, situating the agent as merely a passive receiver of properties transmitted by the environment. Discussion is appropriately contained to aspects of the early 20th century phenomenological movement most relevant to this thesis.

5.1 Introduction

Why consider phenomenology? In previous chapters, the idea that symbolic AI is implicit in agents has been held to be restricting for an agent theory[1] It was the ignorance of the AI community to Heiddeggarian phenomenology that formed the basis for Dreyfus' work attacking the methodology of AI (Dreyfus, 1992). Attempts at reconstructing AI, see especially (Agre, 1997), have been largely encapsulated and subsumed into the category "reactive agents" of which Brooks' work is held as the flag-bearing exemplar.

As justification for this work, we can consider the example given in (Dreyfus, 1996). He recounts an experiment performed in which a junior world chess champion played against a marginally weaker player. The chess game was conducted under conditions reminiscent of bounded rationality, wherein the players had to complete each move in the game in under five seconds. In addition, the junior champion had to add together numbers presented, at a rate of one per second, as rapidly as possible before the delivery of the next number. In effect, the junior master was engaged in two activities; one requiring expertise and skill (chess) and the other (Dreyfus hypothesised) requiring abstract computation. Dreyfus noted that the junior master chess player had no time to engage in deliberation (as posited by a symbolic AI heuristic-search machine) and simultaneously, that any deliberative ability was stretched by the player being asked to do arithmetic. The junior master played adequate chess, winning in a succession of games and competently performing the mental arithmetic task. Dreyfus concluded that chess (like all situated, skilled activity) is directed by mechanisms akin to face recognition and not the rules or algorithms of heuristic search. He concluded that connectionism might one day be able to offer an account of such behaviour if it takes account of embodiment. Dreyfus' work will be explored in more detail later.

Phenomenology, then, takes both behaviour and cognition to be fundamentally "first-person" subjective phenomena, inextricably connected to the environment. It represents a foundation in which the traditional sentential symbolic representational approach is not the only contender for a theory of agency.

It may be suitable at this point to revisit the concept of "routine activity", being essential to an understanding of this chapter. Agre (1997) states: "A routine is a frequently repeated pattern of interaction between an agent and its familiar environment." A routine is an identifiable set of

[1] Recall that an agent theory as construed in this thesis is one that addresses the question of *what an agent is and should be characterised as.* It is not, therefore, a panacea for engineering agents

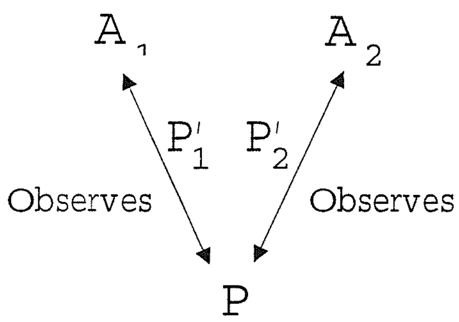

Figure 5.1: Outline of Phenomenological Interpretation

interactions between agent and environment, acknowledging the co-dependence on agent and environment in the definition of a routine. It is the establishing of these routines that informs the discussion of this chapter.

5.1.1 Defining Phenomenology

Phenomenology is concerned with "essences" of subjective experience (Merleau-Ponty, 1962). It sets the datum for analysing cognitive phenomena (of all varieties, including intentionality and therefore representation) to be the individual (e.g. agent). In its most ardent varieties, as discussed in later sections of this chapter, phenomenology refutes representational cognitivism but it should be noted that its foundations largely predate the cognitivist movement of the 1950s and 1960s.

Phenomenology is apparently at odds with the traditions of natural scientific method because a phenomenologist denies the existence of any objective world. Any construction of the world formulated is "polluted" by the individual making or articulating the construction. From (Nagel, 1974) and (Roy, Petitot, Pachoud and Varela, 1999) the position can be articulated as shown in Figure 5.1. In the diagram, two agents A_1 and A_2 perceive some phenomenon or property P. The agents perceive this world as P_1' and P_2' respectively. Using Nagel's example, let A_1 be a bat, and A_2 be a human. The extremity of the two exemplars is for illustrative purposes in that they fundamentally rely on radically different perceptual apparatus (one visual, the other aural). Nagel asks:

> "In so far as I can imagine this [perceiving the world as a bat] (which is not very far), it tells me only what it would be like for *me* to behave as a bat behaves. But that is not the question. I want to know what it is like for a *bat* to be a *bat*. Yet if I try to imagine this, I am restricted to the resource of my own mind, and those resources are inadequate to the task."

Nagel also suggests that no amount of subtraction or addition of "bat-like" apparatus or

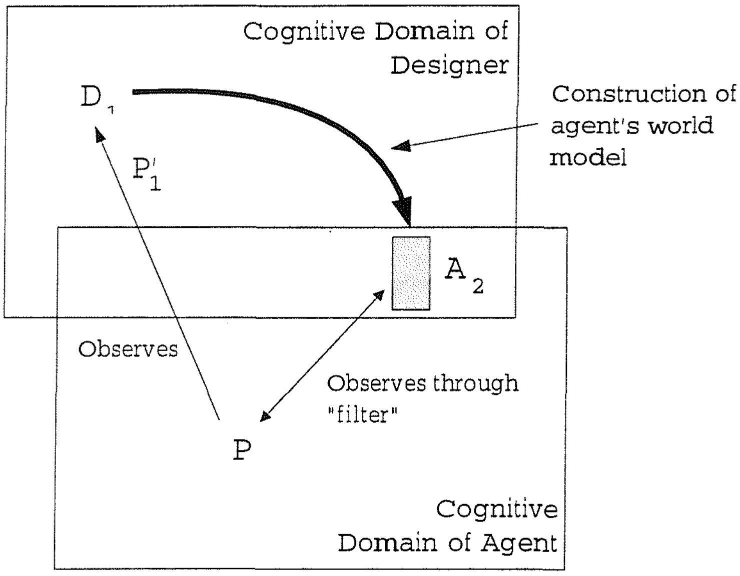

Figure 5.2: Traditional, Sentential-Symbolic View of Agent Theory from the Phenomenological Perspective

---

experience could aid him in experiencing the world as a bat. According to a phenomenological perspective, how the world is perceived by A_1 (the bat) is very different from how the human A_2 perceives its world. If we could enquire of both agents "What does the world look like?", the descriptions would be quite different. This is because the embodiment of each agent is radically different, as are their modes of engagement with the world. An ardent phenomenologist would argue that it is impossible for two human agents to be *identical.* To adopt a reductionist stance, the very connectivity and efficacy of each human's nervous system will vary quite considerably, encoding different socio-cultural conditioning as well as onto- and phylogenetic characteristics (examples of such phylogenetic differences being an inherited disposition to colour blindness or motor deficiency).

The ultimate phenomenology denies that $P_1' \equiv P_2'$. However, a more liberal interpretation might postulate that, to a certain approximation, if structurally $A_1 \approx A_2$ then $P_1' \approx P_2'$ – see for example, (Dretske, 1995). It is this fact that assumes when two people discuss an apple (assuming no significant cultural differences) that they agree on a similar (but never exact) definition or experience. Later in this chapter, phenomenological method (adopting the standpoint described above) will be examined more closely. In the following section, we can see how current agent theory meshes with phenomenology.

5.1.2 Phenomenological Method and Agency

Adapting the picture outlined in Figure 5.1, and supplanting one individual as the agent designer (a human) and the other as an artificial agent (e.g. a mobile robot or software agent)

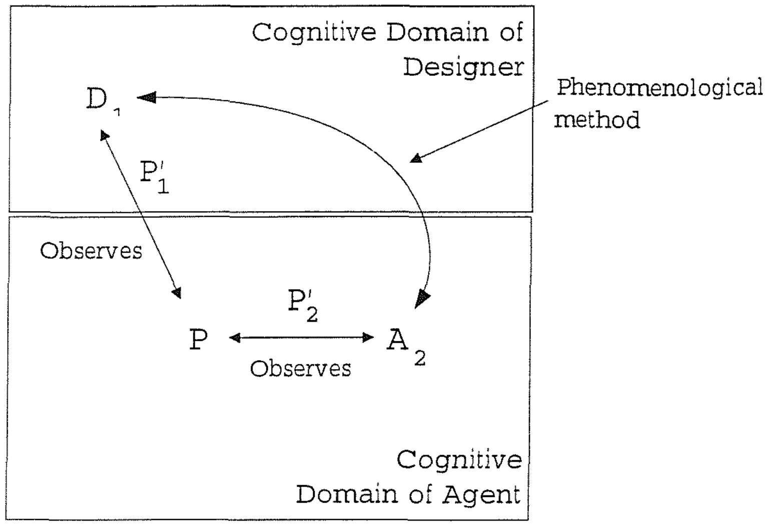

Figure 5.3: Phenomenological Method and Agent Theory

then the situation is described in Figure 5.2. Note that the cognitive domains (Maturana and Varela, 1980a; Winograd and Flores, 1986) of the agent and designer are overlapping. The designer interprets the world the agent will inhabit and then describes this in some way, such as by constructing a sentential-symbolic representation that describes the world, which is then built into the agent's architecture. The agent then experiences the world using an architecture which "filters" the world according to the designer's interpretation. This is analogous to building an artificial bat (with synthetic sonar) and then giving it a "blocks world" description of its environment that describes the world as visual features such as edges, intersections and geometric constructions. This approach would fail to acknowledge the phenomenology of the agent; more specifically, it assumes the embodiment of the designer is functionally relevant to the agent.

Figure 5.3 redescribes this process employing a phenomenological method; the construction of the agent is not guided by an articulated version of the world according to the designer. Instead, the construction of the agent is not in terms of an abstract formal ontology, but the architecture *itself* underwrites experience, perception and intelligent activity in the world. In Nagel's bat experiment, substituting "bat" for artificial agent and "designer" for Nagel, we have a similar scenario. A phenomenological method is required to enable interpretation of the agent whilst not divorcing it (or its structure or epistemology) from the world. Maturana and Varela's phenomenological method will be discussed as a possible bridging principle later in this chapter. The kinds of structural variation exhibited by self-organising connectionist networks offer a method of allowing the agent to find and configure relevant functional states that enable routines of activity to be established (learned) and revised (unlearned or adapted).

5.1.3 Irreducibility of Phenomenology to Naturalism

It would seem (especially from Nagel's position) that science and phenomenology are mutually exclusive. Husserl, Heidegger and Merleau-Ponty were all sceptical of the "naturalisation" project (the term naturalisation refers to an attempt to unify the study of phenomenological ideas with those of contemporary cognitive science, which for Husserl and Heidegger would be the natural sciences). In Husserl's case, phenomenological experience was taken as an irreducible fact. Heidegger, more radically, considered that the ontological commitments of science fundamentally prevents the proper treatment of phenomenal experience. Nagel's position was based on the assumption that science cannot cope with subjective experience. As (Roy, Petitot, Pachoud and Varela, 1999) explain, Nagel's basis for rejecting natural scientific method is incorrect in that the method now acknowledges the epistemologically relational nature of objects and their properties. More recently a kind of "weak objectivity" has permeated physics in the light of the quantum paradigm and uncertainty principles (Roy, Petitot, Pachoud and Varela, 1999) pp. 17. Roy *et al*, consider it "myopic" to assume that a scientific explanation of subjective phenomena is impossible *in principle*.

5.2 Phenomenology and Situatedness

In traditional AI, the internal symbol is taken as an ontological category upon which mental processes and agency are founded, being the building block for intelligent activity. Both Husserl and Heidegger (and their recent adherents) start with a different ontology; a different building block upon which to establish the nature of intentionality (for Husserl) and situated activity (for Heidegger).

> Phenomenology is the way of access to, and the demonstrative manner of determination of, what is to become the theme of ontology. *Ontology is possible only as phenomenology*. – Heidegger, *Being and Time* (Translated in (Krell, 1996) pp. 82)

Heidegger's statement refers to the fact that the methods employed to discuss and define ontology (e.g. the scientific process of categorising phenomena) are inherently defined by the phenomenological properties of the agent (i.e. the embodiment and history of the agent) – in effect, articulation of ontological commitments is "polluted" by the agent's faculties, experiences and current state. Returning to Nagel's example, a bat's version of an ontology of the world must be relative to its experiences of the world.

The unit of analysis in phenomenological method is the individual, from which ontology can then be constructed. As (Dautenhahn, 1997) states, the autobiographic memory of an agent, and its "replaying" of this memory (e.g. by observable patterns in its behaviour) is dependent on its embodiment which also defines the nature of its experiences of the world. If an agent were able to linguistically articulate this autobiographic memory (as for humans), then it is fundamentally important that this embodiment is acknowledged in what is articulated as an account of the world.

Heidegger describes the agent as "pre-ontological". It is therefore a more verbose expression and a more philosophical principle common to the notion of *situatedness* – offered by (Jennings, Sycara and Wooldridge, 1998) as the discriminating line between the deliberation of older AI systems and agents.

However, it differs from the technical connotations of situatedness because phenomenology attempts to reconstruct the basic ontology of experience (which implies intentionality) from the individual. Whereas, situatedness in agent theory means that *a priori*, an ontology of the world is defined and implanted in the agent which the agent then carries as a representation of the world defined from the designers "objective" view of that same world. The agent's world model (often called its ontology) is also predicated on the ontological commitments of computer science.

5.3 Extant Agent Theory and Intentionality

Heidegger's work posits a 'Background', being a set of shared common unarticulated principles which shape everyday activity. Importantly, as (Dreyfus, 1992) observed, Background cannot be constructed and embodied as a set of rules and facts, regardless of the sophistication of the representational language (e.g. modal logic formulations of belief and desire).

This tendency is related to Frege's notion of *sense*, stated broadly as "an abstract entity that, roughly speaking, captures aspects of a representation that determine whether it is true. A sense is thus not itself a representation or a symbol of a component of a psychological mechanism; it is, rather, the content that those things might express." (Agre, 1997) pp. 227. Frege's 'sense' is a theory of cognitive content, however it is not mentalistic. Sense can still indicate or pick out objects in the world (it must do, in order to determine if a representation is true) but it should not be located in the agents "head". Agre (1997) argues that AI has a legacy where the distortion of sense to mean simple reference (e.g. some internal content represents by mere reference an object in the world) and its location in the agent's internal mechanism, is based on the historical relationship of Frege to model theory in mathematics.

5.4 Overview of Current Research

Phenomenology has been of particular interest to other workers in computer science for the past fifteen years. Notably, Winograd's revising of his earlier AI work in which articulated symbolic descriptions of background were predominant. The analysis Winograd and Flores give in (Winograd and Flores, 1986) is not sufficiently detailed for connectionist-based agency, in that it addresses higher level issues of computer construction, design and analysis from a phenomenological basis.

Maturana and Varela (Maturana and Varela, 1980a; Maturana and Varela, 1998) aimed to ground cognition (and hence epistemology) in some fundamental ideas about the organisation of living systems. What is common to Heidegger, Maturana and Varela is captured in the quotation of section 5.2. Maturana and Varela's work represents a fundamentally different approach, in that they aim to ground the whole of epistemology in biological and systems science. However, their acts of *distinction* and the notion of a *cognitive domain* being relative to the point of analysis meshes closely with Heidegger's overall position. In effect, Maturana and Varela's work is most significant to operationalising phenomenology in terms of system-oriented principles inherent in connectionist models.

In due course, semiotic principles are introduced in this thesis. The theories of Saussure (1986) and Peirce (1868) will be central offering an "in practice" analysis of the role of representation and intentionality. Instead of the more analytic methods of Dennett, Fodor and Dretske's work, pragmatists focus on accounting for everyday phenomena instead of trying to decompose or locate the origin of, such as, intentionality. (Sowa, 2000) Chapter Two describes how the "mental mediation" of signs (that is, signs that represent or reference something in the world) in Peirce's theory of semiotics necessarily involves a mediator, or observer. This coincides with Husserl's principle of intentionality as the thing that "directs attention to the object of perception" – (Sowa, 2000) pp. 62.

(Agre, 1997) describes a reconstructed theory of intentionality and representation which focuses on *functional indexical* representation. However, he and Chapman dispute the relevance of connectionism because "they preclude the use of pointers, and so they rule out virtually the whole tradition of symbolic programming" (Agre, 1997) pp.80. It might be said that Agre's work is a symbolic theory of indexical functional representation. What is attempted, and required, in this thesis is a corresponding connectionist interpretation. The implications for agency and connectionist models generally will be presented in conclusion.

Finally, (Preston, 1993) and (Dreyfus, 1991; Dreyfus, 1996) consider connectionism and its role in phenomenology. Both Preston and Dreyfus' insights are particularly pertinent as they offer a way of operationalising phenomenology for an agent employing connectionist principles, in that they are not committed to symbolic principles. Specifically, Preston reconsiders the idea that causal mechanisms might, in part, embody Heidegger's notion of Background. Dreyfus is considered for his evaluation of Merleau-Ponty's contribution from *Phenomenology of Perception* with respect to intentional arcs.

5.5 Origins of Phenomenology

Before proceeding, it is worth revisiting the definition of intentionality. The philosophical term *intentionality* refers to the ontological category of a state which possesses "aboutness" or "directedness". Dennett (1987) defines the terms as:

> Some of the things, states and events in the world have the interesting property of *being about* other things, states and events; figuratively, they point to other things.
> – (Dennett, 1987) pp. 240

Husserl, and Brentano before him, defined intentionality to be fundamentally irreducible, in the same way as an ontological entity in the natural sciences can be. That is to say, Husserl denied that (for example) intentionality can be broken down into constitutive parts which (in the case of neuroscience) we might one day find a neural correlate or locus for. (Smith, 1999) asserts that a contemporary analysis of Husserl's objections would be aimed at reduction of intentionality to either computational or causal explanation. Roughly, intentionality cannot be explained using the computational idea of representation which assumes the structural relationship between properties. Nor can it be reduced to *any* explanation which relies on cause-effect mechanisms. Husserl asserts that the fundamental natural scientific *ontology*, that includes cause and effect as the most primitive first class citizens, is fundamentally the wrong basis for analysis. For example, the causal relationship between hydraulic components is no more capable of explaining intentionality than is the virtual-physical machinery of computer science. Both are predicated on an ontology of natural kinds which includes causation.

Husserl's notion of intentionality is founded on his own ontology of *what there really is*. However, according to (Roy, Petitot, Pachoud and Varela, 1999), Husserl is sure that intentionality has an internal component. He describes two ontological categories of the "internal" which are the "exact" and "non-exact". The latter of these categories is the intentional content described as phenomenological (Dreyfus, 1991) pp. 2–3.

Whenever an agent experiences the world, there are intentional states which (in Husserl's proposal) are irreducible and fundamental to the ontology of science and phenomenology in the same way that causal relationships are taken as a definite ontological feature in Newtonian mechanics. Husserl's method is to debunk the existing (or background) assumptions of what he saw as natural science and replace them with an ontology including his phenomenological categories. In doing so, he hoped to re-frame science to include phenomenology.

This rejection of causality might seem ludicrous as it supplants one well accepted ontology with something completely different, which remains undemonstrated and oppositional to the ontology of scientific endeavour. In fact, complex system's theory (and consequently, multi-agent systems) rely on a notion of emergence, which has received little attention as an ontological feature of engineering or the sciences it draws from. To say that the behaviour of a system *emerges* from the individual agents is akin to accepting that emergence is an ontological category of its own. For a seminal discussion of this kind of epiphenomena see (Meehl and Sellars, 1956).

5.6 A Vertical Model of Heideggerian Phenomenology

The research for this section is drawn from three sources: (Dreyfus, 1991) who provides interpretative commentary on the work of Heidegger (1927), (Krell, 1996) which reproduces Heidegger's original work with commentary and (Winograd and Flores, 1986).

Heidegger claimed that there is a Background, which consists of a set of socially shared conventions which shapes ordinary, everyday routine activity (Winograd and Flores, 1986). Heidegger's opposition to Husserl was that there is something more fundamental than an internal intentional state. He argued (cf. Brooks, Agre and Wilson) that beneath all intelligent activity was routine coping skills devoid of the kinds of representational intentionality commonplace in AI.

Husserl (so Dreyfus claims) was content with a Cartesian explanation, considering intentionality to be something intrinsically mental and divorced from the material. Heidegger strongly refuted this. In either case, there is not an endorsement of the symbolic representational theories of GOFAI. Heidegger rejects representational intentionality where states refer to the world via mental mediation. Husserl, on the other hand, rejects naturalisation (reduction to the principles of the natural sciences). This seems to present an impasse which is addressed later in the chapter. Heidegger and Husserl had no access to recent work in cognitive science, hence, a re-evaluation of their contributions should be undertaken with respect to connectionism and agency.

It is proposed, following (Dreyfus, 1996), that the connectionist models of Chapter 4 would form a suitable substrate for adaptive Background.

5.6.1 Equipment and Understanding

Heidegger's notion of Background is exemplified by what he calls "equipment". In effect, the ultimate relationship between agent and environment is with objects. Heidegger claims that we don't encounter objects *as mere objects*, but rather the agent always encounters them with respect to a purpose and with some pre-conceptions. The ubiquitous example being the hammer. When grasping a hammer, Heidegger claims that the hammer is not encountered by its user as a sequence or complex of abstract intentional states which represent the object and its properties. Rather, the hammer is encountered as being part of a routine activity.

From Dreyfus' treatment, (Dreyfus, 1991) Chapter 4, we can extract some principles central to this thesis. Dreyfus states that:

1. objects (equipment) are understood by that agent using what Heidegger refers to as *manipulating*, which gives us *primordial* understanding. This is, therefore, the primary way of encountering objects and is akin to sensori-motor grounding.

2. "second hand" understanding refers to knowing *what a hammer is*, without having ever used one (perhaps, knowledge gained through communication). Heidegger termed this

"positive" understanding.

A similar procedural/declarative split may be seen in cognitive science (know-how *versus* know-what). Cangelosi, Greco and Harnad (2000) explore the distinction between sensori-motor toil and symbolic theft. The former is the pre-linguistic knowledge that is acquired by an agent through directly experiencing the environment, which is analogous to the Heideggerian notion of primordial understanding. The latter is the kind of information gained by symbolic interactions e.g. linguistically, and is akin to positive understanding. The importance of this connection is that Cangelosi, Greco and Harnad's work firmly establishes the symbol grounding argument (Harnad, 1990) and demonstrates that sensori-motor toil must precede symbolic capability.

5.6.2 Dasein and Circumscription

In Heidegger's phenomenology, there is a sophisticated nexus of concepts about ontology and the world. Fundamental to this is what Heidegger refers to as *Dasein* which is the backdrop against which all else in the agents phenomenology is revealed exists, and any ontological discussion is based. Background is, in effect, Heidegger's fundamental unit of analysis. This has profound philosophical implications, so much so that Heidegger's work deliberately constructs a language vastly different to the usual modes of discourse in philosophy and science. The resulting "modes" which Heidegger elaborates have real practical implications for a theory of agency.

Heidegger's pre-ontological[2] method conjectures that there is a mode of *absorbed coping* which underpins all intentionality. By this, he states that primordial understanding is necessary for everyday routine activity. It might therefore be stated that this encapsulates everything done without any recourse to "thinking about", cogitating or deliberating. For example, the routine activity of locking one's house as one leaves is an example of a sequence of routinised activity which requires only occasional recourse to deliberation, but *may* invoke deliberation when expectations fail. During such activity an individual is not aware (in the sense that one entertains explicit intentional states which are revealed during the process of locking the door) of objects such as keys, doors and locks. Moreover, individuals do not have representations in the way GOFAI and some tenets of cognitivism would propose. The subject/object split is *not* available to the agent during circumscriptive activity whilst embedded in the world. A more vivid example is a *cane* used by a visually impaired person. When in use, the cane is not presenting itself to the user as a long, narrow cylindrical device but is featuring pervasively in cirumscription, and without direct, descriptive (sentential-symbolic) representation.

Heidegger describes examples of activity, such as those stated, as "nonthematic circumscriptive absorption", where circumscription is the name for such everyday embedding of equipment in activities (Dreyfus, 1991). The computational hypothesis that situatedness is key

[2]Heidegger's term for something that precedes and must exist for ontology to be possible

to agency is supported, however there has been no consideration of what that means for representation and intentionality. Fundamentally circumscription is more than *just* direct interaction with the environment, it requires a theory of representation or, more broadly, intentionality.

This notion of routine activity is therefore of paramount importance to agent theory. The implications may be said to be:

1. during circumscriptive activity, equipment is *transparent*. For example, when typing a thesis, the author is not concerned with any representation of the keyboard, it merely exists as equipment in a nexus of related activities, objects and functionalities within the author's world.

2. that a logical extrapolation of Heidegger's phenomenology is that a majority of time is spent in activity which is not reflected upon (in a deliberate fashion) as it is being carried out. Fundamentally, most activity is routine and circumscriptive.

Studies of routinised work-behaviour (for example, the work of Scribner (1984) on factory workers) and practical use of the formal algorithms taught as basic mathematics and numeracy (Lave, 1988) have revealed a common basis; the routinisation of activity is dependent on the environment and the agent mutually shaping each other. For reasons of cognitive efficiency, recourse to the formal descriptions given in work-practice manuals or numeracy textbooks seem to play little role in the day-to-day activities of people's behaviour (except, as will be argued later, in the case of breakdown of routine activity). Hutchins (1993) quotes Vygotsky's postulate that higher mental functioning is a social process, not one emerging from internal mechanics. The internalisation of experience shapes what is internalised which subsequently affects the reproduction or articulation of the social process (see also (Hutchins, 1995) for elaboration).

These last points appear to deny the very kinds of things (e.g. representational mechanisms such as hand-writing or diagramming) that enables deliberative thought. However, Heidegger strives to provide a mechanism which provides for a "layered" kind of activity. Before presenting an investigation of this in the next section, a final important addition to this discussion is that of *comportment*. Heidegger frames cognition and understanding in the service of how an agent might exist with equipment in the environment. The "dealing with" and cognitive aspects of perceiving and sensori-motor tasks that facilitate equipment use are called *comportments*. This is important because comportment must be adaptable if the agent is presented with a changing environment. Since Heidegger's work predates cognitivism, there is little discussion of constructs such as schema theory (Arbib, 1989) and it therefore difficult to say if there are isomorphisms. A schema being an abstract entity which captures the routines and idiosyncrasies of a particular capacity for situated action. For the purposes of migrating phenomenology to agent theory, the idea of a schema is congruent. However we should be careful when importing the representational baggage that accompanies it.

For a given situation (i.e. perceptual state) and available motor activity, a schema acts as a template for co-ordination of some set of actions. Schema are activated by situations and produce behaviour, so are both a causal and explanatory mechanism that exist as internal representational intentionality (Arbib, 1989) chapter 5. More importantly, comportments and schemata must be adaptive. If for example, the world were suddenly devoid of hammers, comportment must adapt appropriately so that another object fills the place of a hammer in routine activities such as driving nails.

This now reveals another facet of phenomenology: the way the world and objects are perceived (revealed) to the agent through routine circumscription.

5.6.3 Using Revealing and Equipment in Agent Theory

After arguing that equipment is transparent to the agent, and that comportment is adaptive, we need to integrate comportment and circumscription in a meaningful way so that this adaptive notion can be given leverage.

Returning to the hammer example, assuming it has two heads (facing in opposite directions) then circumscriptive activity will not require one to pay attention or entertain representational intentionality "about" the hammer.

The proposal now proceeds by showing how Heidegger's phenomenology and routines of activity (circumscription) can be built into a vertically organised agent theory which posits a number of levels of intentional (or representational) "engagement". The term mode of intentional engagement is introduced here to signify the connection of each vertical level to a relevant perspective arising from the traditions of AI (e.g. symbolic representation) through to this thesis' proposal that routine circumscription can be implemented in connectionism. For this realisation to be possible, the notion of *adaptive comportment* is also introduced to enable the theory to incoporate recovery from failures in circumscription (by adapting the substrate of circumscription, which here is connectionist architectures).

5.6.3.1 Malfunction

From the hammering example, if during this absorbed activity the head being used to drive nails breaks, then Heidegger argues for a disturbance to circumscription called *breakdown*. It can be assumed that this startles, or draws the, attention of the agent. However, the agent turns the hammer about its handle's axis, and continues. This sequence can be described as :

1. Absorbed activity (*nonthematic absorbed circumscription*) – hammering

2. Equipment fails (one of the heads is broken), revealing itself to the agent. Heidegger says that the agent becomes aware of the *assignment* of the hammer to the task at hand and enters a state of *malfunction*

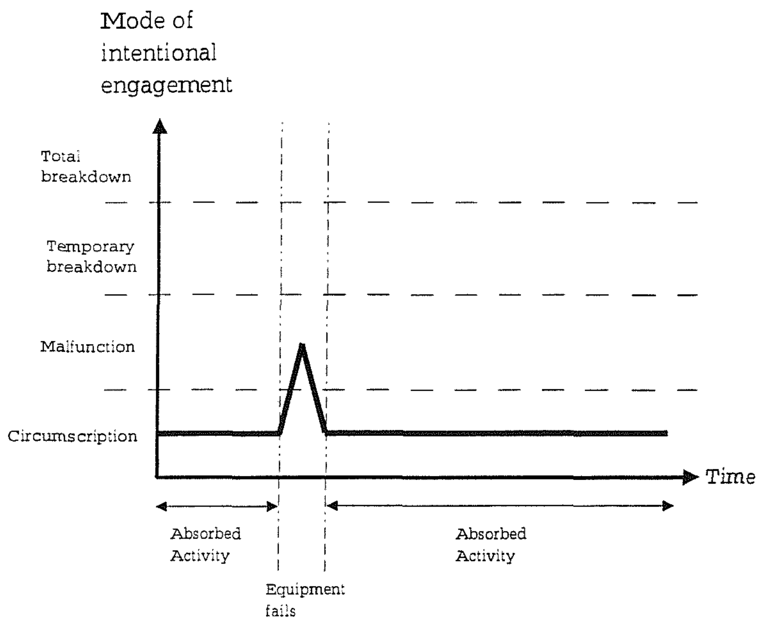

Figure 5.4: Malfunction and Circumscription

3. The agent recovers by correcting and engaging in absorbed activity. The agent might rotate the hammer (correcting the problem) and then re-engages with the same absorbed activity, with the equipment becoming transparent

This is diagrammed in Figure 5.4. The diagram shows time on the horizontal axis and the mode of intentional engagement on the vertical. The latter simultaneously describes the severity of breakdown and the corresponding intentional states which must be engaged to recover. Heidegger's hypothesis is therefore that intentionality as classically conceived is parasitic on circumscription and is prevalent in the agent's activity, as breakdown becomes more severe. Trivially, we might say more abstract representational intentionality is required as we become further removed from the norm of everyday routine activity. Or, to use Arbib's schema theory, when a schema instance fails, another schema instance takes over.

Using Heidegger's principle of malfunction, what is proposed here is that schema-like mechanisms underpin routine behaviour. For example, a schema such as the current weights (context invariant representations) of a network implement the routine action, but produce an undesirable outcome. The network's connectivity is adapted to restore circumscription, for example, by a small change in a number of weights in the reinforcement learning and/or perception network. This proposal of "adaptive cricumscription" is added to Heidegger to enable an agent theory and implementation.

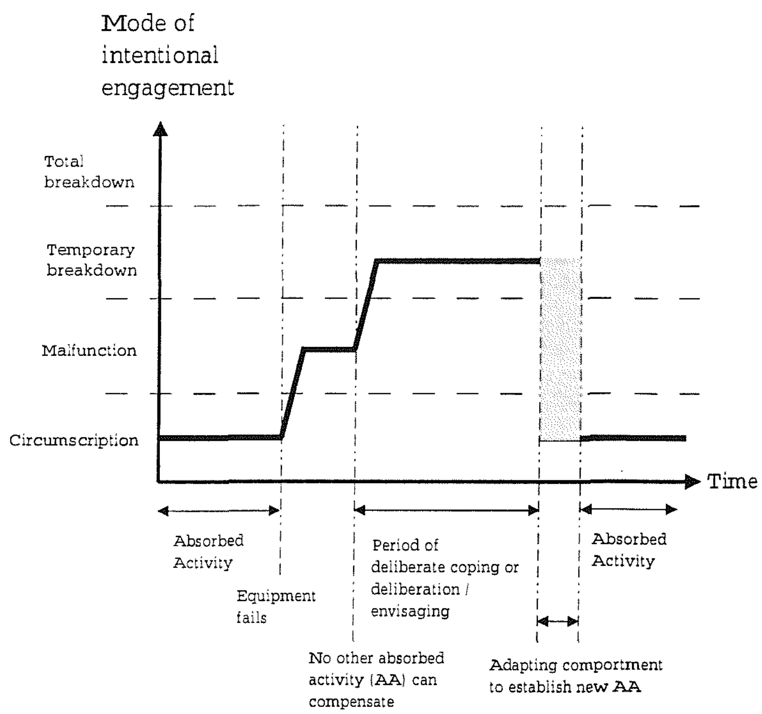

Figure 5.5: Temporary Breakdown and the Re-establishment of Circumscription

---

5.6.3.2 Temporary Breakdown

If, during the sequence of events described above, the agent cannot easily recover from the failure (for example, both heads of the hammer have broken or the handle has fractured) and resume absorbed activity, it is forced into *temporary breakdown.*

This is shown in Figure 5.5. For clarity, the diagram does not differentiate between Heidegger's two classes of temporary breakdown. Firstly, the agent enters *deliberate coping* where, for example, another hammer is searched for. If the agent can continue (e.g. a second hammer is close by) then the problem is resolved without recourse to the second class of temporary breakdown – namely *deliberation.* This more severe form of temporary breakdown forces the agent to engage in behaviour which, in technical discourse on agency, is almost the paradigm example of *computationally reactive* agency.

> "The scheme peculiar to deliberating is the "if-then"; if this or that, for instance, is to be produced, put to use, or averted, then some ways and means, circumstances, or opportunities will be needed" – Heidegger, *Being and Time*, quoted in (Dreyfus, 1991) pp. 72

This meshes with the Pengi system of (Agre and Chapman, 1987) and more precisely with the constructions of "circuits" and the dependency maintenance systems of (Agre, 1997). In Agre's system, a vertical architecture is used to build circuits of "if-then" contingencies which, in common with activation spreading action selection systems, fire when their pre-conditions are met (by pattern matching).

The diagram of Figure 5.5 furthers the proposal developed here. Heidegger does not provide an explicit description of the recovery from breakdown, at least not in a way amenable to computational studies of his phenomenology. It is proposed that in order for the agent to return to circumscriptive activity, it must enter into a period of oscillation between temporary breakdown, malfunction and circumscription that *includes* adaptation of comportment. This is shown by a grey region on the diagram of Figure 5.5.

In connectionist agents, this recovery will take the form of repeated trials of behaviours and consequent adaptation. While this adaptive effort may be mechanically similar to that for the proposal for restoring circumscription after malfunction (e.g. the same learning algorithm is used), it represents a qualitative difference in the behaviour of the agent. The agent tries out *novel* behaviours because the breakdown indicates sufficiently different rules of engagement between it (the agent) and the environment. In classic machine learning terms, we might say that this event and breakdown occurs is when the environment's state transition probabilities shift significantly, a longer period of trial-and-error learning is required to adapt to the new demands and establish new routines[3]. The use of temporary breakdown to build a layered architecture for such environmental shifts is given in Chapter 9, section 9.3.

Also, if we reach deliberation *after* passing through deliberate coping (e.g. failing to retrieve a hammer from a nearby toolbox) then the agent's environment is revealed to it by necessary and sufficient intentional states that enable long-range planning (what Heidegger refers to as *envisaging*).

It is proposed that this must also involve some adaptation of routine absorbed activity. For example, an agent might decide, in deliberation, to take a ready to hand lump of iron and employ this as a substitute hammer. Assuming that, at some time, this substitute equipment is used and engages in absorbed activity, there must be some transition from the deliberation (temporary breakdown) that resulted in absorbed activity being reinstated.

5.6.3.3 Total Breakdown

If circumscriptive activity cannot be reinstated, the agent has recourse to the mode of *total breakdown* which involves a form of detached reasoning traditionally associated with research in AI and agency. Heidegger describes this as *thematising* where:

> "Its aim is to free the intraworldy entities we encounter, and to free them in such a way that they can 'throw themselves against' a pure discovering – that is, that they can become 'objects'. Thematising objectifies." – Heidegger, *Being and Time*, in (Dreyfus, 1991) pp.82.

[3]Again, such mechanical precision is not admitted in Heidegger's analysis but is included here to crystallise the arguments and move forward the proposed agent theory to architectural principles

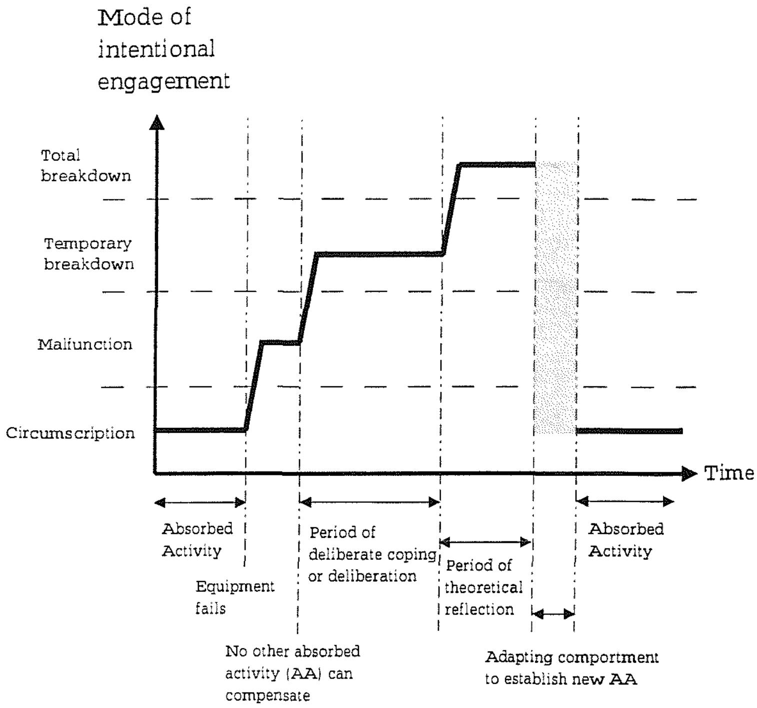

Figure 5.6: Total Breakdown and the Re-establishment of Circumscription

---

Figure 5.6 describes the progression to total breakdown. At this stage of detachment from the environment, the agent is able to conceive of an ontology of objects and reason about them as detached objects. That is, understanding (cf. primordial understanding) is not grounded in embodied action. To objectify and conceive of a hammer, one need not be in the agent's grasp or visual field. At this level, it is proposed that the traditions of AI might mesh in a vertical, hybrid architecture where connectionism implements circumscription and comportment, a generalised dynamical system (acting in some kind of executive role) might influence the behaviour of the connectionist architecture, and finally, a symbolic system (parasitic on the other systems) provides for long-range planning (thematising). An instantiation of such a theory in a concrete, application domain is given in Part IV of this thesis.

Beyond the thematising stage, Heidegger moves from such contemplative reasoning to *pure contemplation*, which is the domain of philosophy, metaphysics and true ontology (that is, the kind of reflective thinking that perhaps Heidegger is engaged in when writing *Being and Time* and when conceiving of Dasein as the fundamental unit of analysis).

An important technical overlap exists at this level; the designer of an agent must surely be involved in pure contemplation – wondering what objects might best be featured in the predicate-based description of an agent's ontology. The agent utilising such an ontology cannot be beyond total breakdown, because on the whole, such contemplation is the exclusive domain of a designer with the capability to engineer such an agent. It is assumed that any reflection and self-modification of a set of predicates expressing an agent ontology is purely mechanical, and really does not involve the kind of reasoning Heidegger is clearly alluding to.

An equally important conclusion to this description is that Heidegger constantly emphasises these "higher" levels of intentionality as being:

- *grounded* in everyday activity – that is the fundamental, primordial way of understanding the world is by absorbed circumscriptive activity.
- *any* deliberation, cogitation or reasoning which is conducted as an incremental detachment from the norm of absorbed activity.
- parasitic on passing through other levels of engagement in order that the world be revealed to it; the actual revealing of the world (to the agent) is strongly dependent on the circumstances that lead to the breakdown of activity.

Heidegger strongly concurs with Wilson (1991) that the starting point for what we know as higher level cognitive ability is lower order, routine embodied activity *in the absence* of reflection. This is perhaps the most important conclusion to be taken from Heideggarian phenomenology and subsequently imported into technical practice in constructing computational agents. However, some stronger mechanical assumptions are necessary to bridge between a plausible implementation and Heidegger's phenomenological stance. These mechanical assumptions will be explored later in this chapter.

5.6.4 The Problem of Externalised Representations

One area which defeats the above analysis is when positive understanding takes the form of some kind of object in the world which is not utilisable. That is, when sensori-motor toil cannot be engaged to ground the symbolic thieving of concepts (Cangelosi, Greco and Harnad, 2000). The classic example being a sign of some kind – or any form of externalised knowledge communicated as an artifact such as a document, a road sign, car indicator light and so on. Heidegger is sceptical of the pragmatist movement of semiotics, since it takes the relational notion of *sign* as a *signifier* for some *object* as the basis of ontology. Heidegger takes Dasein as the basic unit and everything else is considered parasitic or derivative.

Harnad's technical proposal for overcoming this is to suggest that an atomic symbol is connected via sensori-motor toil, and then compositional structures of grounded atomic symbols form the basis of higher-order concepts (Harnad, 1992). For example, Harnad gives the grounding of 'zebra'. We assume that a concept *horse* is grounded in some sensori-motor, iconic representation (again, congruent with cognitivism, this is posited as internal and representational, but not necessarily sentential) as is the concept of stripes, or patterned, textured visual stimuli. However *zebra* is not grounded, so the agent would proceed to represent a zebra, of which it has no direct experience, as the concatenation *horse* + *stripes* (cf. (Sharkey and Jackson, 1994)).

Heidegger is burdened with explaining the phenomenon of signification in a framework which neither had access to, and in fact refuted, the approach advocated by hybridisation of cognitive architectures manifest in Harnad's work. Heidegger's approach is to take signs as another form of equipment. A sign in the semiotic tradition, being anything that stands in relation to something else by means of representation. The agent engages with signs in exactly the same way as it would with any other equipment. He argues that signification is simply part of the nexus of associations between equipment. So, a picture of a hammer refers to a hammer by being equipment, with the particular property of its function being to *indicate a hammer*. Just as a lever is a tool, with the property of being able to move massive objects, a sign is equipment with the property of being able to draw attention to the physical artifact it (for example) looks like.

This is congruent in as much as denoting a sign as equipment means it must, like all equipment, be part of a complex nexus of association in the socially shaped practices of absorbed activity. In essence then, Heidegger's signification is the *function* of the equipment "a sign" just as the function of a hammer is to exert force on nails. Semiotics is therefore a powerful tool for understanding agency, as is elaborated on later in this chapter.

5.7 Husserl and the Symbolic Paradigm

Husserl's contribution is harder to establish in relation to agency, as little work exists on the subject. Recently, (Petitot, Varela, Pachoud and Roy, 1999) have attempted to provide a "naturalisation" of Husserlian phenomenology, but the significance and scope of the project is far wider reaching than the study of agency. However, the analysis of (Roy, 1999) contributes a Husserlian perspective to this thesis in that it isolates intentionality and focuses on competing theories for a computational (or more widely still, a naturalisation) explanation. Roy (1999) explains why a symbolic account is inadequate. Using Fodor's language of thought and its implicit commitment to symbolic, internal intentionality, he states:

> "Now, what Fodor seems to have forgotten is precisely the fact that, inasmuch as it [the symbol] is something standing in for something else, a symbol pre-supposes an element x – an *interpretant* in Charles S. Peirce's terminology – as the *source* of the symbolic relation." pp. 124.

A symbol is therefore an intermediary thing, which requires a referrent. In attempting to naturalise Husserl, Roy is introducing the necessity of signification. While Husserl's work is surely more relevant than simply to debunk symbolic theories, such as language of thought, it remains a question for further research.

5.8 Using Merleau-Ponty's Intentional Arcs as the Basis for Adaptive Comportment and Circumscription

Merleau-Ponty (1962) proposed the notion of an intentional arc, being something that enables and underpins an agent's ability to perceive and act in the world without employing any representation of goal directedness (in short, without reference to a consciously active description cf. the linguistic/logical forms of traditional AI). Especially, he wanted to identify the intentional arc (like Heidegger) as more than just internalised individual experience:

> "... that the life of consciousness – cognitive life, the life of desire or perceptual life – is subtended by an 'intentional arc' which projects round about us our past, our future, our human settings, our physical, ideological and moral situation or rather which results in our being situated in all these respects."– (Merleau-Ponty, 1962) pp. 136.

The most important part is the situated nature of that arc and how it intimately connects with the embodiment of the agent, for example, its specific cognitive faculties and motor abilities. From Merleau-Ponty's work, Dreyfus (1996) proposed that skillful, purposive goal directed behaviour in the form of intentional arcs can be captured in a neural substrate. However, Dreyfus stresses the point that connectionist methods need careful consideration.

According to Merleau-Ponty, embodiment is the mutually affective relationship between perception and action. However, he stresses, indirectly, the role of the sensori-motor loop cf. (Sharkey and Heemskerk, 1997). Discussing routines of absorbed activity he states:

> "In fact every habit [cf. routine or skill] is both motor and perceptual, because it lies, as we have said, between explicit perception and actual movement, in the basic function which sets boundaries to our field of vision and our field of action" – (Merleau-Ponty, 1962) pp. 152.

In Figure 5.7, it is proposed that there are at least three conjectured "stages" of an intentional arc formation – that is to say, the acquisition of skill in manipulating the environment. Merleau-Ponty's notion was not as concrete, nor was it intended to be, but here an attempt has been made to unify Heidegger's notion of breakdown with skilled practice. The assumption therefore being that breakdown and malfunctions are more common during acquisition of skill (e.g. while comportment adjusts to enable circumscription).

Recall Figure 5.4, where circumscription, malfunction, temporary and total breakdown were shown as "levels" or modes of intentional engagement. At total breakdown, intentionality is most closely associated with the sentential-symbolic ideals of AI. However, at the level of circumscription, these are inappropriate. The question then arises about what an intentional substrate might be. Dreyfus proposes that the intentional arc might be implemented by a neural

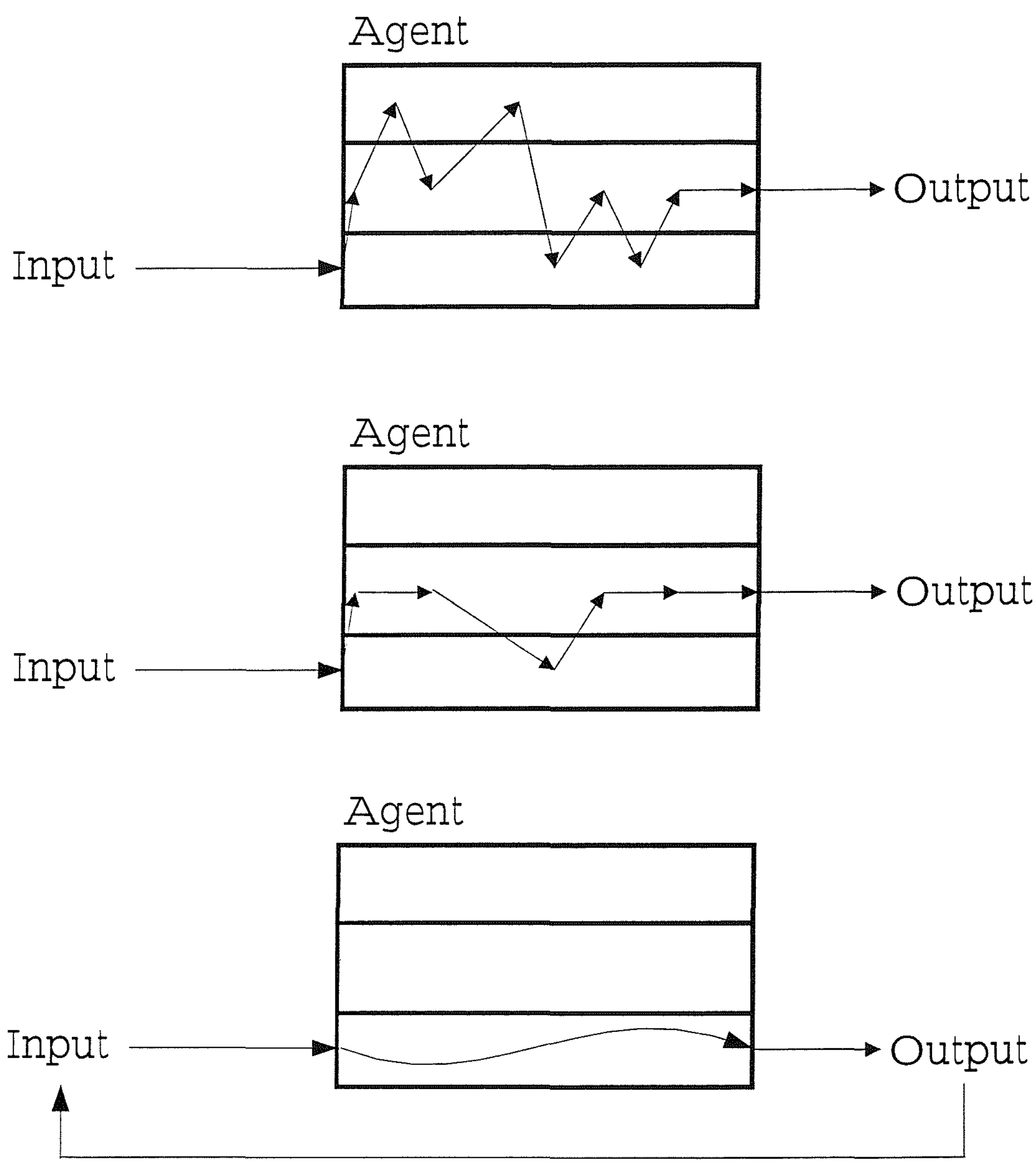

Figure 5.7: Schematic of Intentional Arcs (see text for explanation)

substrate. That is to say, the particular representations which lie beneath intentionality during routine activity, are not of the overtly propositional attitude form; 'x desires that y'.

Figure 5.7 (top) shows an agent as possessing three levels; circumscription, malfunction and temporary/total breakdown[4] analogous to those of Figure 5.4.

The agent encounters the "input" at the level of circumscription e.g. during routinised activity. This input disturbs circumscription, because the input is, for example, novel to the agent in some way such as if the hammer head breaks, or the expectation of some equipment-to-hand fails because the agent has never used it before. The agent is therefore unskilled to cope with this scenario as presented. It is suggested that this causes some vertical oscillation between malfunction, temporary and total breakdown, for example the agent might try and find a way to cope with the situation, fail and resort to higher-level detached reasoning characteristic of envisaging. This, as (Dreyfus, 1996) reminds us, is a phase of skill acquisition, requiring attention and commanding effort. But still, during this process, "Movement is not thought about movement, and bodily space is not space thought of or represented" (Merleau-Ponty, 1962) pp.137. All the while, Merleau-Ponty thinks that any representation (intentional arc) is consciousness "being-towards-the-thing through the intermediary of the body." and that "A movement is learned when the body has understood it, that is, when it has incorporated it into its 'world', and to move one's body is to aim at things through it; it is to allow oneself to respond to their call, which is made upon it independently of any representation" pp. 139. This alludes to deictic representation, which is returned to in Chapter 9.

In Figure 5.7 (middle), the agent has begun this process of adaptation, acquiring some level of skill when the same input recurs. It still forces the agent to occasional temporary breakdown or malfunction – e.g. to use Dreyfus' example, a novice driver might occasionally have to think about the order of pedals in the car. Finally, Figure 5.7 (bottom) shows the intentional arc as something that represents skilled competence (expertise) and that no longer disturbs circumscription. The agent merely reacts, with expertise acquired over time.

5.8.1 Incorporating Intentionality and Motivation: Homeostasis in Agents

Merleau-Ponty emphasises that goal directed behaviour requires a motivation, however that 'motivation' is not represented explicitly. For example, in an agent's everyday activity, it will not suddenly become aware of a desire to achieve a certain goal, e.g. to find food because of hunger. The motivation will emerge in more a fluid fashion through activity, by the deviation of a certain situation from an equilibrium ideal e.g. when certain chemicals such as sugars are so low that a motivation to eat emerges. Congruent with Heidegger, a verbal articulation of this property will be possible by total breakdown when long-range planning (*envisaging*) becomes necessary because of the *unreadiness-to-hand* of food.

This provides further suggestions for the environment and embodiment of the agent being

[4]for clarity, coarser-grained levels are shown

crucial in determining a structure for the agent's internal mechanism. For example, in an environment without water, an agent would be unlikely to be biologically dependant on water, and hence the environment, agent internal structure, and therefore emerging goals, desires and intentions would be very different to that of an agent subsisting on water.

In effect, to propositionally define a goal requires two factors. Firstly, the assumption that the agent's designer is *interpreting* and *articulating* the proposition. Secondly, that the designer is doing so based on some *observation* of the mechanisms of the agent and its environment (its embodiment) and in relation to the designer's own intention. Merleau-Ponty (1962) pp. 50 articulates this in a similar fashion; the motivated phenomenon (eating food) is explained by the motivating internal phenomenon (sugar levels plummeting).

5.8.2 The Intentional Arc and its Neural Substrate

In describing Merleau-Ponty's phenomenology, it becomes clear that it has much in common with earlier phenomenology. What is different, is (as (Dreyfus, 1996) points out) how an account of how memory and perception are integrated with action. In Chapter 3, it was argued that connectionism (as an *in principle* reactive substrate for agency) *does* incorporate a theory of long-term state cf. (Jennings, Sycara and Wooldridge, 1998; Wooldridge, 1999). Dreyfus agrees, and illustrates the relevance to phenomenology:

> "Neural networks provide a model of how the past can affect present perception and action without needing to store specific memories at all. It is precisely the advantage of simulated neural networks that past experience, rather than being stored as a memory, modifies the connection strengths between simulated neurons. New input can then produce output based on past experience without the net having to, or even being able to, retrieve any specific memories." – (Dreyfus, 1996) para. 48.

Despite enthusiasm, Dreyfus remains sceptical because of the explicit need for a teacher (supervisor) in most neural networks research. He alludes to reinforcement learning (para. 52), while leaving open the nature of the kinds of connectionism which might implement intentional arcs. One of the aims of this thesis is to explore those properties of agency and connectionism which achieve what Dreyfus alludes to. For now, the salient fact is that human construction of input/target pairs for a multi-layer perceptron trained by back-propagation of errors, is not situated enough to justify their role as a manifestation of neural intentional arcs.

The key facts of intentional arcs, in Merleau-Ponty's theory, is that they are *acquired* by situated practice. They are also influenced by socio-cultural factors, but fundamentally, they must be grounded in practice, becoming "shortcuts" from perception to action. They are the intentional mechanisms of Heidegger's circumscription. It is this fact which strongly motivates the unification of agent theory with connectionism.

The nature of the environmental feedback given cannot be general pain/pleasure indicators. Merleau-Ponty insists that it will be indicators which enable the agent to achieve equilibrium in coping with its environment being a broader class of feedback than the controlled scheduling of behaviourist conditioning. To this end, a more sophisticated derivative of reinforcement learning will be developed in Chapter 8.

In Chapter 9, Dretske's solution to functional intentionality will be examined, which further emphasises the role of learning and adaptation in acquiring intentional arcs. Dretske's proposal is challenged by Fodor and Dennett *because* of the learning/performance divide in the training of connectionist machinery. However, this argument *can* be defeated if agency and connectionism are united and the type of connectionist machinery employed is appropriate to Merleau-Ponty and Dreyfus' demands.

5.9 Autopoietic Systems and Embodiment

At this stage, Heidegger provides the superstructure for a phenomenology of agents. Merleau-Ponty provides the cognitive notion of intentional arcs (e.g. an intentionality of action) but have yet did not provide a method of operationalising these concepts. To this end, Maturana and Varela (1980a, Maturana and Varela (1998) provide an independent, but similar, body of work[5]

The foundational principle of *autopoietic* systems is that cognition and experience are biological phenomena, and that a system engages with its environment in such a way as to maintain an equilibrium which guarantees its existence. Maturana's earlier work on frog vision (with Lettvin, McCulloch and Pitts) is grounded in neuroscience and forms the basis of an abstract neural theory of phenomenology.

The most important contribution to this thesis, is that it is overtly embedded in phenomenology (although does not suggest connections) and furnishes an interpretation of a "unity" (which is analogous to an agent) where behaviour is explained with reference to an observer and the unity's nervous system and embodiment. While the connection between Heidegger and autopoiesis was first explored in computer science by (Winograd and Flores, 1986), their analysis utilises the general systems theory of autopoiesis to discuss linguistic competence and human-computer communication. However, unlike in this thesis, the specific causal principles underpinning *enactive agent cognition* (Varela, Thompson and Rosch, 1991) were not explored.

5.9.1 Unity and Agency

If we imagine a "soup" of animate, autonomous "things" which exists without ever being observed or studied, and then an observer comes to study and observe the "soup", Maturana and Varela claim that at the very instigation of the observation, a cognitive act of *distinction* takes place (Maturana and Varela, 1980a) pp. xix. This means, that there is an ostensive act of in-

[5](Maturana and Varela, 1980a) is a collection of (Maturana, 1980) and (Maturana and Varela, 1980b) with prefaces and editorial. Page numbers given refer to the 1980 collected works.

dividuating the agent that separates the agent (the *unity*) from its *background*[6]. To examine the soup, and identify individual entities, is to identify unities as something *distinct* from their background. From the agent's perspective, this is causally inert. Identifying an agent in a complex system is something an observer does, and irrelevant to the functioning of the agents themselves.

An agent in its environment (unity and background respectively) is a categorisation present in the epistemology and cognitive domain of the observer. Maturana and Varela further claim that this can be a recursive process. After picking out a unity, closer examination might reveal it consists of a complex, or composite, of other unities. They stress that this decomposition is still observer dependent.

A *composite unity* is an agent where (recursively) the observer has identified subordinate mechanisms, each one individuated from the background of whatever the agent's internal mechanism consists. As an example, a thesis is a composite unity, where its constituents are simpler unities called chapters. The structure and organisation of chapters is essentially linear, so we can say that a thesis (a composite unity) has constituents called chapters, and the property of linearity makes this complex unity a member of the class "thesis". The important issue is that the properties which denote the organisation and structure of a composite unity define its class membership. Should these properties no longer be applicable, then the unity is no longer part of the class. For example if we remove the property of linearity, then a collection of chapters in random order is no longer a thesis.

5.9.2 Structural Coupling

Maturana and Varela say that an agent maintains itself by adapting to disturbances in the relationship between itself and the environment (referred to as *unity* and *medium* respectively in their terminology). The term *structural coupling*, refers to the relationship between the agents structure, i.e. the structure of composite constituents, and the environment. Structural coupling has two components, ontogenetic and phylogenetic. Ontogenesis is effectively the adaptation which occurs during the agents lifetime, whereas phylogenesis is the longer term evolutionary history of the agent. In direct congruence with Heidegger, they stress that the agent is acting in the present, without conscious regard for its past or future. Such awareness (cf. temporary breakdown of circumscription) is afforded only through retrospection.

The principle most illuminated by structural coupling, is that the agents are both effectors of perturbations and the recipients of the consequences of these changes. In effect, the agents 'shape' the environment and the environment (mutually) 'shapes' them. This occurs cognitively, behaviourally and ontologically from the perspective of the observer. The environment is often said to select agent's according to their level of adaptation to the environment. Agents

[6] In Maturana and Varela's work, background refers to the "medium" which is similar to the Heideggerian notion of Background

that do not adapt become some other class of unity.

Similarly, a quantitative description of structural coupling (as key to embodiment) was elaborated upon by Quick, Dautenhahn, Nehaniv and Roberts (1999). They defined the embodiment of an agent in an environment as being the existence of two channels; one where the agent affects the environment and another where the environment affects the agent (and its internal structure).

5.9.3 Cognitive Domains

A cognitive domain is a phenomenal domain. For any given unity, i.e. the agent or any constituent part of the agent, there is a triadic relationship between the environment, the observer and any distinct unity. For example, one might describe an agent as "seeing" a tree. This defines a cognitive domain and, as will soon be argued, an explicit instantiation of a semiotic mediation between observer and the agent. If the observer then picks out a constituent unity (in this thesis, these are artificial neurons) then what it "sees" is defined by a different cognitive domain. Throughout this cognitive domain is a semiotic triple of observer, event and unity.

The implications of this idea are profound. What an observer perceives as an agent encounter is, as Nagel explained, not what the agent perceives. In this way, any one of the agent's artificial neurons is perceiving something different. Take, as an example, an action potential building as the result of antecedent neuron's spiking behaviour. Only when the concomitant neuron's signalling is correlated with an observation do we have ascribed meaning or denotation of internal state to the world, and this necessarily includes the observer. As Maturana and Varela (1980a) state:

> "... the answer to the question 'What is the input to the nervous system?' depends entirely on the chosen point of observation." – pp. 22.

5.9.4 Employing Autopoiesis

The explanation of autopoiesis stated above, does not include a phylogenetic account of coupling, nor does it attempt to explore the communicative or social aspects of autopoiesis (a point upon which Maturana and Varela disagreed). The notions of introspection and self-reference are far reaching (see (Maturana and Varela, 1980a) pp. 30-38) but not directly relevant to this project cf. (Winograd and Flores, 1986).

More important is the idea of embodiment captured by structural coupling. The causal mechanisms underpinning such ontogenetic adaptation are, in Maturana and Varela's thesis, entirely explicated in terms of neuroscience. The connection of phenomenology to autopoiesis, is in the way an agent is treated as engaged in recurrent interactions with its environment. It therefore emphasises the observer's role in defining an autopoietic system, but stresses that the cognitive domain of the observer is *not* that of the agent. Its greatest asset is that it provides a *systematic* model of Heidegger's circumscription, and grounds this project in causal explanation, perhaps not obvious in Heidegger's phenomenology. Discussion now formalises the notion of cognitive domains by explicitly investigating the triple of observer, world and unity.

Autopoiesis is the theoretical framework which suggests the temporal course of events whereby agent and environment are coupled. While not commiting to an internal mechanism for an agent, it strongly coincides with connectionism in its reliance on structural properties such as ontogenetic learning and adaptation of connectionist mechanisms. Autopoiesis is also interesting because it provides a partially worked-out theory of the representational issues of the cognitive domains of observer and agent (unity). We now introduce semiotics as the complementary and fuller theory of representation that will be used in the agent theories and architectures implemented in this thesis.

5.10 Semiotics: Toward a Common Language for Phenomenological Agents

Dreyfus (1991) pp. 5 describes Heidegger's work as a "radicalisation" of Peirce's semiotics. However, it is proposed that just as its philosophical origin in pragmatism suggests, it is *useful* for ascribing intentional content with meaning.

Broadly speaking, the semiotic tradition accepts, *prima facie*, that meaning is relational[7]. That is hermeneutic acts are dependent on the observer. Despite the obvious similarities to phenomenology, semiotics is largely excluded from analytic and phenomenological research, because it is not a general theory of epistemology or ontology in the same encompassing fashion as Heidegger's project. However, it has been illustrated that autopoiesis is semiotic in nature, yet still compatible with notions such as circumscription. This use of Maturana and Varela's work as a bridging principle, requires a language which exists at a different level of explanation. Pragmatically, semiotics presents a language consistent with the philosophy espoused and is able to explain relations between the agent and the world at a variety of levels (e.g. whole agent to constituent unities). Therefore, semiotic explanation may be used without committing to signification as an ontological, analytical unit. (This was Heidegger's objection, that forced the subsumption of signs into the role of equipment). Maturana and Varela explicate this, emphasising that the point of observation for the analysis of what a unity receives as input *is* observer dependent.

Ziemke and Sharkey (2000) note that there is a curious distance between adaptive agents, cognitive science, AI and semiotics. This is surprising given the disciplines shared common interest in relational issues of internal representation of the environment. This is perhaps because of the sentential-symbolic nature of internal representation (and intentionality) implicit

[7]Recall that Heidegger denied this, since signification (the relational part of semiotic theory) is not as fundamental as the analytic unit of Dasein

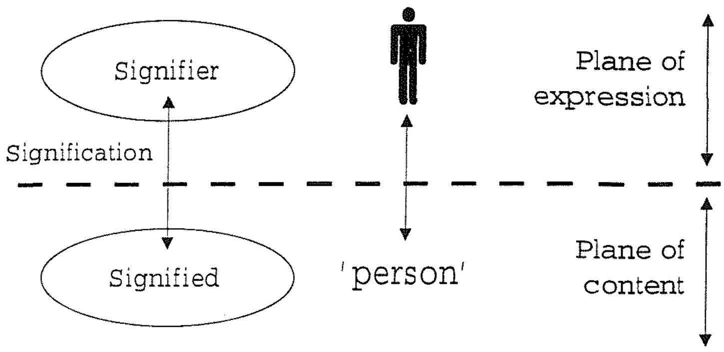

Figure 5.8: European Binary Model

in most AI, because of its model-theoretic origins in the work of Frege and Tarski. More boldly, Sowa (2000) suggests that the field of AI might more reasonably have assumed the title "computational semiotics", because, in essence, that is what AI attempts.

The relevance of computational semiotics to this thesis is that Peirce's triadic semiotics (as opposed to Saussure's binary model) provides a convenient and tractable framework for understanding semanticicity, not constrained to symbolism, sentential explanations or linguistics. Anything that *signifies* something else is given equal precedence, and therefore, incorporates the peculiar notion of internal representation that is presented by connectionism. In Chapter 6 this is also explored in the difficult context of *what* agent epistemology is in a multimedia system.

The two dominant schools of, semiotics (American pragmatist movement of which Charles Sanders Peirce was a key figure) and semiology (European tradition originating in (Saussure, 1986) [8] the *Course in General Linguistics*) will be outlined and then shown to be necessary to the project in hand.

5.10.1 Semiology and Saussure

Although semiotics is the term generally used for the "science of signs in everyday life", the European school of *semiology* originates from the Swiss linguist Ferdinand de Saussure (Guiraud, 1975). Saussure's formulation is the notion of *signifiers* as *expressions* of *signifieds*. For example, the word icon in Figure 5.8 is a signifier which expresses the signified 'person' (a human being). Many different signifiers can be related to the signified, motivating Hjelmslev to describe a *plane of expression* (the domain of signifiers) and a *plane of content* (the domains of signifieds) (Andersen, 1997). The European tradition abstracts away from the user of a sign, in its attempt to deal with the "super-individual" (Andersen, 1997) pp. 11. However, despite

[8] Saussure's work was collected together after his death in 1913, and published in 1916. The English translation used here was published in 1986.

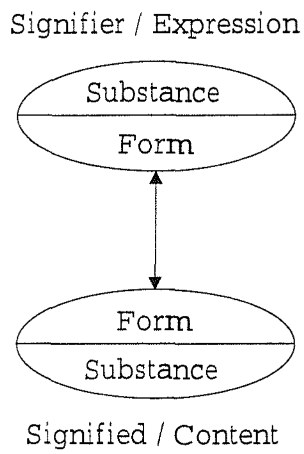

Figure 5.9: Form/Substance in Signs

---

the binary model being less appropriate in the *static* analysis of phenomenological agents and autopoiesis, its utility is in the explanation of the dynamics of semiosis. The planes of content and expression, parasitic as they are on the binary model, offer a convenient explanation of how semiosis occurs by analogy with Newtonian mechanics, energy surfaces and catastrophe theory e.g. see (Andersen, 1994).

The relationship between signified and signifier (content and expression) is more complex than the binary model Figure 5.8 shows. Both signifier and signified have *form* and *substance.* These concepts bring together the raw materials of a sign, and the interpretative aspects in a way that features the social (cf. Background in Heideggerian phenomenology) construction of a sign and its meaning. Figure 5.9 shows both the signifier and signified as composed of two parts. An example from (Andersen, 1997) pp. 80 clarifies how notions of form and substance capture more than a simpler binary model. Two people (A and B) arguing over who will perform a domestic task decide to arbitrate the decision by tossing a coin. The coin's resting position will become a sign which indicates who will undertake the task. The semiotic system, agreed verbally at the time of arbitration, can be described as follows:

- *signifier/expression substance* : the continuum of all possible coin orientations
- *signifier/expression form* : articulation of two interesting positions : heads or tails
- *signified/content form* : the articulation of two possible actors to perform the domestic task, A *exclusively or* B
- *signified/content substance* : the set of all people who can undertake the domestic task

The use of a coin as a sign in an act of arbitration involves and shapes the planes of expression and content into substance and form. However, if we were to examine the role of a

	Quality	Indexicality	Mediation
Material	Qualisign *A quality which is a sign*	Sinsign *An actual existent thing or event which is a sign*	Legisign *A law which is a sign*
Relational	Icon *Refers by virtue of some similarity to object*	Index *Refers by virtue of being affected by object*	Symbol *Refers by virtue of some law or association*
Formal	Rheme *A sign of qualitative possibility*	Dicent Sign *A sign of actual existence*	Argument *A sign of law*

Table 5.1: Peirce's Trichotomies and Theory of Signs – adapted from (Sowa, 2000)

coin as a token of financial value in a transaction, the sign system's substances and forms would be very different. It is for this reason that we can justify the claim that European semiology is about "super-individual" issues, and factors-out the user of a sign.

The difficulty is that while this might be applicable to certain enterprises in social agency, it only partially provides a language for micro-level intentionality in the individual agents. Additionally, it does not map onto autopoietic systems in a convenient way[9].

5.10.2 Semiotics and Peirce

A complete discussion of Peirce's trichotomies of signs can be found in (Sowa, 2000) and (Sowa, 1984). The binary model of de Saussure coincides with Peirce's framework at the level of the *relational trichotomy*. However, the European semiological tradition is embedded in what became the structuralist movement, whereas the American tradition became labelled pragmatism. A unification of semiotics, phenomenology and connectionist agency, requires a methodology which penetrates beneath the relations between signs in the world. Explanations of what relations are supervenient on are required, e.g. an account of lower perceptual processes or sensory information upon which relations such as the signifier and the signified are dependent. The work of Peirce offers just such a framework.

The basic matrix of Peirce's theory of signs is shown in Table 5.1. In terms of sentential-symbolic systems, the third trichotomy is pertinent. Relations between signs and objects are

[9]Although (Andersen, 1994) uses autopoietic theory at a systems level, it ignores the more mechanical levels of composite unities and observer dependent issues of representation that are the concern of this thesis

determined by the conventions of some language, either formal or natural. For example, in writing the word "agent", an author is referring to a concept and this is a *rheme*. If a sentence is formed, such as "an agent perceives its environment", then we are picking out an agent and objects in the environment without specifying what the agent or environment are. This collection of rhemes forms a *dicent sign*. It is an *indexical* reference in that it refers to objects, but does not explicitly name or individuate the very objects by absolute reference. Finally, a collection of dicent signs forms an *argument*, which can be a logical construction such as "an agent perceives its environment, therefore it is cognisant".

The material trichotomy is concerned with signs which are properties of objects, for example a *qualisign* is a pure sensory experience independent of the source such as a visual or aural stimulus. A *sinsign* is similar to a qualisign, except that it indexes, i.e. directs attention or individuates, the source such as a telephone ringing. A *legisign* is a habitual association between such as a telephone ringing and the fact that a telephone rings if some other person is trying to communicate. Clearly, this has similarities to associative, classical conditioning. If a stimulus is presented consistently (qualisign) contemporaneously with another stimulus, then they are associated by the learner. Legisigns can be primitively thought of as expressions of regularity in sensory input and observed behaviour.

The relational trichotomy builds on the material trichotomy. In the relational trichotomy, the familiar signifier/signified model coincides with the semiotics of Peirce. Each relational category determines some kind of relationship between signs and objects explicitly, as for semiological models. An *icon* refers to an object because of some similarity, the most obvious example being the visual shape or form of a sign to its referrent. An *index* refers by being *causally affected* by another sign, for example, smoke is a sinsign which is causally connected to combustion or fire. In this way, the relation between smoke and fire, as two sinsigns, is indexical. Finally, a *symbol* is only associated to its object by convention. For example, road signage are symbols, their meaning is not on the whole communicated by any material property such as qualisigns, sinsigns or legisigns. Neither does it refer by relational icons or indexes. Being arbitrary, the meaning of a road sign is in the convention of its use, and therefore, documents such the highway codes are required to communicate meaning.

In this more complex system of semiotics, we can summarise the key components which directly reference and fit with autopoiesis. The situation is further complicated by the problem that learning, or as Peirce called it, "scientific intelligence", shifts the semiotic relationships around the matrix illustrated in Table 5.1.

Figure 5.10 shows the basic outline of any semiotic system based on Peirce's triadic model, coinciding with autopoiesis. The *sign* is the focal point of semiotic analysis and can be as simple as a Material property, such as a qualisign, and as complex as a law such as an argument. The *object* is the thing in the world to which the sign relates in either the Material, Relational or Formal trichotomies and via its sign qualities, indexicality or mediation. The *interpretant* is similar to the Fregean notion of *sense*, and the Husserlian ideal of *intentional-*

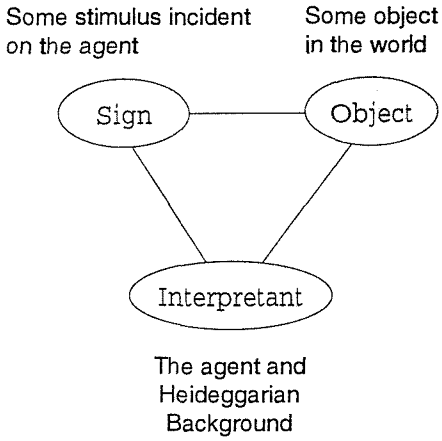

Figure 5.10: Peirce's Triadic Model

ity. The interpretant therefore correlates the sign and object, but is more than *just* a relation in the set-theoretic sense. It is often as sophisticated as the entire agent observing the sign, including predispositions included in its Background. To clarify, it is necessary to identify the point of semiotic analysis, whether agent-centric or autopoietic. If the system of sign creation (semiosis) and understanding is to be autopoietic, the observer must be present (under Maturana and Varela's definition). Alternatively, as the diagram shows, we can take an agent-centric perspective, ignoring the observer, and treating the agent as the interpretant. The difficulty here is that it is the agent's Background which is relevant to the hermeneutic act. This excludes an autopoietic perspective since, as Nagel's allegory shows, we can never know what it is like to be an artificial agent. Hence, the interpretant's form, content and meaning will always escape an observer's interpretation.

If a strictly autopoietic interpretation is warranted, the *interpretant* is the observer, yet this naturally brings with it the vast array of Background, and the cognitive and consensual domain of the observer. The *sign* will be what the observer locates as some stimulus incident on the agent. The *object* will be the concomitant activity (artificial neuron activity) of a constituent of the composite unity (the agent's internal mechanism). Harvey (1992) made a similar point in his paper on the misinterpretation of connectionism. Harvey's dictum is to apply the formula "P is used by Q to represent R to S". It is then necessary to specify who Q and S are, with respect to the "tokens" or arbitrary signs P and R. It is argued that Harvey's approach is overtly semiotic. Later, in the work of (Andersen, 1994), the dynamical systems approach to semiotics was given a computational grounding.

It should be noted that semiotics is not an internalist theory of representation. It is uncommitted, in either variant, to the location of the signifier, signified, interpretant or sense. Just as Frege did not locate intentional content "in the head", neither does semiotic theory. It is plausible to locate part of the interpretant's hermeneutic role within the agent, while acknowledging the "thirdness" of social and law-like conventions which impinge on semiosis and interpretation from the Formal trichotomy.

5.10.3 Heidegger and Semiotics

The utility of semiotics, is in its ability to describe and categorise phenomena. It is important to Heideggarian phenomenology, however Heidegger's dislike of signification as an ontological requirement meant that he largely ignored semiotics until shortly before his death (Sowa, 2000).

The point of contact for Heidegger and semiotics is in the notions of circumscription and the levels of breakdown. A coarse-grained definition of breakdown (Winograd and Flores, 1986) affords two categories :

- *ready-to-hand* which is how the world *is* during circumscription (and describes the *readiness-to-hand* of equipment).
- *present-at-hand* which is how the world is *disclosed* if breakdown occurs (and describes the *presentness-to-hand* of supplementary equipment when circumscription fails e.g. if an alternative hammer is easily available).

Circumscription is routine activity, cf. intentional arcs, that pre-supposes Background. Background is shaped by the complex of social, cultural and personal experiences. The Formal trichotomy (thirdness) is implicit in defining the relationships which are expressed by the Relational trichotomy, and these supervene on the Material trichotomy. Returning to the example of a hammer, it may be said that a hammer has Material signs, which are largely properties such as the visual appearance of the hammer. The Relational trichotomy is concerned with signification, that is the assignment of the hammer to a role. For example, an icon of a hammer suggests or connotes the multiple uses of a hammer as a means of applying force. This relational assignment of signs to the hammer is dependent upon the Formal trichotomy, in that the hammer becoming accepted as a tool is a socio-cultural phenomenon. In essence, the present-at-hand (during breakdown) *requires* the Formal and Relational trichotomies *a priori* and the explicit relational aspects (icons, indexes and symbols) become disclosed through or during breakdown. Otherwise, they are part of the system that shapes circumscription (readiness-to-hand).

5.10.4 Heideggarian AI

In this section, we begin to connect Heidegger with AI more directly, and move towards a treatment of connectionism being the focus of this thesis. (Preston, 1993) asserts that Background (the everyday pre-ontological things which are pre-supposed by circumscriptive activity) cannot be a set of facts and rules. In her terms, *knowing-how* cannot be reduced to *knowing-that*. Background cannot be present in an agent (or AI generally, for that matter) as a set of rules

UNIVERSITY OF SOUTHAMPTON LIBRARY

available for reflection and introspection. The burden is therefore to demonstrate a naturalisation or computational implementation of a phenomenological agency. It is proposed that instead of the rules and facts being *the subject of* an action selection mechanism (cf. a divide between a world model and a separate device which takes action) that the architecture of the mechanism is the substrate of Background in an agent. In effect, circumscriptive activity in agents will be a function of the architecture and will be an embodiment or implementation *of* that background. This may seen in the shaping of the nervous system alluded to by Maturana and Varela's structural coupling.

Implicit in the proposal above is a concession to the causal power of the mechanism which implements behaviour and "encodes" or manifests Background for the agent. Preston similarly allows for a causal explanatory role for internal mechanism, contrary Dreyfus and Heidegger. Preston then poses the question of how much of the explanatory burden internally-held representation can bear, and how much must be based on relations, such as concomitant activity between the agent and the environment. To put this another way, how much of the internal mechanism accounts for meaningful purposeful behaviour and how much is present in the environment? For example, can the environment be the *complete* bearer of intentionality, leaving the agent as a simple wiring of stimuli to responses? Or, is there more of a balance between internal mechanism and the environment? These questions will be addressed in technical work presented in subsequent chapters.

Preston helps clarify her stance by looking at intentionality as conceived by eliminativists such as the Churchlands (Churchland, 1986; Churchland, 1979; Churchland, 1996). The Churchlands maintain that intentionality is an internal phenomenon, but reject sentential languages of thought. By implication, symbolic intentionality (even if *not* sentential) is also rejected in favour of the state-space interpretations given to connectionism cf. (Smolensky, 1988; Rumelhart and McClelland, 1986) under super-positional representation of micro-feature inputs/stimuli. However, these theories of representation are predicated on composition *as concatenation* (van Gelder, 1990). This concatenative compositionality is the classic propositional logic approach to building systems of propositions from atomic tokens and formulae.

Both Dreyfus and the Churchlands (according to Preston) are aligned with internalist accounts of intentionality. However, Preston's own position is similar to that espoused here, namely that an *interactionist* locates intentionality not *only* on the basis of internal representation, but also on ascription of meaning cf. (Dennett, 1987). Intentionality is parasitic on the environment, giving credence to the meaning of internal content just as a sign has no divorced capacity to "mean" anything in the absence of Peirce's referrent and interpretant.

Preston emphasises that connectionism offers some insight into how phenomenology might be more appropriately characterised in computational terms. She also insists that if the environment is seen as a *provider of training sets*, then the ascription of meaning to intentional states collapses. The environment must always be present, re-emphasising the role of circumscription and comportment with respect to the agent and environment. Consequently, the computational paradigm of agency forces us to concern ourselves with ongoing, situated interaction – decaying the traditional boundary between training and performance in connectionism. In isolation, connectionism without its embedding in an agent is insufficient, but a combination of the two offers a synergy which simultaneously defeats the sentential symbolic assumption of classical agents whilst giving connectionism the "grounding" it requires.

5.11 Unifying a Phenomenology of Agents

Heidegger indicates breakdown as invoking, or disclosing, levels of intentionality such as those traditionally assumed in AI and deliberative, sentential-symbolic agents. As stated earlier, it is only in degrees of *total breakdown* that Heidegger recognises the kinds of intentionality which are normally sentential-symbolic envisaging.

During circumscription, assignments of properties to *ready-to-hand* objects are transparent to the agent. Deliberation, or envisaging, is *not* the substrate or foundation of circumscription. It is proposed here, that a means of progression with a phenomenological theory of agents is to use the causal mechanism *itself* to underwrite (to use Preston's phrase) absorbed activity, and seek a theory of intentionality with respect to this mechanism and its functional role in the agent/environment interactions. A deliberative mechanism, employing a sentential-symbolic world model, cannot meet these needs because mechanism and representation are separate. Both Dreyfus and Preston elaborate this argument with reference to computational complexity growth when Background is "captured" in sentential-symbolic models.

For agency, however, we can infer that if situatedness is a feature of circumscription, and disturbance reveals levels or modes of intentional engagement, then: "Fact and rules are, by themselves, meaningless. To capture what Heidegger calls significance or involvement, they must be *assigned relevance*. But the predicates that must be added to define relevance are just more meaningless facts; and paradoxically, the more facts the computer is given the harder it is for it to compute what is relevant to the current situation" (Dreyfus, 1991) pp. 118. This borne out by Shoham having to prevent complexity explosion by limiting the embedding and expressiveness of his BDI modal logic.

The conclusion most influential to technical practice, drawn from Heidegger's phenomenology, is that deliberation is parasitic on something lower-level, being Background and Dasein. To progress, autopoiesis, Merleau-Ponty's intentional arcs and semiotics have therefore been used to advance a theory of representation which is not inextricably tied to sentential-symbolic ideals.

Autopoiesis is a relational semiotic theory such as Peirce's second trichotomy. It does not acknowledge the "firstness" or "thirdness" of the *material* and *formal* trichotomies respectively. Maturana and Varela insist on the cognitive and consensual domain of the observer (analogous to Background) being the interpretant of any hermeneutic act over internal mechanism, *at any level of granularity*, within the complex unity (agent). An artificial neuron in a connectionist system "tells" or "carries content" for the agent depending on the sign (the stimulus), the object (an artificial neuron as part of a complex unity) and the interpretant (the cognitive and consensual domain of, and including, the observer). A similar study to that conducted here is presented by (Goldspink, 2000). He proposes a meta-model of agency which can be used in the study of social systems employing multi-agent simulations, using a strictly autopoietic perspective. His principle motivation is that social simulations, based on sentential-symbolic assumptions of representation, do not give enough credence to the role of environment in the emerging patterns of behaviour in a complex society of agents.

The power of autopoiesis, is that it frames the semiotic relationship in biological and cognitive language – strictly enforcing a relational view which is congruent with the irreducibility of phenomenological experience expressed by Nagel. However, as Ziemke and Sharkey demonstrated (using von Uexküll's theory of biological semiotics) the agent's perceptual and operational worlds (corresponding to domains of perception and action respectively) can be taken as signifiers where the agent is the interpretant. This creates a kind of agent-centric semiotic theory, in which the agent interacts with the world, whether or not an observer is present.

Ziemke and Sharkey provide an account of what Peirce calls *firstness*, which is captured in the *material trichotomy* of qualisign, sinsign and legisign. To further the assignment of meaning in this context (as with the relational approach), a qualisign such as a stimulus might become a sinsign, a stimulus which acquires meaning for the agent's activity if it is repeatedly followed by harmful consequences. A legisign becomes present when the agent responds to the stimulus in order to avoid the harmful consequences. Learning from operant conditioning is therefore analogous to the semiotics of firstness. Semiotics provides a framework for discussing agent intentionality, just as autopoiesis does, but it advances beyond the relational model of Maturana and Varela to include Heidegger's Background and Dasein. In technical practice, this means that we can integrate both an agent-centric view (cf. Ziemke and Sharkey) and an observer-relational sign system. Although there is no concrete mapping between Heidegger, autopoiesis and semiotics, Sowa (2000) has suggested that firstness and thirdness are analogous to ready-to-hand and present-at-hand in circumscription.

The final bridging principle, is that autopoiesis provides a cognitive approach within the confines of neuroscience. Maturana and Varela insist on systems of neurons being the fundamentally reductive basis of phenomenology. This enables the connection with Merleau-Ponty who, like Preston and Dreyfus, insists that acquisition of skill is underwritten by intentional arcs. The "arc" is something that "maps" a particular perception to a particular action during circumscription. The mechanism is the intentionality, if at all. In principle, what Merleau-Ponty alluded to is what is now called reactive agency. However, Brooks, Agre and Chapman did not explore the adaptive requirements of a self-organising system that facilitates intentional arc acquisition, without explicitly symbolic representation.

To connect intentional arcs with computational agents, it has been suggested that connectionism be considered as it allows for adaptation, and is the mechanism which underwrites activity for an agent. The weights (structural loci of adaptation) and activation functions of artificial neurons (dynamic causal mechanisms generating action) are neither symbolic, nor sentential theories of intentionality. Connectionism is also (in parallel distributed processing terms) a paradigm example of reactivity and stricly bounded rationality.

5.12 Design Requirements for a Phenomenological Agent Architecture

In this thesis, it is proposed that phenomenology is the superstructure which provides an agent with a meaningful relationship with the environment. It articulates (especially under Heidegger's version) a set of necessary relationships in terms of intentionality, circumscriptive activity and here, these have been set against a mechanical backdrop which posits adaptation as the device which *returns* an agent to circumscription. In effect, phenomenology is a principled exposition of an agent theory, which also constrains the internal structure of the agent. In addition, (and because of its lineage and development around the time of the rise of cognitivism) Merleau-Ponty's work provides a psychological-oriented perspective on embodied skill as circumscription, but further, enables the principle of intentional arc to manifest such routines. Heidegger's *Being and Time* pre-dates such speculation on the manifestations of circumscription.

To summarise:

- circumscription is the fundamental mode of engaging with the environment. The computational notion of agency enables this view, although phenomenology challenges the means of implementation
- circumscription is not symbolic, or indeed just internal. It is interactionist, and so meaning and internal mechanism is parasitic on the environments relationship to the agent.
- any higher-order activity must be couched and grounded in circumscription, cf. (Wilson, 1991; Harnad, 1992; Cangelosi, Greco and Harnad, 2000).
- in approaching phenomenology from a computational perspective, it is necessary to adopt an agent approach as well as a principled mechanism of intentional arcs and circumscription.

Perhaps the most profound implications of this work are that goals, desires, intentionality (representational or otherwise) are encoded or embodied in the very structure of the agent (Preston, 1993). This structure is determined by recurrent interactions with the environment (Maturana and Varela, 1980a).

When developing a strong agent theory, the following foundations must be emphasised:

- no higher-order epistemology (e.g. sentential symbolic structures) can be present without *structure determined, circumscriptive absorbed activity.* In Heidegger's words, an agent cannot have the ability to *thematise* without non-thematic circumscriptive activity. cf. (Harnad, 1990) and (Wilson, 1991).
- agent implementation (adhering to such a theory) needs unpolluted (or as close as possible) intentionality sufficient for circumscriptive activity
- something more fundamental than symbolic intentionality is required (hence, connectionism as a plausible substrate).
- the architecture that embodies routine activity *could* employ a vertical structure of breakdown and disturbance to facilitate any kind of thematisation.

One significant implication is that a theory of representation is required which is embodied in internal architecture, and shared with the environment and observer. (Agre, 1997; Agre and Chapman, 1987) accounted for this by introducing indexical-functional representations. Simultaneously, this eliminates static representation (although their's was still symbolic, but curiously not sentential because it disallowed variable binding) and acknowledges the *referential* (hence indexical, cf. the relational trichotomy) and the causal aspects (functional) of the internal symbols. While powerful, and clearly semiotic, it is symbolic, and therefore the complexities of connectionist intentionality required more careful examination.

In terms of producing a concrete architecture that supports this theory, it has been suggested that skill acquisition and the formation of intentional arcs is key to realising circumscription. Naturally, connectionism and its learning algorithms are a candidate.

5.12.1 Requirements

To achieve any incremental step towards the theory of agency espoused here, the following key points must be carried forth into the design of an agent:

- intentional arcs
- adaptation as a key to establishing routine activity (i.e. circumscription)
- a model of breakdown
- an account of the agent-centric and observer-centric versions of intentionality cf. Peirce's trichotomies *versus* autopoietic interpretations
- a model of adaptation which is plastic enough to account for the maintenance of routine activities, in environments commanding different and changing modes of comportment

These will be incorporated in the following chapters as design principles.

5.13 Conclusion

The contribution of this chapter has been in reconciling computational agency with phenomenology. The approach was to consider Heidegger's phenomenology, and employ Merleau-Ponty and autopoiesis as means to operationalize Heidegger's theory of agency. To facilitate these diverse approaches, a commonality exists in the pragmatist theory of semiotics. It was shown that these approaches are congruent, especially if the embodiment of an agent is brought to bear on the philosophical problems and disparities. Dreyfus reminds us that Heideggarian phenomenology is a radicalisation of the pragmatists approach, that is the American and European traditions of semiotics and semiology respectively. In effect, bringing a number of different philosophical theories of intentionality to bear on the principle of *enaction* which began in autopoietic theory (Varela, Thompson and Rosch, 1991).

A recent position and overview given in (Dautenhahn, 1997) emphasises the social basis of phenomenology. By analogies with social behaviour in primates, she describes how a unification of computationalism and phenomenology might work. Her approach suggests embedding and enaction (cf. autopoietic structural coupling) and a variety of other AI techniques to provide the "high effect" features of social interaction such as the use of language to influence other agent's decisions. The approach here is similar, but is a 'micro-complement' to Dautenhahn's macroscopic strategy.

The contribution of phenomenology to this thesis, is that it presents a foundation for a strong agent theory which acknowledges the situatedness of agency, the primacy of experience and interaction with the environment and (by the operational principles used here) enables a framework for routine activity in agents to be established.

What is perhaps unusual about the study of agency and connectionism is that either cannot be said to be sufficient, as (Dautenhahn, 1997) states: "Even neural network approaches are mostly decoupled from the (social) dynamics of the environment". Recently, computer science has begun to treat agents as the grounding principle enabling systems to be thought of as the 'nearly decomposable systems' of (Simon, 1985). However, without a suitable substrate for intentionality, it has been argued that this is incomplete. Connectionism can be viewed as an abstract mechanism for implementing internal representational intentionality, but is susceptable to the symbol grounding arguments (Harnad, 1990). This suggests that connectionism is nothing more than a *different* implementation of cognitive architecture that really seeks to provide a systematic language of thought (Fodor and Pylyshyn, 1988). Merging the two contributes to both. By developing an agent with an adaptive mechanism, which captures the notions of circumscription and intentional arcs, connectionism is legitimated by placing it in an enactive or interactionist context. From an agent theory perspective, an agent gains a plausible substrate of circumscription which is absent from sentential-symbolic theories.

Part III

Applications and Simulations

Chapter 6

Agents for Multimedia Retrieval and Navigation

This chapter represents the first attempt at making concrete the ideas expressed in the preceding chapters. It begins by exploring the emerging trend to view multimedia information in terms of low-level media objects and high-level components; the former being feature-based and the latter expressing content information – e.g. semantics, intended meaning/interpretation or symbolic meta-data – about what is portrayed by the media object. Traditionally, this has been viewed by employing analogies with generative linguistics (e.g. compositional semantics). Recently, a new perspective based on the semiotic tradition has been alluded to in (Smoliar, Baker, Nakayama and Wilcox, 1996) and used by Tansley (2000) as the basis for the multimedia thesaurus. It is proposed that this is a more appropriate approach. Semiotics presents a critical theory which forces reflection on the nature of information as it both stored and processed in multimedia information systems (MMIS). Partly, the need to describe these agents from a semiotic perspective arose from a frequently asked question which took the form of how agents could tell apart (say) a horse from a chair, since both have four legs. This question arises (in part) from the belief that an agent is in some way more powerful than any other machine recognition technique (possibly because of the anthropomorphic language used in discourse on computational agency). However, to explain the source of the error in this question as well as the hermeneutic 'circuit' which clearly identifies the source of interpretation, the semiotic perspective proved useful.

Later in the chapter, the first attempts at generalising connectionist models to an agent architecture are outlined. Three classification networks were implemented and tested – a standard MLP with back-propagation of errors based on the classic sum-of-squares error, a statistically robust version of the same network (using an error measure based on cross-entropy error) and a 'heuristic' agent-based network, derived from the principles of Chapters 3 and 4. The agent theory from Chapter 5 is used as a means of exploring epistemology of the MAVIS2 agents.

This chapter contributes a prototype implementation of an agent system exploring architectural issues developed in this thesis, a compatible framework for understanding multimedia information in relation to computational agents and knowledge representation and finally, serves to highlight future work which shapes the remainder of the thesis.

6.1 Introduction

Multimedia content-based retrieval and navigation (CBR/N) relies upon relatively low level feature extraction techniques to provide representations of the media objects. For example, an image might be represented in a retrieval database by features representing texture, colour distributions etc. More recently, the annotation of media objects with symbolic metadata has provided an alternative, "semantic feature" based description.

(Lewis, Davis, Dobie and Hall, 1997) describes the research on the MAVIS project, which represents initial work undertaken in the Intelligence, Agents and Multimedia research group on integrating hypermedia and multimedia content. A direct extension of this work was also indicated which became the MAVIS2 project. This attempted to further increase content based navigation and retrieval capability by integrating the low and high level information pragmatically, using an intelligent agent approach.

In the MAVIS2 system (Dobie, Tansley, Joyce, Weal, Lewis and Hall, 1999) this high-level information is captured by means of a restricted conceptual (associative or semantic) network. While the architecture and design intends to capture the full gamut of associative knowledge structures, a restricted version (termed the multimedia thesaurus) was implemented; see (Tansley, 2000) for details. It aims to capture certain key taxonomic relationships between concepts as thesaurus-like associations and simultaneously, associates exemplar multimedia data items usually as processed images. One contribution of the MAVIS2 system is that navigation can be performed on multimedia content (e.g. content based navigation and retrieval) or by using content methods augmented with data contained in the taxonomy.

The first philosophical problem is that the word "CAR" denotes a concept "CAR" in the associative structure. However, an image of a car is equally as representative, but perhaps, more difficult to process algorithmically. In this instance, the choice of the text as representing the 'concept' is a purely technological one. In MAVIS2, the principle architects (Mark Dobie, Robert Tansley and Paul Lewis) chose to define a "preferred representation" which might be the text "CAR". The associated images, for example, are then subservient to the preferred representation but can be shown or queried upon by substituting the textual representation for an "equivalent" or representative multimedia representation. This is the second interpretation of the multimedia thesaurus; it provides "synonyms" for any one of the multiple representations of concepts.

Returning to the epistemological questions around what agents in MAVIS2 learn or represent, the answer lies in what is represented and captured by system of symbols and signs. This thesis attempts to provide agents to aid the integration of the high level and low level features. If we adopt the agent paradigm, an explanation of the multimedia problem compatible with that paradigm must also be offered. Otherwise, the paradigm of agency has no inherent explanatory value in the multimedia context.

With this in mind, this thesis departs from the knowledge engineering perspective, and adopts a semiotic approach. The fundamental problem illustrated in the preceding paragraphs was that of commitment to a concrete definition of "concept" and its unique representational relationship with its denoted "real world" object. It was this traditional knowledge engineering approach that led to the exploration of alternative theories, and finally to a semiotics approach.

Other workers have similar noted this. Of direct relevance is the work of (Tansley, 2000) which adopted the European model of semiotics (semiology) espoused by (Smoliar, Baker, Nakayama and Wilcox, 1996). Saussurian semiology (Saussure, 1986) abstracts away from the user-agent-multimedia object relationship. In part, this is because structuralism (the intellectual doctrine descendant from Saussure and his contemporaries work) attempts to explain the relations in terms of institutional use of signs (cf. Saussure's focus on *synchronic* relations in linguistics – see (Saussure, 1986) foreword and Chapter 1 – the complete relational nexus of signs, their production, use and reproduction). As in Chapter 5, semiotics can take an autopoietic stance, factoring in the observer's cognitive and consensual domain when discussing the epistemology of unities (agents). Peirce's relational trichotomy is therefore highly applicable in the context of multimedia agents, where the multimedia object plays the part of a sign, the user the interpretant and the 'real world' object the referrent. For this reason, C.S. Peirce's model is adopted[1] which is similar to the triadic model of (Ogden and Richards, 1949).

6.2 Introduction to MAVIS2

Figure 6.1 depicts the data structure at the root of the MAVIS2 system. The design intentions for MAVIS2 were to allow the user to navigate and retrieve from multimedia information spaces using traditional, content-based methods or, using the taxonomic relations described by the conceptual layer or a combination of these. Traditionally, text-based systems enable navigation by keywords, and the associative network of links between keywords is taken as a navigational structure. Likewise, conventional content-based retrieval is feature similarity based only. The novelty in MAVIS2 is the integration of both levels, the higher knowledge-based level with lower-level feature-based methods. This chapter describes intelligent agents that facilitate the process of navigation and retrieval between and within these two distinct levels. The full detail of the networks implemented can be found in (Joyce, 2001a), which for

[1]A useful first reference for Peirce's ideas is J. Sowa's book (Sowa, 1984)

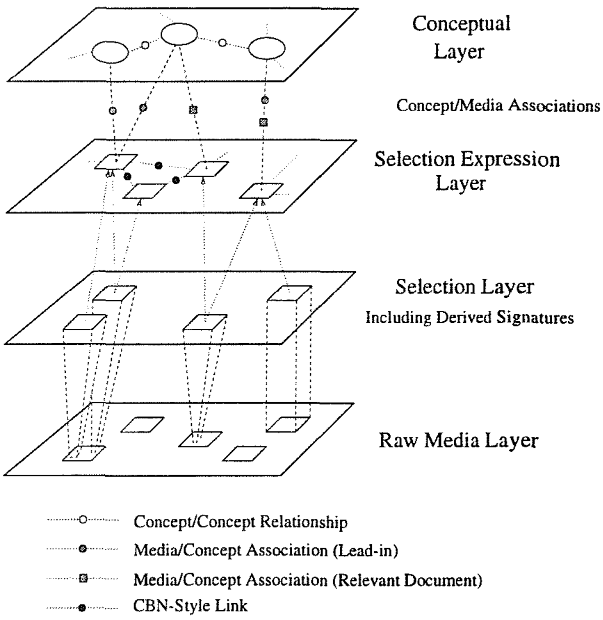

Figure 6.1: MAVIS2's Four Layer Multimedia Data Architecture

convenience, is included in Volume II of this thesis. Essentially, the agents provide 'fast-track' connections between classes of low-level data and the symbolic representations of classes to which they are deemed to belong. This has one significant engineering efficiency in terms of speed of search. The agent's learn (by training) associations between the levels and respond quicker than if, for each query, the whole of the mediating layers (the selection and selection expression layer) were involved in an exhaustive search. Also, the pattern classification abilities of the agents enable queries which were unseen (but transpire to be sufficiently similar to learned categories) to be plausibly processed and for the agents to 'suggest' likely classifications based on feature similarities. Another potential advantage is in the use of multiple agents trained on different feature sets, which enables a fusing algorithm to be used to enhance classification.

6.3 Semiotics and Multimedia

A simple interpretation involving the user of the sign is given as follows; see Figure 6.2:

- the *sign* : refers to something other than itself (e.g. the word "car" or an image of a car)
- the *object* : the referrent of the sign (e.g. an actual car or class of car-like things in the world)
- the *interpretant* : the effect or causation of the concept in some user's mind (e.g. the effect of observing the sign causing a person to mentally conceive of a "car" based on their experiences of cars)

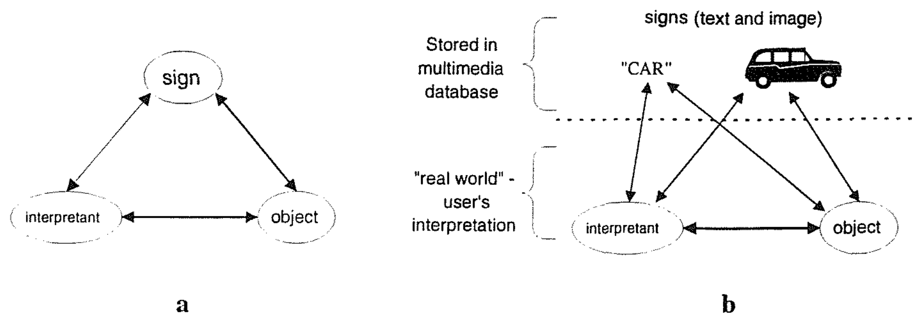

Figure 6.2: (a) the Peircean triadic model of signs (b) an example of "cars"

For a moment, we turn our attention to how a sign might manifest itself in a database or any other organised data structure. Figure 6.2 (a) shows the Peircean model of relational signs and 6.2 (b) shows the concrete example of a relationships for the car example. For ease of exposition, we might say the "real-world" depicted is the user, but it should be noted that Pierce intended no such mentalistic or individual *possession of concept* interpretation cf. Frege's concept of 'sense'. Instead, semiotics emphasises the holism of social and cognitive factors in the whole process of sign creation (semiosis) and interpretation. Analogously, Heidegger describes this as Background – the 'pre-ontological' conditioning which shapes the agent's perception.

Now, the scenario depicted in Figure 6.2 b shows how different signs are interpreted, aggregated and classified together despite the very different manifestations in the database (image and text respectively). The fact that human cognition and social experience enables such classifications leads us to the suggestion that both high-level representations (as abstract as the word "car") and what are seen as lower-level perceptual features (those present or conveyed by the image) are used conjointly in the hermeneutic act. This is significant, because it emphasises that the multimedia object (e.g. picture of a car) has no intrinsic or original semantics i.e. what (Searle, 1980) calls original and derived semantics cf. (Dennett, 1987). All artefacts of semiosis possess derived semantics.

6.4 Multimedia Integration as Semiotics

Figure 6.1 shows the type of data structure used to capture taxonomies in MAVIS2. If everything is considered a sign, then the conceptual layer contains signs which conveniently represent the objects (the referrent). In MAVIS2, the conceptual layer consists of a preferred representation of each object captured by the ontology. We use textual (symbolic) signs in the conceptual layer, primarily for convenience in authoring and displaying the conceptual layer.

At lower levels, we move to the selection layer. This contains signature (feature-extracted representations) of media-level objects such as images, video and audio. The vertical link between the conceptual and selection layers is mediated by a selection expression layer, which represents the connection between what might loosely be called the "concept" and its representation in a variety of media types (image, audio etc.).

The signs present in the conceptual layer are textual, symbolic descriptions – classically called high-level knowledge representations such as the word "car". In Peirce's semiotic framework, these are called *symbols*. We emphasise this term because it has technical meaning: a symbol *signifies* its object *relationally* – that is, by some agreed law or association (see Chapter 5). Other possibilities in this classification of signs include signification by *icon*, where the relation is by virtue of the sign's apparent perceptual similarity to the referrent. However, as far as textual labels for referrents in a multimedia system, they most closely resemble symbols. Symbols were (in Peirce's terminology) signs by virtue of habituation – that is, the association is learned through continued and persistent use; there is no intrinsic reason why the sign acts as a signifier otherwise. This derivative nature of written words led Saussure to denigrate writing as secondary to the purity of spoken language. They differ from *icon* and *index* signs because their visual appearance (material and indexical properties) do not affect the interpretation of the sign. It is, therefore, an arbitrary association between sign and referrent.

Associated with each symbol are the multimedia representations of the same object. For example, the word car is commonly associated with the image of a car. Such images are iconic symbols. However, we need to consider another level of signification, because the machine will assess retrieval based not on the icon, but *feature* signs extracted from the icon. This makes a semiotic model more complex, but not intangible.

Let I_c be an image or multimedia object which is a member of an aggregated collection of multimedia objects to form a conjoint representation of some real world object, captured in the conceptual layer [2].

Let S be some function which yields a *signature* or feature-based representation of the image. Then $S_c = S(I_c)$ denotes the formal relationship between the original sign (the image of the car) and a new, feature sign which will be used in processing this image (e.g. to assess its similarity in a search). Any traditional matching algorithm will be utilising the feature signs, not the original sign. Two principles emerge:

- the *value* of the feature sign in referring unequivocally to the original sign (the image of the car) is given by its ability to represent features which enable the matching algorithm to qualitatively agree with human judgements of perceptual similarity. For example, a "poor" feature set or extraction algorithm is one which represents many perceptually different signs as the same or a very similar feature sign.

- the matching algorithms may not have access to the original sign, only the feature sign.

[2]We use images as an example, but the semiotic framework presented here applies across all media. It should be noted that the nature of the original sign might be different – a sound might reasonably be a qualisign; a purely material property pertaining to the car – but the processing performed by the machine will not be conducted on this qualisign either, but a feature sign derived from it.

Therefore, it is the responsibility of the feature sign to represent information (for the matching algorithm) relevant to the process of signification of the original sign.

If we hypothesise a function $\Gamma(a,b)$ which measures the similarity of two signs, and then we are presented with a query sign (e.g. another image of a different car) Q_c, then we might say that: a human assesses similarity of two images by "computing" (Marr, 1982) $\Gamma(I_c, Q_c)$ whereas the algorithm implemented in a multimedia retrieval system computes $\Gamma(S(I_c), S(Q_c))$. Hence, the machine is using a feature sign not the original sign.

Figure 6.3 shows how this is modelled in a semiotic framework. The signature generator S produces a vector such as $\langle s_1, s_2, \ldots, s_p \rangle$ where each of the p elements is a measure of a particular feature of the original sign. For this reason, we may treat them as "raw" indicators of specific content in the original sign (the image) such as "frequency of horizontal edges". This means that as signs, the features each represent elements of a *qualisign* – a material quality of the object it refers to. The reader is reminded that the object it refers to is in fact the sign or image of the car, not the car itself. Then, the interpretant has no intuitive meaning, but for symmetry, becomes the sum of the features and algorithms which relate the object (the image of the car) to the sign (the signature).

6.5 Related Work

In terms of retrieving multimedia information, Santini and Jain (Santini and Jain, 1999) recognised that the semantics of an image are relative to the retrieval task and the user's goals, however, they did not connect this observation with semiotic notions. The novelty in Santini and Jain's work was in the way the user constructed a display space which visuo-spatially reflected the user's assessment of image similarity and relevance to the search task. In the semiotic model advocated here, the user is explicitly reflecting to the search engine their assessment of the role of the object and interpretant (signified) in relation to other signs. This is a radical advance acknowledging the role of hermeneutic acts in the user's retrieval task, since most other systems ostensibly retrieve based on assumed objective geometric measures of sign similarity and a simple correspondence to object and interpretant. Other systems have attempted the kind of integration of multimedia data presented in MAVIS2 (Hirata, Mukherjea, Okamura, Li and Hara, 1997). However, only Smoliar *et al* (Smoliar, Baker, Nakayama and Wilcox, 1996) have connected multimedia generally with research in computational semiotics.

6.6 Bridging the Gap

In MAVIS2, we have collections of media objects which are taken to be in some way representative of a set of concepts. Bound to these collections is a taxonomy describing relationships between the concepts. The conceptual layer represents domain specific information about concepts – the signs are textual descriptors. The lower selection layer consists of qualisigns of the corresponding media objects such as images of cars. In Figure 6.3, this division is shown as a vertical divide in the top layer of the semiotic hierarchy. We wish to bridge this gap effectively, allowing for:

- media to symbolic retrieval – the system is presented with a sign, e.g. the query image, which is converted to a set of qualisigns. Retrieval tasks take place over the set of all available qualisigns (e.g. signatures of images) and the the system returns (using the associations stored in the conceptual layer) a set of symbols indicating the conceptual association of the query image. The user then has a route into the ontology described by the conceptual layer purely by symbols (text). A user may then navigate in the conceptual layer.

- cross-media retrieval – once a query has been classified, alternative media representations (signs at the top of the bottom layer of Figure 6.3) can be presented to the user.

In relation to the MAVIS2 data-structure and the role of the agents, Figure 6.3 shows the complete semiotic system of signs in MAVIS2. The reader is reminded that the signatures are effectively signs, where the object (referrent) is the original image. The exception is text; where we have shown the text signature, in fact, it is the *preferred representation* for the MAVIS2 conceptual layer. The signature is, for example, the stored numerical representation of the word (sign) "CAR".

6.7 A Multi-Agent System

Given this conceptual description of MAVIS2, we now describe the implementation details. As shown in Figure 6.4, each agent possesses knowledge of (and specialises on) :

- specific concept classes (signs that are symbols, signifying the object such as "car")

- specific signature types (e.g. texture-based features, spatial colour histograms)

Hence, during incremental authoring of a concept-based multimedia application, new concepts and any signatures associated are presented to an applicable, capable agent or a new agent is spawned. Dividing the agent's responsibilities by concept is shown in Figure 6.4.

During the query process, agents receive messages requesting classifications. These are signatures and their associated type (e.g. sound or image and the feature algorithm used). The agents return confidences and classifications which are then fused (using a simple weighted sum) and the MAVIS2 system then presents the agent's suggestions to the user. The user can then (if desired) provide feedback, which reflects their satisfaction with each declaration provided by an agent. This information is used to assign credit or blame to individual members

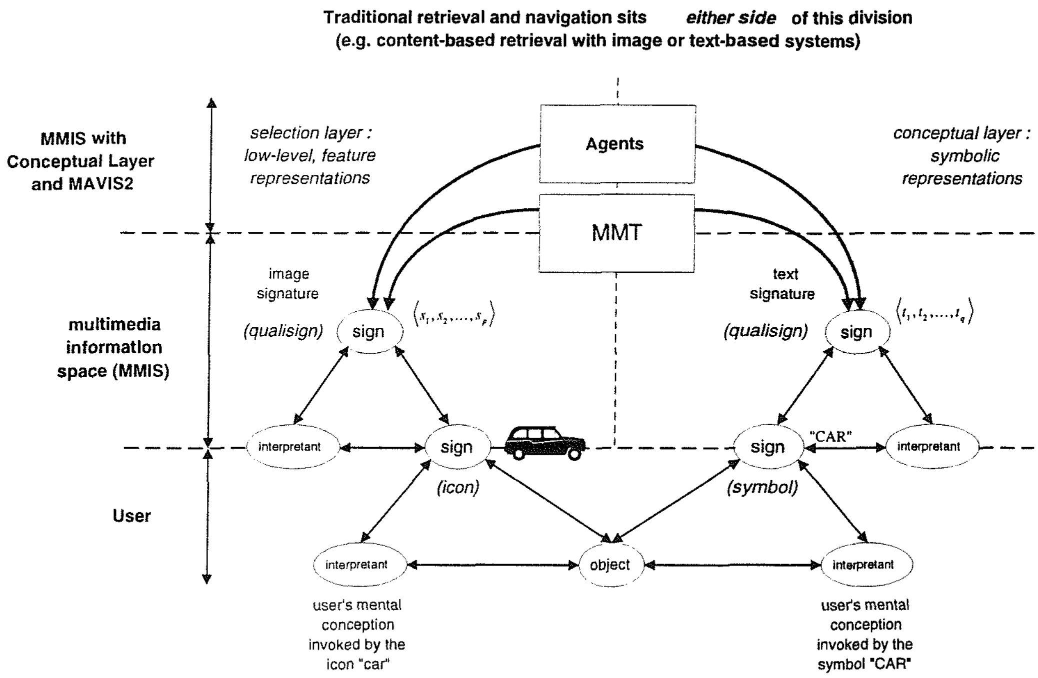

Figure 6.3: A Semiotic Framework for Multimedia Retrieval and Navigation

of the agent population. In future, during results fusion, the trust associated with each agent is used to weight its declarations. Selectionist mechanisms such as these have been used in other multi-agent systems, and agent-based multimedia environments (Minka and Picard, 1997).

Here, agents should attempt to answer with suggested symbols (classifications and conceptual associations) when prompted by a query from the remainder of the MAVIS2 system. This may be the result of a user submitting a query.

6.8 Classifiers and Connectionist Techniques

As described above, each agent is a single unit coping with a particular region of the concept-space. The agents "know" about members of the concept-space and the preferred sign (text descriptions) and has learned associations with qualisigns at the computational level of signatures. This represents the application of inductive learning to concept acquisition (Michalski, Carbonell and Mitchell, 1983) in an ill-defined problem. Connectionist classifier techniques were considered and implemented.

6.8.1 Requirements and Constraints

The theory of connectionist techniques for robust statistical classification is well documented (Bishop, 1995; Rumelhart and McClelland, 1986; Jordan and Bishop, 1996; Haykin, 1999;

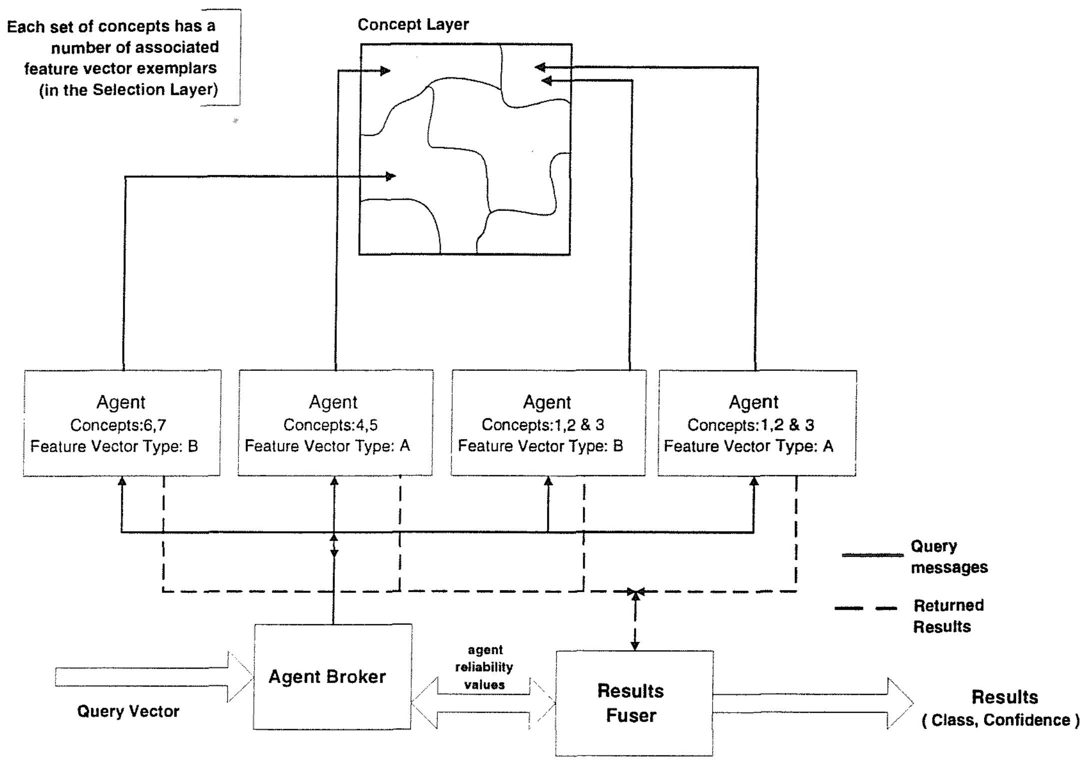

Figure 6.4: Multi-agent System for Media-to-Concept Classification

Lippmann, 1987; Bridle, 1990; Ruck, Rogers, Kabrisky, Oxley and Suter, 1990). Neural networks have an appeal because despite being semi-model based, they can be used (perhaps, misused) as black-box devices for classification tasks. In the absence of any real information on the feature space, we can train a neural network and observe its performance. If the trained network classifies well on both the original ensemble feature vectors and unseen vectors, we are satisfied. However, the resulting classifier may be coercing the implicit model to fit the data, usually resulting in phenomena such as over-fitting of the model to the data (sometimes called "memorisation" – where the network will *only* classify vectors in the training process) and lack of predictive ability (that is, the network cannot generalise the classification schema obtained through training to unseen vectors).

The following requirements were determined for agents in MAVIS2:

1. be *constructive* : the concept-layer and selection-layer may change as a result of incremental authoring. Hence, a classifier must be able to learn new exemplar / concept associations requiring the network to grow.

2. be capable of *on-line training*: the network must be able to switch between training and performance phases.

3. be robust to noise, but sensitive to novel data: the ability to classify outlier feature vectors

as either new classes or extensions to existing classes.

4. enable a control *superstructure* to be integrated : the network must be modular enough to be controlled during the training process so that correct components are trained correctly.

5. be as far as possible, model-free : the implicit statistical assumptions should be minimal, so the classifier network is able to model heterogeneous distributions and classification schemata.

6.8.2 Techniques Implemented

The following network architectures (based on those surveyed in of Chapter 3) were implemented and tested as candidates:

1. An MLP using sum-of-squares (SSE) error measure.

2. An MLP using the statistically robust cross-entropy error (CEE)

3. A localist network, which divides the problem of classification into two parts; modelling the distribution of the feature space and mapping categories onto output classes (symbols)

Neural network-based classification systems usually offer sub-optimal but time efficient performance (for example, when compared with optimal solutions such as first nearest neighbour) and partial freedom from explicit statistical models. This makes them candidates for a multimedia content based retrieval or navigation system, where a compromise between optimality and speed favours time performance for user satisfaction.

The decision to test two MLP training rules was because the agent system generates training data which consists of the exemplar feature vectors and classifications coded as one-of-c vectors. Therefore, target vectors for classes are given as binary vectors such as $\langle 0, 0, \ldots, 1, \ldots 0, 0 \rangle$. Such target vectors form a space where each class is the vertex of a hypercube. This distribution is far from the normally distributed assumptions which the sum-of-squares error measure assumes when it attempts to shape the classification function implemented by the MLP. So, the MLP using SSE may form a skewed model of the classification task, whereas the CEE error should provide a more robust model (Bishop, 1995).

The localist network is based on a system which effectively models the distribution of the M dimensional feature space in $N << M$ dimensions by categories, and a module which learns to predict class/concept associations by mapping from the N dimensional space to classifications. The feature space map is formed using GART – Gaussian ART – of (Williamson, 1996). This provides an constructive algorithm. The mapping from category nodes to output classes is achieved using a linear associative network and search. The motivation was the conditioning network of (Chang and Gaudiano, 1998) based on previous conditioning models of Grossberg *et al*; see also (Joyce, 2001a), included in Volume II of this thesis.

The training of any internal classification mechanism for an agent would require that it cope with statistically sparse data – multimedia applications may not generate enough batch data to train a reliable classifier. If this is the case, an agent should be able to perform *some* classification and be prepared to refine its knowledge when more data becomes available. The agent should be able to suffice with available information but be able to train further. For example, a general rule (Foley, 1972) is that to obtain robust classification, there should be three times as many examples (signatures) per class as there are elements in the feature vector. Any autonomous agent responsible for the tasks defined above will need an approach grounded in principles of learning which the classic MLP does not accommodate. For this reason, a general neural network and agent architecture to cope with such problem domains (Joyce and Lewis, 1999a) was designed and implemented.

6.8.3 Agent Architecture

Congruent with Nigrin's proposals, it was decided to explore a model of reinforcement-based learning in order that the architecture can be generalised to situations unlike the MAVIS2 system where input to the agent is not accompanied by explicit teaching vectors detailing the correct classification or output. The MLP networks were simpler to implement, since no external machinery was required beyond that which processed queries (e.g. the application-specific layer).

In the localist network, the architecture is complicated somewhat by the need to "internalise" reinforcement so that the target vectors are not presented as is, but used to generate an indicative reinforcement signal showing correct/incorrect response choices. This emulates the MAVIS2 'environment'. The MLP networks are simply presented with the target vector (as is usual). Figure 6.5 shows the architecture designed for MAVIS2 agents. Note the use of train/query vector.

In Figure 6.5, the bottom layer is the connectionist network (the localist network architecture is shown). This divides in two with a mediating category choice layer (called the C layer). The response network receives input from the C node causing one rC node to be activated. In effect, the response layer must learn a mapping from binary vectors on the rC layer to classes on the rR layer using associative search.

A complete description of the mathematics of both layers is given in (Joyce, 2001a) along with the description of the two error measures used in back-propagation training (see Volume II of this thesis). At any moment in time, the agent can be switched between training and performance modes. The algorithm for processing an input is identical in both modes. The application specific layer generates reinforcement and indicates whether the current input is a query or a training vector. If the the latter, a switch causes the response network to behave differently causing it to search for outputs until reinforcement is rewarding.

During training, the agent sends no output to the MAVIS2 system. In performance/query

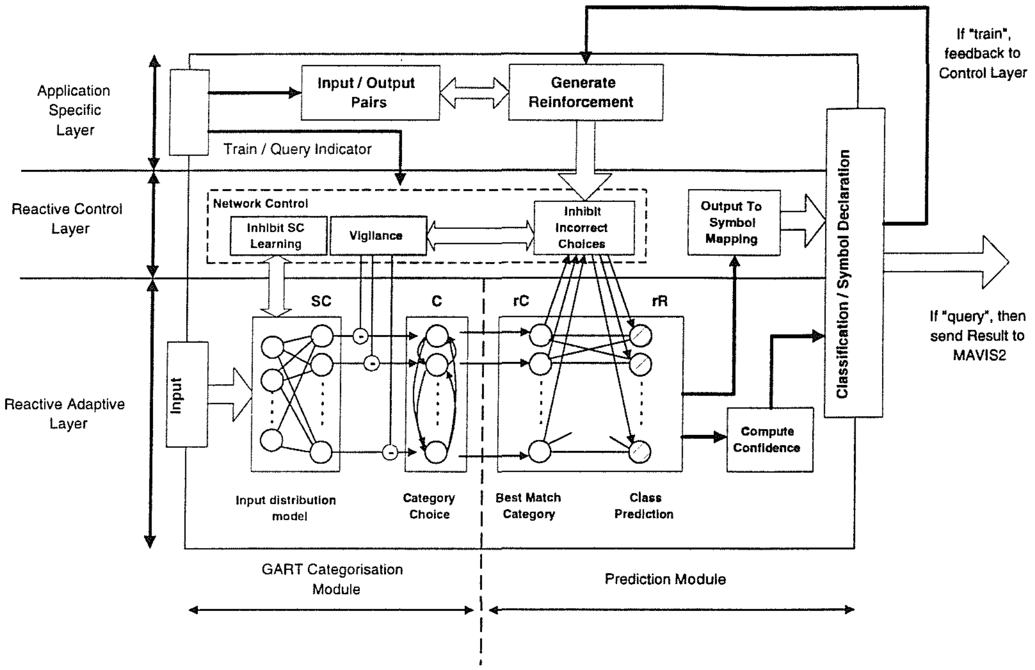

Figure 6.5: Agent Architecture Incorporating Neural Network for Classification Tasks in MAVIS2

mode, the agent learns no new associations, but sends the resulting *rR* signal vector (indicating the agent's decision on the most likely classification) to a confidence assessment and symbol look-up table. The confidence assessment normalises the highest output (most likely classification) against the mean output of all nodes. If all nodes in *rR* are outputting roughly the same signal strength, then the classification declared will have low confidence (because the network could not decide decisively). However, if the winning node is significantly higher than the mean of the others, then confidence is reported to be high. The symbol look-up table serves to map one-of-c classification vectors to MAVIS2 internal labels which have no meaning except to individuate a class in the MMT to the rest of the query processing system (e.g. they are unique but meaningless strings of alphanumeric characters).

6.9 Comparative Classification Performance

In this section, the pattern classification performance of the localist and MLP networks are compared. This testing had a number of purposes. Firstly, the ability of the three techniques to classify based on simple measures of correct classifications reported over training data and unseen queries. Secondly, the reliability of the classifiers. Part of the difficulty in training MLP classifiers is knowing when to terminate training to achieve optimal performance. In MAVIS2, this is decided based on generalisation ability on unseen vectors. The parameters for the back-propagation algorithm are also difficult to establish in a uniform way. This can be seen functionally as the output of a classifier given by:

$$\mathbf{O}(\mathbf{W}^*;\mathbf{x}) \tag{6.1}$$

where $\mathbf{W}^*$ denotes a matrix of optimal parameters for the classifier (the weights of the neural network) and $\mathbf{x}$ is an input vector (e.g. a query). Linked to this, the parameters $\mathbf{W}^*$ supervene on other parameters. In the case of neural networks, the final form of the weights depends on (at least) three factors: the scalar learning rate during learning of $\mathbf{W}^*$, the number of hidden units and the duration of training (that is, if iterative approximation methods are used, how many iterations are allowed before we deem that $\mathbf{W} \approx \mathbf{W}^*$). So, for every classifier, the training of that classifier and its consequent performance supervenes on the the learning rate (η), number of hidden nodes P and the number of iterations τ of the training algorithm.

The MLP requires the correct setting of the number of hidden nodes in the network. This well-understood parameter of the MLP has two significant ramifications:

1. too few hidden units, and the network will over-generalise, similarly to kernel-based density estimation techniques, too few kernels results in a poor approximation of the function

2. too many hidden units, and the network will over-fit the data – the network will classify vectors present in the original data set but fail to generalise to unseen vectors which are close to the training set in feature space.

Firstly, then, we have the problem of determining the number of hidden units. Universal approximation theorems inform us that any arbitrary real function can be approximated by a neural network, (Cybenko, 1989) cf. (Kosko, 1992) pp. 199, but do not tell us the required number of hidden units to approximate a function to any accuracy degree. The limiting case being one hidden node for each feature vector in the training set.

Hidden units define the size of the weight matrix $\mathbf{W}$. As a result, the complexity (e.g. dimensionality) of the weight space is usually vast and the error surface (expressing the change in error E with respect to each weight) equally as difficult to search for the global minima $E = 0$. Now, we are presented with another task. Finding a set of optimal parameters η^*, P^* and τ^* for each collection of multimedia data presented to the system. In effect, for any given instance of a problem, we could apply the classic machine vision methodology and manually build and tune a neural net classifier i.e. manually engineer η^*, P^* and τ^* from an image ensemble.

An option to automate this process is to search this parameter space, assuming that there is some function governing them and that they are independent of each other and the data set. As an example, we might consider employing an evolutionary algorithm that optimises the network for a given problem. This is, however, computationally expensive and the assumption that each parameter is independent of the others and the data set is highly suspect.

Other works have focused on particular components of the problem. For example, Messer and Kittler (Messer and Kittler, 1997) review methods and propose a novel technique for determining optimal numbers of hidden nodes in a given problem. The results indicate that using a combination of input feature and hidden node saliency in contributing to the output performance, a network can be derived which performs within a certain tolerance and has a minimal number of hidden units. A result (a set of parameters for the network architecture) obtained by Messer and Kittler's algorithm on one data set will not generalise across different feature sets, since each data set represents a different non-linear optimisation problem.

In the tests performed on the MAVIS2 agents, no attempt was made at automating the task of parameters choice for learning. During multiple trials with test data sets, the time taken to reach an optimal level of classification performance was measured and used as an indication of reliability which might be informative in future work that attempts to automate such parameter setting.

6.9.1 MLP Learning Parameters

The parameter space of the MLP techniques is taken to be the number of hidden nodes and the learning rate parameter. Tested combinations of 2, 4, 8, 16, 32 and 64 hidden nodes with learning rates of 0.01, 0.1, 0.5 and 0.9. In total, both SSE and cross-entropy error functions are tested with one of 24 possible network configurations.

6.9.2 GART-based Network Parameters

The local network agent has a simpler parameter space, namely the vigilance parameter, which dictates how precise categorisation should be. For comparative tests, the vigilance parameter was set at ten values in the range [0,1]. Note that because the algorithm is constructive, no other parameters need be set by hand. The GART network has a standard deviation parameter (or radius) for the Gaussian kernels/category nodes. However, all data was normalised between 0 and 1, and this parameter set to 1. The GART algorithm then gradually shrinks the kernels (rapidly at first) to fit the data as sufficient exemplars have been presented. The learning rate for the response network (since it is simply an associative single layer net) can be set according to the application domain and in some cases, a fast 'one shot' learning rule might suffice (e.g. setting all weight learning to be instant and the learning rate is therefore unity). Here, the signal correlation is the strength and direction of the update and a learning rate of unity was used. Weight normalisation after periods of learning activity enabled the weights to be kept bounded. In principle, both the radius and the learning rate for the associative network can be optimised. The vigilance is the only parameter which is varied. Recall from section 6.8.1, that the agent must perform even in the absence of "good" data which would enable robust classification. An autonomous agent should attempt to provide the classification service even when the data is weak. It should be capable of learning incrementally, so it can improve if or when more data becomes available. For this reason, no attempt was made to optimise the local GART/associative search network parameters and it was tested 'as if' against the more robust MLP-based classifiers.

6.9.3 Data Sets

The data sets chosen were: a) iris data set from UCI(Blake and Merz, 1998) b) HSV-based histogram features from a MAVIS2 data set c) RGB-based histogram features from a MAVIS2 dataset d) 3 clusters of Gaussian distributed data. The rationale was as follows : a) is a standard machine learning data set, and is of relatively compact feature dimensionality with adequate samples. Both b) and c) are grossly under-determined datasets, each containing 208 samples (one per image) of 27 and 96 features respectively – this represents a case where the agent must try its best to perform classification in a difficult multimedia problem. The actual data is HSV (27 features per image) and RGB (96 features per image) colour histogram feature data obtained from the Victoria and Albert image data set, which was used as a test bed application for the MAVIS2 system. Test set d) is a simple two-dimensional problem, which contains enough samples (150) to train on and is present to test whether the classifiers and agent are working as expected.

6.9.4 Performance Measures and Training Schedule

Each result presented is the average over 10-fold cross-validation testing (Cohen, 1995); this enables us to assess the "memorisation" of the training set and (more significantly) its generalisation performance on unseen data. Validation sets are generated from 10 independent samplings drawn from the original data and are kept back during training. This ensures that generalisation performance on unseen queries is not attributable to fortuitous choice of a hand-picked validation set. Each data set is shuffled into random order and then divided into 10 sets (folds). The classifier is trained on 9 of these. The classifier memorisation performance is tested by presenting the data from the 9 training folds and generalisation is tested by presenting the held back fold. Then, a new classifier is generated and trained and a different fold held back for validation. This is repeated for each of the ten sets, and classification performances averaged over these 10 different training conditions. For each setting of MLP learning parameters, there are ten classifiers trained and the performances are averaged.

Performances are measured as follows. For each configuration of folds, the agent is trained by passing through the 9 training folds. One pass through each training pair in the 9 folds of training data is called an *pass* through the data set. Training proceeds by propagating errors back after each training pair is presented (instead of batch learning where errors are accumulated over the whole pass). An epoch is 50 passes through the training set. After each epoch, the 9 training folds *and* the generalisation test (the held back fold) are presented to the classifier and the classification performance measured on both. This is repeated after each training epoch. Performance measures are therefore on *actual* classification performance, rather than any internal error measure (e.g. by summing the output errors over an epoch). Generalisation and training performances are continually recorded and the result presented is the best performance obtained during the entire training period (which was 50 epochs).

Naturally, such an exhaustive search of the parameter space and training in the fashion described is impractical for an agent in practice, but it reveals the classification performance and training properties while also illustrating the difficulty of automatically training classifiers in multimedia environments. These tests are to assess the classification performance rather than suggest a reasonable method for training MAVIS2 agents *in situ*.

We present the results in the table below. The first column indicates the network used; cross entropy error (CEE), standard sum-of-squares error (SSE) and the local GART-based agent (AG). The second and third columns are best training and validation set performances over the whole parameter space tested. The remaining columns report the number of hidden nodes and learning rate (or vigilance in the case of the local network technique) which gave the best performances, plus the mean number of epochs required to train the networks and finally the standard deviation (s.d.) of this value. A low s.d. indicates that almost always, the best performance can be achieved in the mean number of iterations, whereas a high s.d. indicates a large variation in the number of iterations required over multiple trials. It could be argued that for autonomous agents, the training behaviour must be predictable and this value should be low.

ANN	Data Set	Best TS Performance	Best VS Performance	Hidden Nodes	Learn Rate	Best Performance at epoch	Standard Deviation of Best Epoch
CEE	Gaussian	100%	100%	64	0.01	1	0
SSE	Gaussian	100%	100%	64	0.01	1	0
AG	Gaussian	98.9%	100%	n/a	0.6	1.3	0.4
CEE	iris	94.5%	99.3%	64	0.9	4.4	3.6
SSE	iris	94.7%	98.6%	32	0.01	12.7	9.8
AG	iris	95.5%	99.3%	n/a	0.7	3.2	2.2
CEE	HSV	50%	50%	64	0.1	9.5	6.5
SSE	HSV	45%	49%	32	0.01	17.7	12.5
AG	HSV	24.4%	39%	n/a	0.2	3.9	3.1
CEE	RGB	28%	46%	8	0.5	20.3	15.3
SSE	RGB	32.6%	45%	32	0.01	9.6	11.9
AG	RGB	29.9%	38%	n/a	0.7	3.2	2.2

6.10 Discussion

The table shows that for all data sets except the artificial pure Gaussian data the standard deviation of the best performance epoch is lowest for the GART/associative network technique.

However, the classification performance of the technique in its present form is below the MLP networks and further tuning is required. As expected, all three techniques worked well on the simple Gaussian data set. The HSV and RGB features sets indicate that for low numbers of features, the vigilance should be low and conversely, high for high numbers of features. For multiple-independent attributes and sum of squares error, the learning rate parameters seems to be less predictable.

More importantly, the implications for implementing classification agents in MAVIS2 are that despite poorer performance, the local GART-based network with a simple associative map is easier and quicker to train in situ. In the prototype implementation of the MAVIS2 system, the GART-based agent was implemented because of its ability to perform reasonably and only the vigilance parameter needed setting. In responding to queries, the most important problem that arose was while classifications were correct, the confidences reported were usually very low which would present problems for the future development of a fusing method based on probabilistic measures of class membership (if sufficient agents were trained over a variety of media types for each class, then voting could be used and this confidence measure is less significant). In addition, the prototype of MAVIS2 required agents to train quickly on the data available at the moment the system starts. The local network agent trained consistently faster in the comparative tests above and even with the arbitrary setting of radius (for the GART network) and the learning rate (for the associative learning net) the agents still performed well. Future work will need to focus on how an MLP network might be configured to cope with such "sufficing" tasks, where good performance can be guaranteed without the exhaustive search of the parameter space. While the data set available during the development of the MAVIS2 prototype is believed to be typical (e.g. under-determined in terms of class exemplars) it is possible that the GART-based agent performed well in a unique case. The exhaustive testing described above was an attempt to refute this claim, and performances obtained would suggest that the heuristic approach used to build the agent's classifier might be a fruitful basis for future development.

Further, some diagnostic data was collected to ascertain how difficult picking parameters for the MLP or the vigilance parameter for the local-network based agent might be. The dataset used for the example presented below was produced from a subset of the Victoria and Albert Museum Artifacts image base. This resulted in 100 individual feature vectors (samples) of 23 features, unevenly distributed amongst 12 classes. Foley's heuristic suggest that three times as many samples *per class* are needed as there are features. Hence, for a well-determined dataset we seek over 800 samples. However, this serves as an interesting problem to examine the conditions of the parameter space which must be searched and is not untypical of the scenario we might be presented with in a real multimedia retrieval problem.

The surface shown in Figure 6.6 represents the best classification performance on the validation and training data sets. Note the scale of the Hidden Units axis. Recall that the validation performance is a better indicator of the network having learned a model of the data

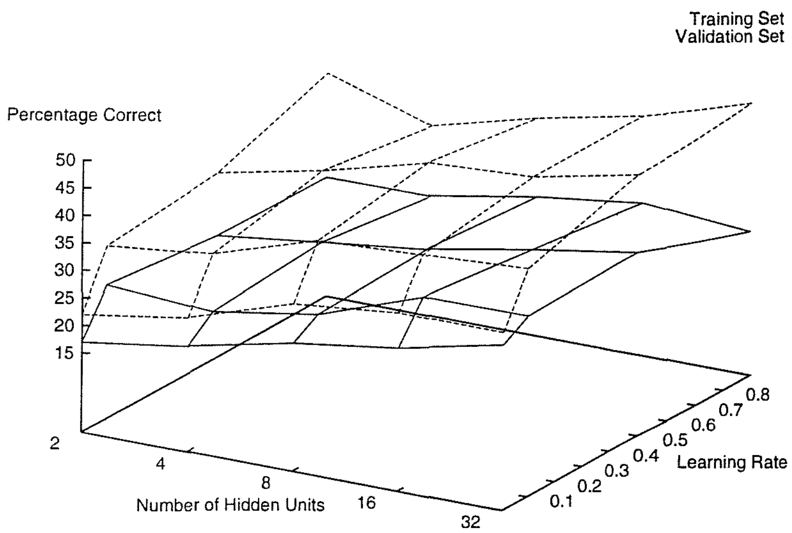

Figure 6.6: Validation *and* Training Set Performance – Parameter Space for MLP trained using back-propagation with Cross-Entropy Error Function

as opposed to simply "memorising" the training set (which results in poor performance on unseen inputs). The qualitative properties of the surface are that the best results appear to lie at the extreme of the parameter space, with the exception of one other peak result at $\eta = 0.4$ and 8 hidden units. The surface certainly does not indicate a desirable global maxima, clearly distinguishable from the other network configurations.

This suggests that a further search process (other than exhaustive search) over the space of MLP parameters might be difficult (and more so if extensions to back-propagation are used such a momentum) due to the lack of obvious maxima in the classification / parameter space above.

A similar parameter space for the localist technique on the same data subset is shown in Figure 6.7. Note that at $\rho = 1$ the network is over-fitting, and therefore should be discounted. The optimum value would appear to be 0.75 since the network performs equally well on both training and validation data.

The trend in behaviour is more predictable on the same data set. While certainly not conclusive, it suggests that on sparse data sets, the relationship between generalisation and memorisation of the training set is easier to ascertain.

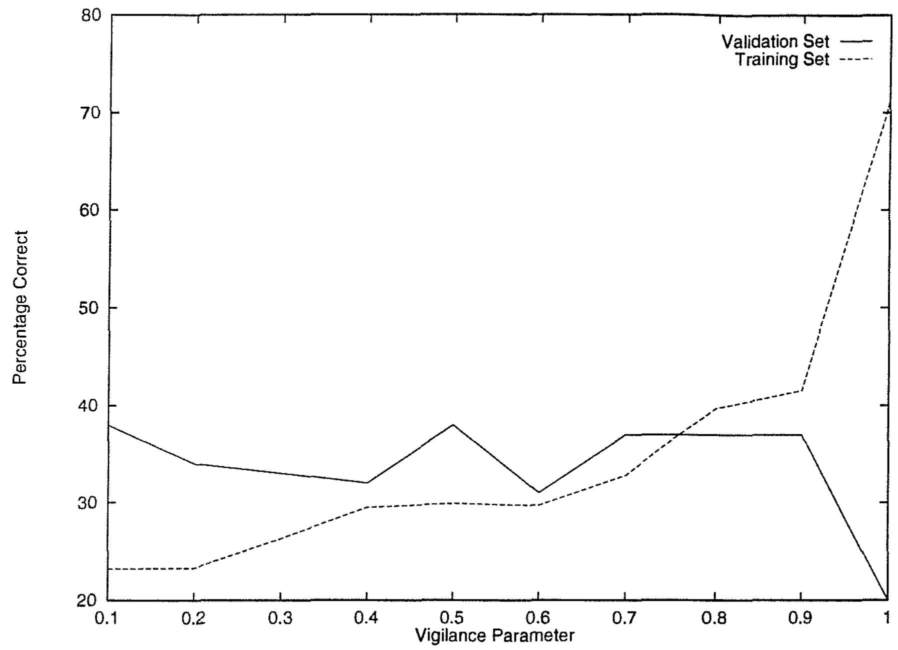

Figure 6.7: Parameter Space for Gaussian ART Network with Local Learning Associative Search Classification Network

6.11 Conclusion

The semiotic framework presented enables the treatment of multimedia information and the over-worked term "semantics" in a principled framework. It highlights opportunities for novel treatment of different media and demarcates philosophical boundaries that exist and have been typically encountered in engineering of multimedia systems. It differs from (Smoliar, Baker, Nakayama and Wilcox, 1996) in that the triadic model, which explicitly involves the user and the system's processing in the analysis (e.g. the relationship between the referrent and the intermediate feature vectors). The agent's presented were integrated in the MAVIS2 architecture and their performance assessed, enabling comparison and reflection on the design of connectionist architectures in a scenario where the agent must decide upon and train the network without human intervention (usually, a crucial part of the performance tuning process).

While much work remains to be undertaken on how to automate this further, some evidence has been presented which supports an approach where performance is dependent on one significant parameter, and the resulting training times are more predictable. The major import from computational agent principles is the notion of an autonomous device which undertakes routine work in the query process. Neural network based classification was offered as one possible method for implementing the kind of sign generalisations that humans are typically capable of.

Further work would need to focus on improving the classification performance and the confidence measures used. For example, if voting is to be used, then large numbers of simple classifiers might be used, for example, simple sub-space classifiers might be more applicable for training using single-layer networks. Using heuristic methods (such as the combination of GART and a linear associative network) achieve some of the design goals of connectionist agents, but from initial results, robust statistical performance is compromised.

The final contribution was a first attempt at defining a realisable agent architecture which enabled "internalisation" of training e.g. was flexible enough to use associative search or supervised training in the form of complete feature / target vector pairs. The associative search network implemented was a first attempt similar to (Chang and Gaudiano, 1998).

Chapter 7

A Simulation Environment

7.1 Introduction

In developing the proposals and theory of agency outlined in Chapters 2, 3 and 5, reference was made to the notion of situatedness, the embedded, continuous participation of the agent in its environment. In an attempt to study the relevant phenomena of situated agency (e.g. the establishing and manifestation of routine activity and intentional arcs), a simple environment was designed. This chapter describes that implementation, and concludes by illustrating how this simulation can be used in exploring reactive agency.

7.2 A Simulated Environment for Reactive Softbots

Most reactive architectures are designed to cope with goal directed behaviour in spatial, topological environments. (Brooks, 1997) surveyed 42 papers appearing in the journal *Adaptive Behaviour* between 1992 and 1995. From these papers, he found that 5 papers dealt with agents without spatial location or a notion of topology e.g. adjacency, proximity and neighbourhood. Brooks' motive was to demonstrate the lack of "real" situated robotics research. However, this equally demonstrates that simulations and environment models which are most closely suited to the type of agents considered in this thesis, are also somewhat scarce.

The principles motivating the choice of model are as follows :

- *Absence of spatial topology or environment geometry* – softbots may not have spatial locations. Imagine a software agent situated in its environment, but (as (Etzioni, 1993) illustrates) connected to the environment by discrete actions having no intuitive spatial interpretation such that to compute the Euclidean distance between two agents may be meaningless. An information retrieval agent will effect actions based on perceptions of the environment, although there is no logical interpretation or analogy of spatial context for the agent.

- *Adjacency (if any) is arbitrary* – softbots will, more than likely, inhabit domains where the analog of multi-agent spatial adjacency is defined only through communicative or social actions, and the medium for transmitting that information. Only if such media exist are agents aware of their peers or social groups.

- *Uncertain and dynamic environments* – although in this category, the manifestation of uncertainty is different. Firstly, the agent's virtual sensors are unlikely to be noisy. There will, however, be uncertainty in the communication medium which might reasonably be treated as noise for both the sender and recipient of communicative activity. Environments "lie" such that an agent might be informed that some action is permissible, but in fact this is an inaccurate reflection of the consequences of that action because of implicit latency of the information arriving at the agent's sensing faculties. Dynamic environments are those which affect and are affected by the agent – this property underpins the establishing of routinised activity, since this provides opportunities for sequenced actions with utility which can be repeated.

- *Softbots must be situated* – the agent's action has an effect on the environment and the environment, which is sometimes reflected to the agent. This condition was used by (Jennings, Sycara and Wooldridge, 1998) to differentiate GOFAI from agents research. (Varela, Thompson and Rosch, 1991) describe this as characteristic of *enactive* cognition. However (Maturana and Varela, 1980a) encapsulate the notion as structural coupling, being the recursive reproduction of systems (e.g. the relation between agent and environment as a couple) which are mutually affective.

- *model simplicity* – the environment must be as simple as possible to enable analysis. Brooks' criticism of 'toy' environments parallels that of 'micro-worlds' explorations criticised by (Dreyfus, 1992). These arguments are not really criticisms of methodology, but of the generalisations and illusory abilities of agents which are extrapolated from those simulations. Similarly, Agre (1988) used blocks-world simulations. In this respect, the simulation presented here enables analysis while containing the necessary features of a model enabling the study of reactive agency. To quote:

  > "Following the tenets of interactionist methodology, the focus is not on complex new machinery but on the dynamics of a relatively simple architecture's engagement with an environment" – (Agre, 1997),pp.105

These ideas are not difficult to imagine, particularly in the context of distributed information systems. Communication substrates or infrastructures are rarely fault free, so the agent must be tolerant to "noise" or uncertainty in the information conveyed to it by such a substrate or infrastructure. The source of this noise is irrelevant. The fact remains that either in the sensor or the environment, the information used for perception is not completely certain and reliable.

One model which *does* appear to comply with the principles above is (MacLennan and Burghardt, 1994). Their environment is designed to explicitly model evolutionary aspects of communication. It is an example of an environment which lacks an explicit topology. The model developed for this thesis was explicitly tailored to enable the exploration of artificial neural network techniques and architectures, migrated from situated robotic agents research, to be tested for viability in the 'virtual' or softbot domain. This can help answer the question of whether softbots are actually co-extensive with robotic agents in terms of a general theory of agent science as defined by (Huhns and Singh, 1998).

7.3 The Environment

The experiment here is based on the classic Skinner Box conditioning apparatus (Skinner, 1938) hybridised with the k-armed bandit models used in machine learning studies of reinforcement. An agent must learn that acting in a certain way will result in some resource or food (an appetitive stimulus) being delivered, but that alternative actions can provide punishment (aversive stimuli). Before further discussion, we note the following terminological points: a *resource* is something the agent is designed or goal-directed to collect or acquire, while a *reward* is a reinforcement that indicates no aversive consequence of an action. Therefore, this agent simulation is concerned with one goal directed behaviour (to acquire resources) while actually learning how to use its response/action repertoire to achieve this, *and* this learning process is operant in nature such that the agent can receive rewards and punishments.

In terms of *equipment* which can participate in absorbed circumscriptive activity, the agent is presented with a button which has toggle-state; it can be on or off. The agent will be goal directed by design, and it should try to collect resources from the environment. These resources are delivered when the button is depressed (on) and not when it is off. However, the agent has no model of the relationship between button states and consequences.

With this very simple experiment, the agent would (trivially) have to learn that depressing the button and leaving it in this state would result in maximal resource collections and, presumably, it would receive no punishments. This is, perhaps, the simplest stationary deterministic environment and easily soluble by traditional reinforcement learning and artificial neural network techniques. Similar experimental apparatus were described by Spier and McFarland (1998). They additionally tested an outcome-devaluation effect using a Skinner box-like experiment, but were concerned with the relationships between necessary cognitive capacity and performance. They also use drive reduction (e.g. a lack of arousal or agitation) as a modulator of reinforcement value – see Chapter 8.

However, the following changes are made in this thesis. Whenever the button is depressed (i.e. in the 'on' state), a resource is returned. However, there are (at any moment in time) a finite number of resources available, which is decremented each time the button is in the 'on' state. By analogy, imagine a softbot agent trying to secure time on a multi-tasking operating system. A finite number of resources (time slices or quanta) exist and one agent, on behalf of a user, cannot possess them all despite its requests (the button being in the depressed state).

The environment avoids being maximally exploited by trivial greedy policies (e.g. press and hold the button indefinitely) by delivering both resources *as well as* rewards/punishments. At any moment in time, the environment can deliver a resource, but in addition, will deliver a neutral or punishing stimulus (by analogy, the agent might be given a shock in addition to the resource). This is governed by a simple rule; as the resources are depleted, the environment is more likely to shock the agent to prevent greedy policies from being learned and established as routines. Thresholds dictate the levels at which the environment will deliver punishments as well as resources.

The level of resources available is reflected to the agent via a communication medium. In the Skinner box analogy, it might be a light indicating the level of resources currently available. The metaphorical light represents the level of resources in the environment, and the agent must learn that it is indicative of the probability of punishment (aversive stimuli), since as the consumption of resources increases, the agent is more likely to receive a punishment.

In the operating systems analogy, in order that the environment (the operating system) is not drained of resources, a sustainable number of free quanta are kept available to prevent a single task dominating the processor.

However, as noted before, this 'light' is not entirely reliable, and does not provide deterministic evidence of the chance of aversive stimuli being delivered. Figures 7.1 and 7.2 show the apparatus of the environment and the "agent requesting" metaphor respectively.

The fact remains that the agent's action should be informed by a perceptual indicator – the light. The probability of a punishment is proportional to the light's state, but with stochasticity. This perceptual information is inherently noisy and not necessarily enumerable. If the button is on, and the resources are below a threshold, a punishment is delivered. A resource might also be delivered, if any remain, but the agent is clearly "told" that this resource came at a price.

Finally, the environment is regularly refreshed. The original level of resources is replenished and the environment state reflects this, both in terms of what the agent can perceive via the 'light' and the punishments delivered along with resources.

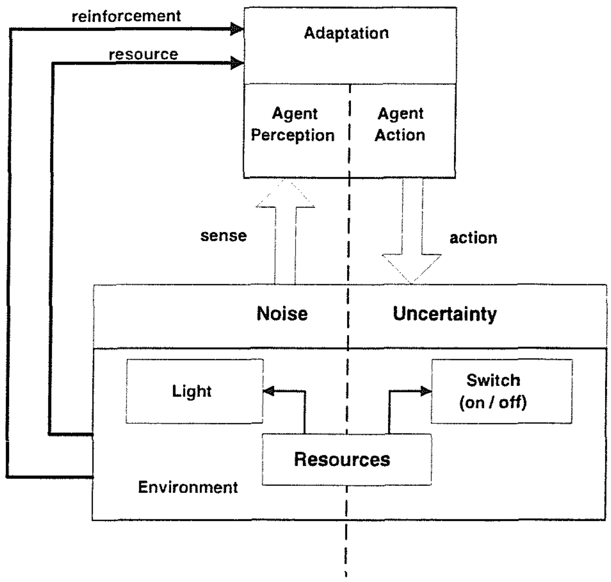

Figure 7.1: A schematic of the simulation environment

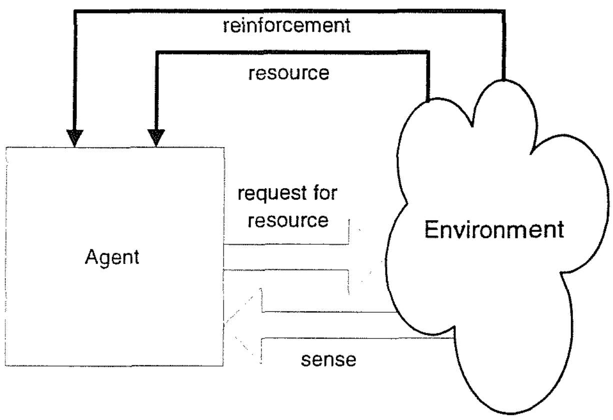

Figure 7.2: The traditional "agent metaphor" of resource allocation in an uncertain distributed system

7.4 Demands of the Environment

The agent must therefore adapt and learn to exploit the environment to attain its goal. This implies that:

1. the agent must learn relationships between the perceptual information available to it (i.e. the light state and the button state) and the consequences of acting
2. the agent must also learn the 'mechanics' of available equipment. That is to say, the button has state and the light reflects probability of punishment
3. the agent's actions will directly affect the success of a learned policy
4. the agent establishes routines of regular behaviours which reflect the risks of taking actions in different environmental states

7.5 Mechanics of the Simulation

As stated before, the button is really a switch and has two states; on or off. If the button is on, a resource is delivered and if the button is off no resource is delivered. In terms of disembodied softbots, we might reasonably see the button as a mechanism which causes a request for a resource. The resources are diminished linearly, with every resource dispensed.

The light or indicator of the resource level state, is given by a function of the resource level. Let $rs \in [o,n]$ denote the integer number of resources currently available, $ls \in [0,m]$ denote the integer-valued light state and $\tau_{danger} < \tau_{neutral} < \tau_{safe} \in [o,n]$ denote the thresholds such that:

$$ls(t) = f : rs(t) \to [0,m] \tag{7.1}$$

where f is a discontinuous function of rs mapping the resource state to the light state according to the thresholds, and m denotes the maximum integer indicating the "safe" state of the light for example with $m = 2$ then $\langle (danger, 0), (neutral, 1), (safe, 2) \rangle$. Figure 7.3 shows an example where $\tau_{danger} = 5$, $\tau_{neutral} = 10$ and $\tau_{safe} > 10$.

The following environmental rule is defined: if the light indicates "safe", then the environment will almost certainly *not* deliver a punishment with the resource (analogously, the food is delivered with no aversive stimulus). If the light state is "neutral", then the probability of a shock being delivered (alongside the unit resource) is approximately 0.5, so an agent would be taking a chance if it opted to continue requesting resources when the light shows "neutral". If the light state is "danger", the agent will almost certainly receive a punishment and *possibly* a resource (depending on rs).

This is implemented as sigmoidal function of $ls \in [0,2]$. The implementation is as follows. Using a uniform random variable v, and denoting shock (aversive stimulus) as $r = 1$ and no shock as $r = 0$, the probability of $r = 0$ being delivered is:

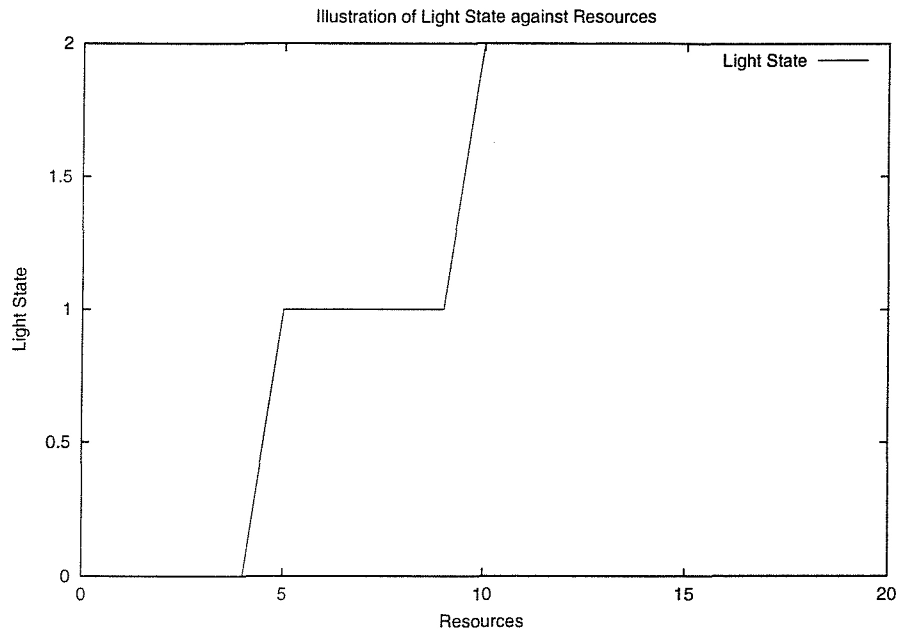

Figure 7.3: Example of light state as a function of resources

$$\Pr[r=0 \mid ls(t)] = \begin{cases} 1 & \text{if } B(ls,c) > v \\ 0 & \text{otherwise} \end{cases} \tag{7.2}$$

where the function B is a transformed sigmoidal function, and c is a variable indicating the certainty of the environment such that as $c \to 0$ then:

$$\Pr[r=1 \mid \text{'danger'}] = 1 \tag{7.3}$$
$$\Pr[r=1 \mid \text{'safe'}] = 0 \tag{7.4}$$

As c increases, these probabilities shift until the light state reflects very little about the probability of shocks being delivered for example:

$$\Pr[r=1 \mid \text{'danger'}] \approx 0.5 \tag{7.5}$$
$$\Pr[r=1 \mid \text{'safe'}] \approx 0.5 \tag{7.6}$$

This is implemented as:

$$B(ls,c) = \frac{1}{1+\exp\left(\frac{-(ls-1)}{c}\right)} \tag{7.7}$$

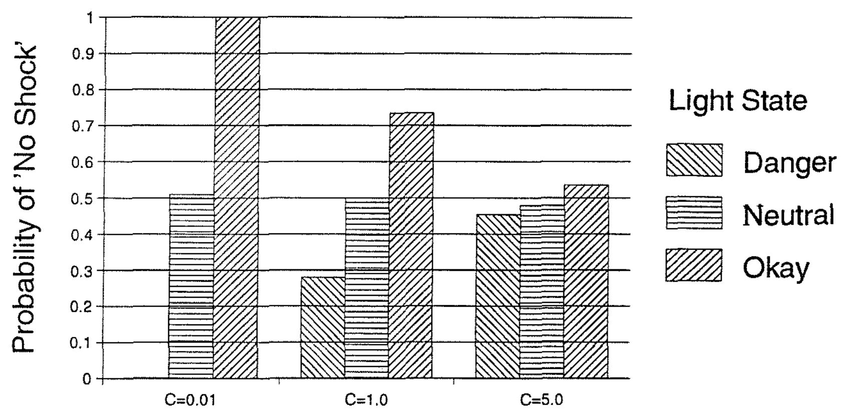

Figure 7.4: Histogram showing the probabilities of 'no shock' for different certainty values

---

Figure 7.4 shows the probability of $r = 0$ against values $c = 0.01$, 1.0 and 5.0 respectively computed from 1000 samples. Note that at $c = 5.0$, the light state has practically no utility in aiding an agent in deciding whether a shock is likely to be delivered, since the probability of shocks is practically equal with all light states.

There is no dispute that this environment is Markov. However, the utility of the light can vary, and in addition agent's actions change the probabilities of shocks. An agent must therefore assess both the utility of the light with respect to the environment, and then decide whether to proceed with resource requests or wait. The probability of getting a reward, given a light state, can also be shifted over time by altering c. Most learning algorithms assume that the environment will be stationary. This simulation enables non-stationary qualities to be explored by: varying c, changing the thresholds τ or altering the operational properties of the equipment (the button).

7.6 Agent Goals

We are now in a position to state what the agent's goals are. The agent must establish a set of behaviours which enable it to 'survive'. That is, some internal measure of goal completion (e.g. energy) which measures how successful it is in negotiating the environment to collect resources. Recall that it *cannot* simply maintain a greedy policy for two reasons. Firstly, this will incur a number of punishments. Secondly, we wish to encourage the agent to learn a maintainable routine, which might (for example) be something analogous to the proposition: "collect resources by holding the button on until the light state indicates (from experience) that it is too risky to do so; then adopt a routine which minimises energy loss until the environment is refreshed". Further details of how the agent's control architecture needs to be configured will be given in later chapters.

7.7 Agent Actions

The agent is provided with three actions :

1. NO-OP : take no action

2. PRESS : activate switch, causing it to be in the "on" state if not already

3. RELEASE : de-activate switch, causing it to be in the "off" state if not already

The agent does not understand the relationship of its actions to the environment and equipment state. For example, there is no implicit coded knowledge about the reciprocal relationship between the PRESS and RELEASE actions. The agent will also pay a cost for each action. To effect either PRESS or RELEASE, incurs some penalty in terms of internal energy, but NO-OP costs nothing. The aim of this is to encourage the agent to recognise situations where doing nothing is the best policy to preserve internal energy and not incur further punishments.

7.8 Performance Measures

It is necessary to assess the behaviour as a repeatable phenomenon, since the design goal is not to learn and perform in the environment once, but to sustain performance *and* necessary learning during the agent's lifetime. The agent must acquire learned sequences of actions, and be able to repeat them as and when appropriate. The agent must minimise punishment, but attain its goal of collecting resources. In order for this to happen, the agent will receive some punishments, because they are one of the indicators of an action's consequences. However, we require a joint measure of performance that includes the punishments received and the resources collected.

Since the agent is attempting to sustain homeostatic state, measures based on optimality of resource collection are inappropriate. This will be evident when the agent's control architecture is explored. Briefly, the agent's internal energy is used as an indicator of goal completion, but to encourage the establishing of routine actions and deter greedy policies, the secondary reinforcement (a joint measure of goal attainment and punishment) is modulated by an outcome-devaluation mechanism. This means the agent will not quickly learn to adopt state/action pairs associated with resource collection when the drive is satiated. Analogously, if an organism has just eaten, it is not as motivated to find food.

A *cycle* is defined to be a period of time between refresh periods where resources are present in the environment to support the agent. For example, if the environment is replenished every 100 iterations of the simulation, then a cycle is 100 iterations in length. For each cycle, we measure the agents punishments and mean drive/agitation level. These criteria will measure both the agent's ability to learn its task while ensuring it does not simply learn the trivial solution. The trivial solution will yield a first cycle performance, and then zero punishments and zero resources in all subsequent cycles. Also, in assessing agent performance, we may wish to discount the first few cycles, since the early exploratory phases of learning and participation may skew results.

7.9 Conclusion

This chapter described the simulation which is used to test the hypothesis that routine activities can (and should) emerge from agent/environment interaction. It is necessarily simple and contrived, in order that these properties can be observed and analysed. While simple, it enables flexibility and can be mapped onto computational tasks such as bidding for resources in a time-sharing system, or scenarios where there are finite resources over a given period of time. The details of necessary and sufficient architectures to enable an agent to participate in such an environment are investigated in the next chapter.

Part IV

A Phenomenological-Connectionist Agent Architecture

Chapter 8

Integrating Perception and Action in an Agent Architecture

This chapter attempts to show how the principles and mechanisms of chapters 3, 4 and the theory of agency in 5, can be brought together into a coherent control architecture for an agent. It begins by describing the issues that need to be addressed, and then advances a principle of 'metalearning' (Doya, 1999a) which attempts to partially explore the reactive component of the kind of architecture that (Sloman, 1996; Sloman, 2000) describe, that is one that implements and is functionally dependent on control state.

Experiments in the simulation environment (see section 8.8 below) show that the problem of controlling an embedded agent requires more than constant setting of the parameters. Thus far, two variants of reinforcement learning have been explored. The computational (e.g. machine learning) basis of these techniques has been well explored, but the connectionist and 'bio-behavioural' aspects less so; see e.g. (Doya, 1999b).

This chapter moves towards an architecture to support such connectionist agents using bio-behavioural and cognitive principles. However, identifying goal directed behaviour and its implementation is more complex, (unless these are all implicitly coded for in a formal logic). Then, as for SAT, a computable machine is compiled to execute accordingly. In this thesis, both goal directed behaviour and learning were examined as being constrained by cognitive and biological correlates, for example the collection (Levine and Leven, 1991). Therefore this work can also be seen as a first step towards validating Doya's 'metalearning' proposal, which relates neuroscientific and cognitive evidence to autonomous agent behaviour. A't various junctures, reference to and descriptions of computational models will be made. Specific models from the literature of cognitive neuroscience will be explained where necessary. The chapter concludes by describing the aspects of agency which have been specifically addressed and those which could form the basis of future work.

8.1 Introduction

In this chapter, a vertical layered architecture is described where higher levels exert control over the behaviour of lower levels. However, this is not the kind of vertical arrangement described in (Müller, Pishel and Thiel, 1995) and especially (Müller, 1996) pp.52, where vertical layers take overall mutually exclusive control of the agent. In the architecture developed here, the control layer presides over the reactive-adaptive layer, but both are simultaneously active and influencing each other. This is similar to the 'fully parallel' models of (Selfridge, 1959) cf. (Bryson, 1999) and the DAC architecture of (Verschure and Voegtlin, 1998). The architecture developed here was first applied to MAVIS2 (see Chapter 6) but reinforcement learning was implemented by associative learning based on (Chang and Gaudiano, 1998).

In terms of control architectures, (Sloman, 2000; Sloman, 1996) explore systems which have multiple levels of behavioural abstraction, ranging from purely reactive to reflective meta-management layers. In parallel, there are fundamental mechanisms, called alarms, which broadcast signals to all layers. These represent fundamental survival criteria for the agent, which require attention from all layers. Sloman also casts an alternative notion of representation on such architectures, arguing that control state is effectively a functional representation which provides an instantaneous picture of the whole system. Each part of this 'picture' is functionally relevant to the agent, and in the model developed in (Joyce, 2001b) has ascriptional content properties. For the reader's convenience, (Joyce, 2001b) is reproduced in Volume II of this thesis.

In the architecture developed, behaviours are acquired by the reactive-adaptive layer, but the higher layers are configured according to the specifics of the problem domain. Later, in Chapter 9, this is mapped onto routine activity and circumscription. In essence, the architecture can be seen as similar to the reactor/planner of (Agre, 1997; Agre, 1988) which he described as the 'running arguments' (RA) system. Agre's RA has a set of rules about behaviour which can be dependent on each other, and effect long chains of causal 'firing' based on input symbol pattern matching. This layer called the dependency network, uses the 'wires and gates' metaphor in a LISP-like language (for examples, see (Abelson and Sussman, 1985) Chapter 3). However, when inputs are similar (that is, there is no change in the pattern matching of the applicable rules in the dependency network) and no new 'reasoning' is required, the agent constructs a prototype which represents the typified input and the stereotypical output, creating a short cut from perception to action. This is then available in the next step of the simulation (in blocks world) so that there is an implicit cognitive efficiency over time (e.g. routine firing of rules causes a shortcut to be established which can be used when appropriate).

The control architecture described here can be thought of as a similar mechanism, but without the sentential-symbolic substrate. The agent learns to establish an analog of intentional arcs through adaptively modifying connectionist circuitry until a stable patterning has emerged in the agent/environment dynamics coded in the network. This is then 'reused', without as much adaptation the next time it is applicable as the agent recognises a similar input from the environment. The control layer is analogous to a device with regulates the construction of these 'shortcuts' (intentional arcs). The whole architecture is a parallel, conceptually synchronous, ongoing adaptive system.

Both the MLP and localised RBF network where implemented, but in later work (described in Chapter 9) only the local network was considered so that behaviour was made more transparent. Note that the concatenative dependencies of Agre's RA system are an artefact of the sentential-symbolic system. In connectionism, such cognitive systematicity (as (Fodor and Pylyshyn, 1988) labelled it) is not present as a matter of nomic necessity e.g. as (Sharkey and Jackson, 1994; van Gelder, 1990) argued when expanding upon the work of Smolensky (1988, Smolensky (1990). There is no ontological certainty that concatenative composition of logical statements is the only manifestation of compositionality. Agre's dependency network relies on, and is shaped by, concatenative composition of rules, whereas a connectionist system has compositionality via a variety of methods, notably the superposition of activations and the stored context-independent components coded for by the weight matrices (Sharkey and Jackson, 1994). Cognitive effort expended in running and firing of rules is Agre's empirical measure. The less the number of rule firings, the higher the use of shortcuts. Such a measure is more difficult to find for the connectionist architectures here because, in principal, all parts of the network are activated, but the relative magnitudes of activation and subsequent signalling are used to effect decision making.

8.2 Requirements

So far, the learning parameters have been set constant across the duration of experiments. An agent architecture should seek to control these parameters in response to interactions with the environment. To this end, research was undertaken to find means of achieving this from a connectionist perspective. Taking (Doya, 1999a; Doya, 1999b) as a starting point, he identified four neuromodulatory mechanisms which might be responsible for metalearning. The exploration of 'metalearning' (the control of learning parameters) led to some literature on the neuroscientific and cognitive basis of such phenomena. Unfortunately, there is some dissonance between the levels of analysis. Much neuroscientific evidence for the effects of neuromodulation divides in two; effects which modulate learning (e.g. synaptic dynamics) and effects which affect the neuron's bio-physical dynamics (e.g. increase its tendency to spike or fire).

In addition, studies at the neuroscientific level are often specific to classes of neurons found in specific regions of peripheral or central nervous systems. These results are not generalisable, and are at best suggestive of qualitative facts at the systems level (i.e. the level of connectionism). The cognitive studies which employ neural networks and apply such network theory to neuromodulatory effects, often model specific facets of systems. For example, the gated-dipole networks of (Grossberg, 1972a) which model reinforcement effects assuming two inputs which compete to activate an output in a certain direction, either high or low. The model is interpreted in terms of serotonergic systems in conditioning. The interpretation of serotonin as a neuromodulator, therefore varies depending on the neural model's architecture, its level of abstraction and the cognitive processes involved. Given this, relevant material is presented below, and an attempt made at reconciling this with the agents described.

8.3 Goal Directed Behaviour in Neural Networks

In this thesis, the connectionist machinery is constrained, regulated and monitored by a control system. This constitutes what (Taylor, 1994) alluded towards as his 'goals, drives and consciousness' of agency (from a neural networks perspective) and what (Doya, 1999a) has more succinctly called 'metalearning'. The kind of model required is one where some kind of internal mechanism relays information about goal attainment to the adaptive layer. The so-called 'control layer' undertakes this task. The implementation which follows is designed to facilitate routine activity for the simulation environment.

8.4 Outline of an Architecture

8.4.1 Motivation and Explicit Goals

In a connectionist-based agent, the specification of purposeful action is more difficult to locate, that is, it is not propositional. (Balkenius, 1995) considered a number of means to implement apparently motivated behaviour. He used the conception of a *drive* as an aversive state caused by deprivation (Hull, 1943), and further, these drives act as *intervening variables* which modulate the evoked *reaction potential* given a number of input variables; see also (Hilgard and Bower, 1966). These 'reaction potentials' motivate the agent to act. Balkenius identified a number of classes of drive, of which the following are most relevant here:

- *Exploratory drives*: similar to default drives, except these drives encourage new behaviours to be explored. Such a mechanism can be envisaged as "try something new in the absence of stimuli" or as an augmentation to the mechanism that tries to respond when an unknown perceptual category is encountered (e.g. a novel stimulus).

- *Homeostatic drives*: these are drives caused by the violation of some internal state which must be maintained by the agent. A simple example if that of hunger. If certain physico-chemical balances are not maintained, then a drive exists causing the agent to find food. This is particularly relevant since the agent will exhibit purposeful action defined as motivated by internal state.

- *Noxious drives*: if a given unconditioned stimulus, taken as a reward, is considered an indicator of satisfaction with the consequences of an action, then punishing reinforcement signals (by definition) represent incentives (motivations) for the agent to avoid the behaviour which resulted in the punishment. We might say that *a priori* we have determined certain stimuli to induce a state of dissatisfaction, and hence *implicitly* there exists a drive to avoid them.

The contribution of a theory of drive and motivation is in the pairing of agent design goals to drives. The agent in the simulated environment has a design goal, "collect resources", which requires motivational analogs. The homeostatic drive might be defined by a temporal measure of resources collected in a certain time interval and metaphorically, this may be called "hunger" for resources. In violation of this homeostasis, a state of arousal causes the agent to act. The nature of this arousal is given by an exploratory drive which forces the agent to find new behaviours. In addition, such implementations in a connectionist substrate add further to the claim that the mechanism underwrites behaviour cf. (Preston, 1993). Such variables constitute alarms (cf. (Sloman, 2000)) which are part of the functional control state.

It is worth noting the interactions between drives and motivated behaviour. In Hull's work, the reduction of a drive (by some stimulus) acts as primary reinforcer, congruent with reinforcement learning's notion of an adaptive heuristic critic, which 'combines' multiple inputs into a single modulating signal r_t. Interpreting Hull's *habit strength* as the probability of an action being selected, and the *effective reaction potential* as the overall probability of an action being selected, then Hull's systematic theory of behaviour states that the reaction potential is the product of habit strength and drive strength, causing all actions to be amplified or more likely to occur. This is analogous to exploration in machine learning terms. Bolles (1967) defines the role of drive differently, claiming that drives will increase the probabilities (i.e. make more likely to occur) of those actions which have *previously* satiated the drive. In terms of Boltzmann temperatures in action selection, Hull's generic increase in action potential is a more appropriate formulation and, in effect, is an exploratory drive.

8.5 Homeostatic State

We define a simple model of homeostatic drive, based on a metaphorical notion of hunger. If the agent collects resources, then this drive is satiated. If the agent does not, it becomes progressively more "agitated". A similar model to that used here is given in (Taylor, 1994). The key difference here, is that the energy level of the agent is unbounded and can grow to whatever level the environment supports, and routinised behaviour demands, until there are no resources left to the agent.

This is defined as an internal energy:

$$E_{t+1} = E_t - d_E E + in - C_o \tag{8.1}$$

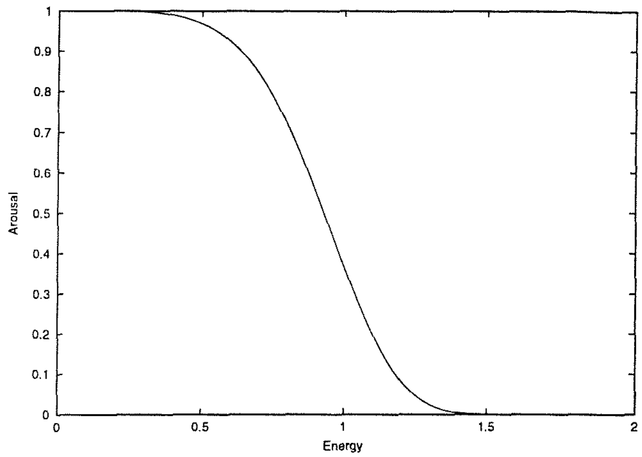

Figure 8.1: General Arousal as a function of Internal Energy; $s = 5$

where d_E is how quickly the energy decays, *in* is the last intake from an energy source (e.g. the collecting of a resource, or other measure of goal attainment) and C_o is the specification for action costs: $C_{NO-OP} = 0, C_{PRESS} = C_{RELEASE} = constant$ where PRESS and RELEASE are the agent's options for operating the toggling device that causes the environment to deliver a unit resource depending on the state of the environment. Both "active" actions cost the agent 0.25 (a quarter of a unit) of energy. Naturally, NO-OP (do nothing) costs nothing, except the passive decay of internal energy. The most an agent can gain for any one action is one unit of energy, from a successful interaction with the equipment and the environment. The passive decay rate is set according to how quickly any goal attainment fades.

A drive motivating the agent can be defined as a general arousal resulting from a lack of energy (e.g. poor goal attainment).

$$A_G(t) = \exp(-E_t^s) \tag{8.2}$$

where s is the *sensitivity* of the agent to the internal energy approaching zero. If $s = 5$, then the agent's arousal is fairly slow as $E_t \to 0$; see Figure 8.1. The agent can be made more 'anxious' about its energy approaching zero by increasing s by orders of magnitude.

Finally, the relief obtained from the last action is defined:

$$rel_G(t) = \frac{dA_G}{dt} \tag{8.3}$$

such that $rel_G(t) < 0$ implies a *reduction* of drive.

The agent has the goal of acquiring resources (e.g. time quanta in a time-sharing operating system or goods available in a market) under a constraint that it must not extinguish its energy. Each action costs the agent a certain amount of energy, and the agent perceives the environment and takes action receiving stimuli about the consequences of actions along with the next functional stimulus. An agent that takes no actions from the very start will simply experience exponential decay of its internal energy. For convenience, we will refer to the instantaneous energy at some time t as E_t.

8.5.1 Interpreting Energy and Arousal

To provide a more concrete analysis, the following analogies are given:

- E_t is the absolute level of goal-achievement of the agent at some time t
- d_E is how quickly previous achievements are devalued. For example, the agent successfully bids for or obtains a product which is consumable or loses value over time. The gain of this product is analogised by internal energy, and the devaluation or "grace period" between successful goal achievements is governed by passive decay.
- s is how easily the agent is aroused by its lack of success in achieving the goal or, keeping the energy above a certain level.
- rel_G is the amount by which the agent's arousal is satiated by a recent success in achieving the goal.

Note that we do not stipulate optimality constraints on goal achievement. The agent is defined as being a recurrently interacting entity in its environment, governed by certain internal mechanics which are goal related. However, no optimal level is specified for example, by an error measure such as the difference between desired energy to actual energy. Instead, the concern is to find observable routine behaviours, and understand the relationship between goal attainment (homeostatic state) and the environment.

8.6 Perception and Homeostatic State

In ethological explorations, such as (Tyrell, 1993), it is routine to include 'internal stimuli' which might legitimately form part of the agent's stimuli-space. Another motivation for this is that the internal energy level (representative of goal attainment) might act as a key affective state indicator, which dominates the form of the stimuli space. This effectively acts as a significant cue to the agent's internal mechanism which implements internal state selection.

In order to achieve this, an augmented input was defined so as to include goal attainment (energy) as part of the compound percept that influences internal state. Figure 8.2 shows this, where each P node receives the augmented stimulus $\langle \iota \mid E_t \rangle$. This can be similarly implemented for the MLP since both use a feed-forward architecture employing fan-out connectivity from each input node to each P node.

The input from the energy node is un-normalised, and as an effect, skews the distribution of P nodes forming the map. This was considered desirable as it enables the construction of

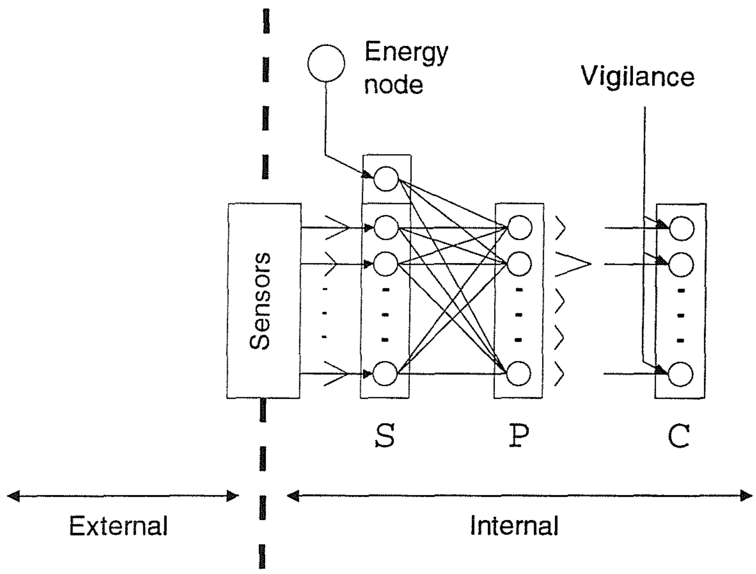

Figure 8.2: Internal Energy Node and Percept Network

perceptual schema for stereotypical percepts, which are dominated by the agents internal homeostatic state. However, it may have a questionable effect on MLP implementations. Weighted-sum activations *outside* the range $[-1,1]$ tend to drive the signal function to either the positive or negative asymptote, which for a sigmoid is 1 or -1 respectively. The derivatives of the signal function are thus very low, and weight changes, e.g. using back-propagation, tend to have little effect on weights.

In addition, a feature to implement additive noise to each of the independent environment sensors was added. As illustrated, one sensor detects the indicator light and the other the state of the switch. Each sensor contributes a certain level of uniformly distributed noise, independent of the strength of the incoming stimulus and adjacent sensors accordingly. Therefore, the final input to each of the P nodes is:

$$\langle \iota_1 + e_1, \iota_2 + e_2, \ldots, \iota_{|S|} + e_{|S|} \mid E_t \rangle \tag{8.4}$$

where e_i is typically distributed in a range which is sufficiently large to span the maximum size/diameter of the category formed by some P node. Adding sufficiently strong noise can therefore disturb one input in stimuli space so strongly, that it moves outside of its normal category P node and into another. This causes incorrect perceptual identification of the input. Analogously for the MLP, this perturbation will cause different hidden node activations to the noise-free signal.

The network will recruit new P nodes whenever a novel input is recognised. That is, the new stimuli fails to take any of the existing, committed category nodes supra-threshold for a

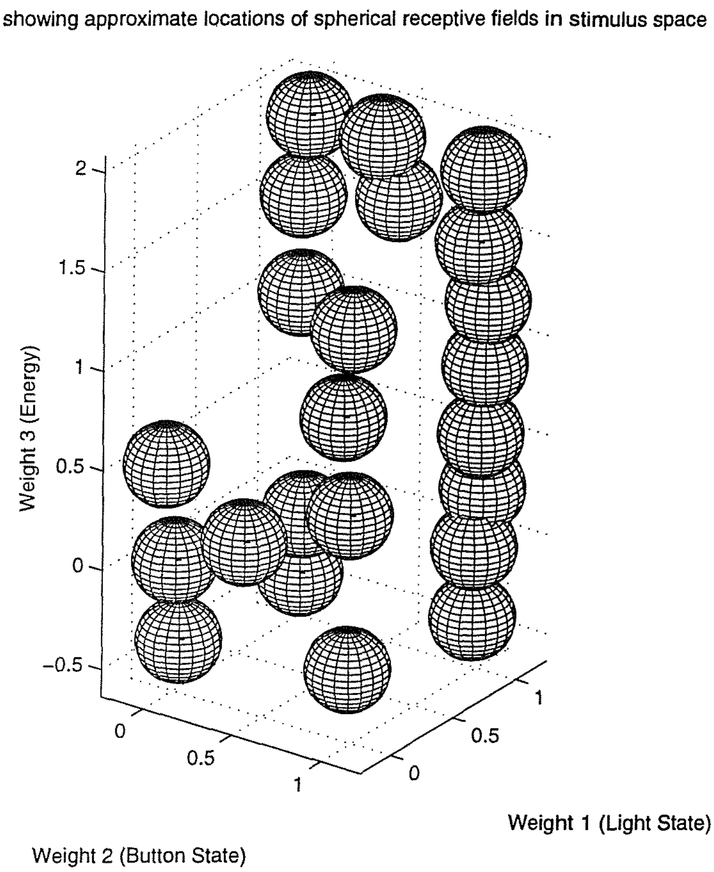

Figure 8.3: Weight space and node centres with low sensor noise

---

given level of vigilance, ν. However, the MLP uses finite resources to construct a stimuli space map, so this analogy does not apply.

Recall that any given stimulus ι generates a vector of activations $\alpha_P(t)$ in the constructed network. The number of P nodes will grow, so it is difficult to show the activation state space diagram when $|P| > 3$. Instead, figure 8.3 and (less informatively) figure 8.4 shows the locations of category node centres formed. Note the dominance of the internal energy stimulus in shaping this space. The node radius was experimentally determined and set constant at 0.2 for the experiments reported herein. The vigilance was similarly fixed at 0.2.

This approach combines all somato-sensory and external input into one 'collective' internal state category. A different approach to fusing multiple sources of information is given by (Billard and Hayes, 1999), where recurrent associative memories are used. In the strictly feedforward model presented here, the summative or collective internal state is a vector of activations on a layer of P nodes. In Billard and Hayes, this is implemented as stable activations of the recurrent fields of neurons. Their approach also enabled a meta-learning principle to be implemented for mechanisms which estimate certain internal regulating parameters for learning,

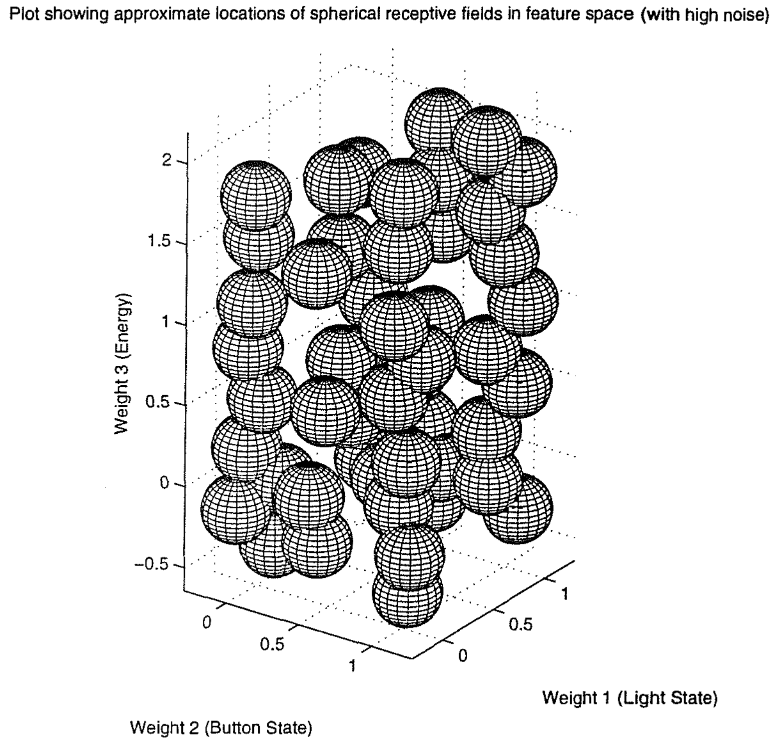

Figure 8.4: Weight space and node centres with high sensor noise

and the duration of short term memory i.e. how quickly short-term percept activations decay.

An important empirical factor in the construction of this mechanism, is that the creation of category nodes should be asymptotically bounded. After a period of adjustment and adaptation to the environment and through a sequence of different affective states (e.g. the dominance in the weight space shaping due to internal energy fluctuations), the agent should stabilise its stimuli-space map. This occurs so that novel category nodes (e.g. internal states) are not generated unless absolutely necessary. Experiments performed in both uncertain and noisy environments showed that certainty of the environment had some effect on the number of nodes formed over the duration of the simulation. This is because the agent is unable to use the information contained in percepts to take actions (e.g. the light is too unreliable) hence novel states are formed as it increasingly finds itself in progressively lower energy states, because any learned activity is not generalising over time; fuller descriptions can be found in (Joyce, 2001a) included in Volume II of this thesis.

8.7 Connectionist Implementations of Drive Effects

Given the simple drive model presented above, it is now necessary to ask how this will affect the formulation of reinforcement learning and its connectionist realisation. Congruent with connectionism's origins, as constrained by neuroscience, neuroethology and cognitive psychology, literature pertinent to this work was sought and analysed. Much of the literature is at the cellular level in that it provides analysis of effects of modulation and transmission in synapses, specific to certain functional regions of invertebrate and vertebrate animals. At the systems level appropriate here, e.g. (LeDoux and Fellous, 1995; Hestenes, 1991), there is the problem of under-constraining with implementation. The work of Taylor (1994) is very similar, particularly in his experimental method. He states that one must avoid experimental work which becomes bound to finding appropriate parameters, in favour of finding, postulating and testing more general principles which govern goal directed behaviour.

What is sought, then, is a system which enables inferences to be made about the functioning of learning parameters in reinforcement learning (Doya, 1999b). In order to proceed, some outline of the basic functional properties of biological systems is required.

8.7.1 General Functional Architecture

A recent and useful synthesis of these diverse fields was undertaken by (Hestenes, 1991) which attempted to use connectionism to explain behavioural abnormality and mental illness. His conclusions for mental illness are irrelevant here, except in connection with (Pribram, 1991), where behavioural 'traits' can be used to infer mechanisms and principles which govern model building experiments. Neither Hestenes nor Pribram reported experimental work on implementation.

Before proceeding, a functional overview is given in Figure 8.5 which is based on (Alexander, 1995). In this diagram, two major cortical areas (the limbic, motivational system and the associative cortex) are show as topographically mapped inputs to the striatum. (Schultz, Dayan and Montague, 1997) call these 'cortical modalities', and they are analogous to the input from the perceptual mechanism for the agent architecture. Parts of the striatum and pallidum/nigra structures are collectively named the basal ganglia. The output of the basal ganglia is the globus pallidus, where a variety of different sub-systems converge to either inhibit or enable certain actions or behaviours. The output of the basal ganglia is also passed through the thalamus which, according to (Berns and Sejnowski, 1996), form representations of action at different time scales. These thalamic representations are fed back to the cortical regions.

The functioning of the basal ganglia, in functionally related regions, are shown in Figure 8.6. There are clear 'plan selection' and 'plan execution' pathways which are biologically and functionally constrained. It is these interactions which are of most interest.

Hestenes architectural overview is significant because it elaborates functional detail of the basal ganglia region *but also* gives evidence for the modulatory effects and their interactions. The model of Schultz *et al* gives important instruction on how to interpret Hestenes work. The complex of labels '5HT', 'DA' and 'NE' represent the major modulator/transmitter used in the pathways, denoting serotonergic, dopaminergic and norepinephrine (or noradrenergic synapses

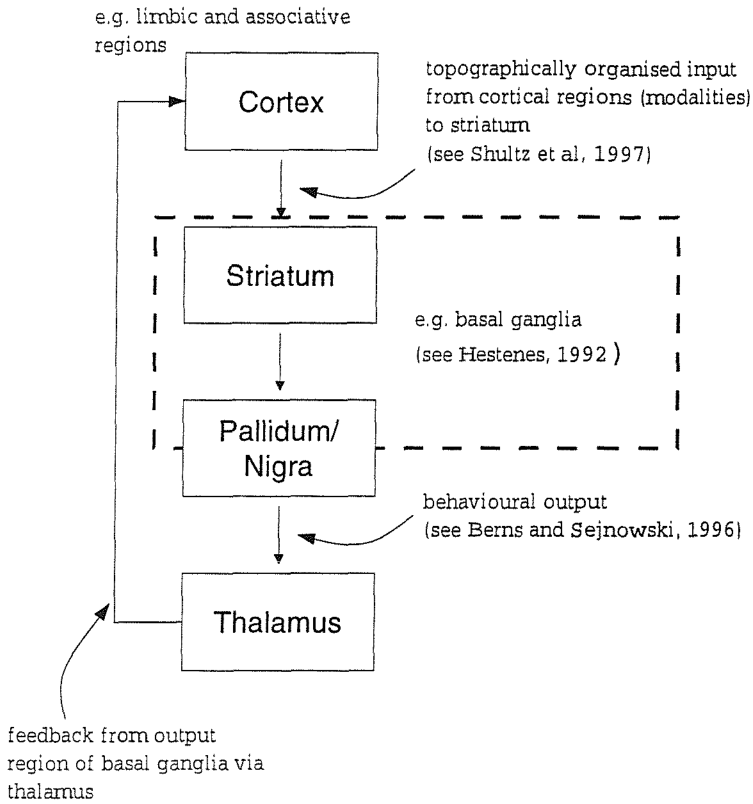

Figure 8.5: Functional organisation of basal ganglia-thalamocortical loops

and neurons). For example, the raphe dorsalis and the ventral tegmentum massively innervate the nucleus accumbens with 5HT and DA neurons respectively. The nucleus accumbens is responsible for 'gating' and selectively shutting down responses selected by the limbic motivational system and the amygdala (both responsible for motivational and emotional aspects of learning). This gating effect (Hestenes, 1991) is due to a fine balance of activity in the modulating effects of DA and 5HT neurons, emanating from the ventral tegmentum and the raphe dorsalis respectively. Hestenes also notes that this balancing effect is subtle; neither 5HT nor DA can be seen as 'cancelling' effects, with 5HT playing the role of inhibiting and DA exciting neurons targeted in the nucleus accumbens.

(Schultz, Dayan and Montague, 1997) postulate a system which mimics (amongst other things) the use of rewards that bees use when learning which flowers to return to in order to maximise pollen acquisition. Taking Schultz *et al*'s work, as well as Hestenes, the diagram given in Figure 8.7 can be deduced.

The reward, e.g. r_t, arrives at the ventral tegmentum area (VTA) (which includes the substantia nigra (SN) – a significant 'output' of DA in the basal ganglia). Simultaneously, the two cortical modalities combine at D to compute a temporal difference (e.g. as for Sutton's

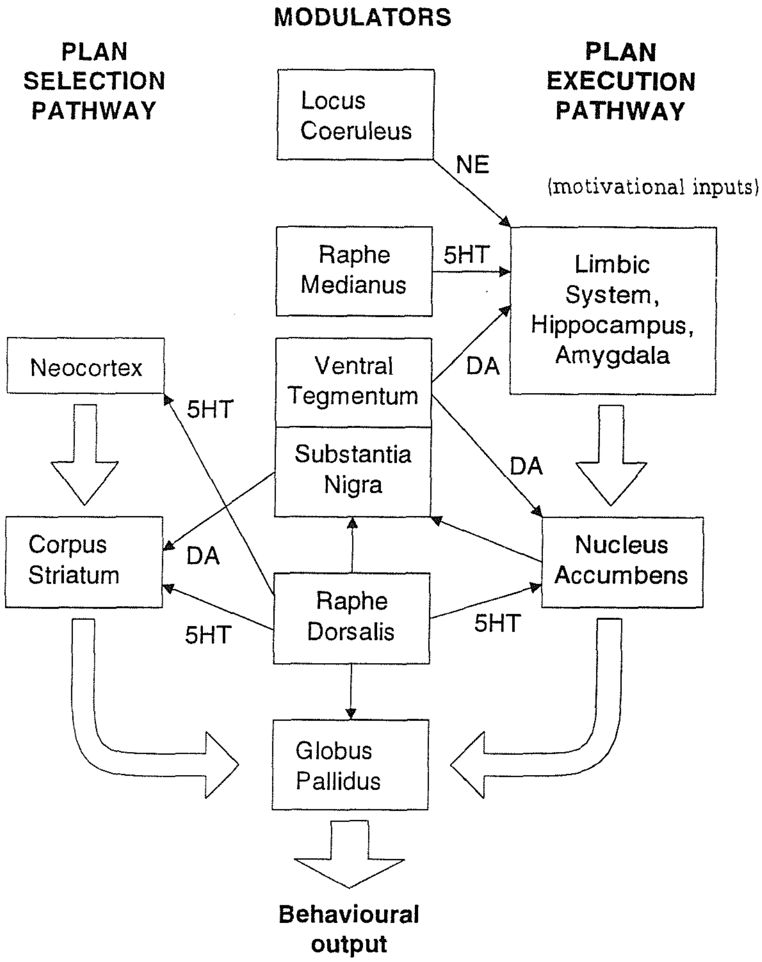

Figure 8.6: Functional organisation of contributing regions and their relationships to modulators, plan selection and execution pathways; after (Hestenes,1992)

$TD(\lambda)$ method and Q-learning) between the two states the agent encountered (at the current instant in time and the one previous). Over time, Schultz *et al*'s model adapts the weight descending from the cortical modalities to represent the value function. The output of the VTA is what Doya refers to as the *relative merit* of the last state transition (i.e. that caused by the last action).

Although lacking an interpretation in terms of modulatory effects, the networks of (Chang and Gaudiano, 1998) build on the notion of reinforcement effects without explicit reinforcement signals. These ideas were used here to build the outcome-devaluation effect used in relief calculations. Relief modulates the value of a potentially rewarding reinforcement; effectively decomposing the real-valued reward/punishment r_t.

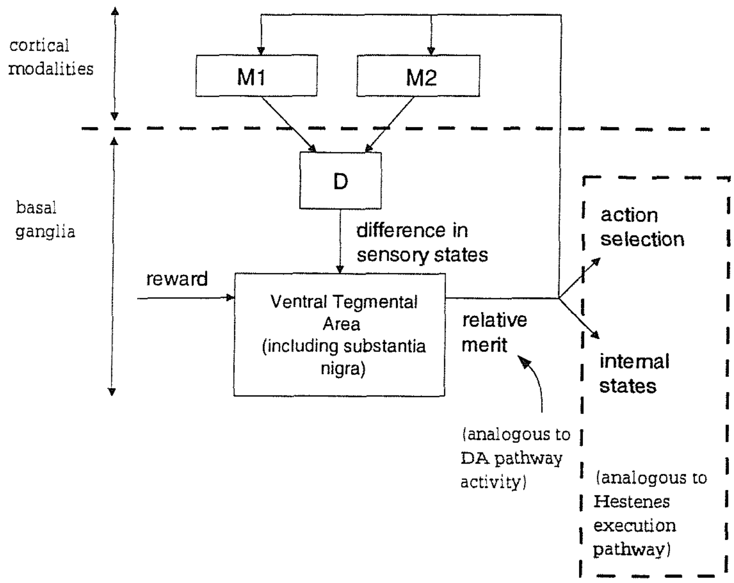

Figure 8.7: Reward Network of Schultz, Dayan and Montague

8.7.2 Metalearning

In a short position article, Doya (1999a) proposed the following principles (but unfortunately, elaborated very little on how to proceed with a model):

- the dopaminergic systems are responsible for the relative merit calculation

$$\delta_t = r_t + \gamma V^{\pi}_{t+1}(y) - V^{\pi}_t(x) \tag{8.5}$$

- the serotonergic system controls the time scale of evaluation – i.e. the amount of future reward discount γ
- the noradrenergic systems controls the temperature variable in the Boltzmann distribution used to choose actions
- the acetylcholinergic system controls the learning rate η

Using this as a guide, the literature was reviewed to establish cognitive, behavioural and neuro-anatomical correlates. Although some of Doya's proposals are crystalised in (Doya, 1999b), he has yet to produce implementations or evidence for the relationships outlined above. Initial results were positive, but it is clear that the functioning of these transmitters/modulators is far more complex than perhaps Doya had allowed for. For example, it was difficult to ascertain exactly how the acetylcholinergic system influences base levels of learning, in correspondence with the temporal dynamics of the other named systems. To proceed, some assumptions where made and the evidence gained from initial experiments with the MLP and local networks led to the model below.

8.7.3 Major Qualitative Relationships

From Hestenes (1991), Berns and Sejnowski (1996), Schultz, Dayan and Montague (1997), Shepherd (1994) and (Levine, 1991), some principles emerged. Namely:

1. the separability of the four modulatory substances is disputable; their interactions being significant. For example, (Brown and Chatterji, 1995) asserts that dopamine is a possible regulator of learning in the Hebbian paradigm, principally because of so-called complex synaptic effects (Shepherd, 1994). This suggests that cellular mechanism will hold little significance for systems-level theory because the interactions occur on scales of abstraction far beneath those of a system. (Alexander, 1995) also suggests a role for DA as a regulator of long term potentiation and depression in the striatum (the major topographically mapped input to the basal ganglia responsible for motor control). Schultz *et al* also describe how the DA neurons of the VTA have activity correlated with *only* rewarding events, hence a decrease in DA activity suggests poor predictive capabilities. In Schultz *et al*'s model, random action selection ensues. The functioning and relationships are not well established.

2. from (Usher, Cohen, Servan-Schreiber, Rajkowski and Aston-Jones, 1999), the relationship between norepinephrine (NE) (e.g. the noradrenergic systems) and the locus coeruleus (LC) is one of general attentional arousal (suggesting a tonic function) correlated with phasic 'bursts' of activity in response to novel or aversive stimuli. Usher *et al* show how the LC also affects behavioural 'output' in changing the dynamics of the output neuron (a single response unit in their model). In essence, this models the bio-physical effects on neuron activity, but ignores learning. Additional evidence can be found in (Devor, 1995) that suggests that NE reduces the activity of neurons, and makes them more susceptible to inputs from other sources. In this way, NE in the ascending brainstem can be seen as 'blocking' the receipt of punishing stimuli (e.g. pain) back from the peripheral nervous system to the brain (note the similarity to an increase in Boltzmann temperature which causes random exploration of actions). Devor notes that this represents a kind of 'stress-induced analgesia' in certain animals.

3. the nucleus accumbens acts as a 'gate' for limbic/motivational responses. This occurs when a quick autonomic response should be suppressed in favour one driven by other cortical processing, such as that produced by learned goal-directed behaviour (Hestenes, 1991) pp. 215. However, functioning of this gate relies heavily on appropriate 5HT and DA balance, suggesting a closely correlated activity implicating the LC, raphe dorsalis and medianus, ventral tegmental and substantia nigra regions. A conclusion for system-level theory is that DA and 5HT assume complementary roles which (at best) can be assumed to be in some equilibrium during neutral experience (e.g. in the absence of aversive or appetitive stimuli). In experiments reported in (Fickbohm and Katz, 2000), swim

motor pattern generators in the mollusc *Tritonia diomedea* were treated with serotonin precursor, causing the release of 5HT in the neural circuits (including the synaptic transmitter *and* the cellular mechanisms of the neurons). The effect of such treatment was to raise the excitatory post-synaptic potential, showing that 5HT as an effective transmitter, increased. Secondly, action potentials were raised, leading to increased spike rates of individual neurons, suggesting a gain-like function similar to the model of (Usher, Cohen, Servan-Schreiber, Rajkowski and Aston-Jones, 1999) in which activation function gain was increased. However, further treatment of the circuit caused an almost complete shutdown of the pattern generating circuit, supporting the claim that equilibrium levels of 5HT are necessary to maintain motor performance. Evidence beyond pattern generators and invertebrates is harder to obtain. However, Fickbohm and Katz's work suggests that both synaptic and cellular effects are present, and saturation, i.e. large deviation from an equilibrium, will eventually shut down activity. Similarly, the evidence gained from experiments with the local network in the three environments of Chapter 7, suggested an equilibrium was necessary between η and γ. This is more usefully thought of as an heuristic, rather than as direct evidence for the proposal of (Doya, 1999a).

4. the ventral tegmentum area (VTA) regulates gain in the plan execution pathway of Figure 8.6. The VTA is implicated in reward processing by being the (hypothesised) generator of the relative merit signal i.e. equation 8.5. The VTA is therefore largely implicated in DA activity. Also, in Shultz *et al*'s model, a drop in the relative merit below zero is correlated with an *absence* of DA activity; only rewarding experience caused DA activity. Further, if this relative merit is below zero, the network takes random actions, otherwise it takes the learned action. Schultz *et al*'s model suggests that in the absence of DA, random actions ensue. Correlating this with Hestenes, we find support for the idea that 5HT activity which increases phasically with LC/NE activity, thus affecting the NAC's balance of DA and 5HT, plays a role in shutting down the NAC and prevents learned action selection from taking place in favour of random responding. In Hestenes' model, vigorous limbically driven responses. The connectivity of the NAC to the VTA and the RD supports Hestenes' proposal that 5HT and DA are tightly coupled, and LC/NE projections to the RD completes the causal connectivity.

5. Both the RD and LC have widespread connections to other brain regions (the RD less so, being confined mainly to the neocortex) and importantly, the LC and RD activity appears to be correlated. From the evidence of (Hestenes, 1991) pp. 218, we can list properties which a model must possess:

   (a) RD and LC activity (5HT and NE neurons respectively) are coupled

   (b) the effects of RD/5HT activity are slow to start and terminate

   (c) RD neurons have a slow tonic output rate

(d) some network circuitry is required to simulate the effects of attentional arousal and 5HT modulation

(e) the LC has a tonic baseline activity (see section 8.10.2) and there is a phasic increase in activity during stressful situations. This suggests a joint model of LC activity and gating effects, as well as modulation of learning.

6. the functional grouping suggested by Doya (of dopaminergic, serotonergic, noradrenergic and acetylcholinergic systems) responsible for effecting reinforcement learning is given some support by (Davis, 1992) cited in (Shepherd, 1994) pp. 614. Davis' experiments illustrated that the amygdala (the proposed locus of emotional conditioning) projects to the VTA and LC invoking effects, implicating that the aforementioned neuromodulatory systems, and corresponding behavioural evidence, was arousal and vigilance increase.

Other work based on models not directly descended from the mid-brain studies of modulation (cf. Hestenes and Schultz *et al*) focus on other brain structures. (Rolls, 2000) describes the role of the orbitofrontal cortex structures as the locus of convergence for multi-modal inputs, i.e. different senses, where emotional significance is attached to otherwise unconditioned stimuli by the amygdala. (Balkenius and Morén, 2000) describe a simple network which simulates a two-process learning theory, where a stimulus is associated with emotional effects, and this association then shapes the stimulus-response association.

Both theses are important, since they separate primary and secondary reinforcement. A primary reinforcing stimulus is one where the effectiveness of the stimulus as a reinforcer is independent of another stimulus (Catania, 1992), that is, a reinforcer which satiates a drive such as thirst or hunger. Secondary reinforcers (e.g. those which are typically modelled by r_t in reinforcement learning and stochastic learning automata – see (Barto, Sutton and Anderson, 1983) for discussion) function only because they are related to primary reinforcers. So the signal r_t is secondary because it has been *a priori* assigned significance by the designer (e.g. if $r_t < 0$ then the last action was harmful and had negative consequences) who acts as the interpreter of primary reinforcers, and specifies the function which maps consequences to $r_t \in [0, 1]$.

Primary reinforcements are incident on the amygdala (see (Rolls, 2000) pp. 180) and the orbitofrontal cortex. Secondary reinforcements, those which help shape the nature of learned stimulus-response relationships in the motor system, are hypothesised to occur in the basal ganglia region (Balkenius and Morén, 2000). Here, this evidence will be used to separate different aspects of the usually conjoint reinforcement signal r_t. The fuller models of emotional conditioning are suggested as future work.

Environment	Resources	Refresh	Danger	Neutral
Hard	25	50	5	15
Moderate	50	100	10	25
Easy	100	100	20	40

Table 8.1: Environments

8.8 Results in the Environment

The local and MLP networks, with minimal control architecture *including* the outcome devaluation mechanism described below in section 8.11.2, were experimented for the simulator described in Chapter 7.

It was decided that to evaluate the local and MLP networks with their various configurations, three experiments would be designed, and the agents tested in each environment with varying parameters. This helps establish an empirical basis for the metalearning theory outlined above, and further advance understanding of the parameters that need to be governed. Extensive results are reproduced in (Joyce, 2001a) – reproduced in Volume II. The three experiment parameters are given in Table 8.1. The environments are characterised as 'easy', 'moderate' and 'hard', according to the perceived difficulty, or 'harshness', of the environment.

In each environment, the agent was given the goal parameters as follows:

- The passive decay rate for energy was $d_E = 0.1$
- The action costs for PRESS and RELEASE were $C_{PRESS} = C_{RELEASE} = 0.2$
- After initial experiments with the localised (RBF/ART-like) network, vigilance and basis-function radius were both set 0.2 (other values caused the network to behave unpredictably and the results were far worse than the MLP tests)
- After initial tests, the MLP proved to be most successful with 5 hidden units. This strongly suggests that agent is 'assigning' hidden units to regions of the stimuli space in a more localised fashion. Other numbers of hidden units were tested (3 and 7) but the resulting performance was more erratic than with 5 units. That is, the gradual increase in performance expected after the start of the simulations was slower to be generated.
- The action-selection temperature was kept constant at 0.01.

The agents energy decays comparatively slowly, but redundant actions are quite expensive; recall, action costs are *subtracted* from the energy level.

To obtain trends, the parameter space for the reinforcement learning component was explored exhaustively. Each trial consisted of setting η and γ at varying levels and testing in the environment for 8000 iterations. The parameters were set at a value between 0.15 and 0.95, at increments of 0.10, resulting in 9 values for both η and γ which combined to give 81 combinations. Each combination of parameters was repeated five times (totalling 405 experiments), and the results presented are averages of these five trials.

During each trial, the mean drive level (arousal) and mean punishments over each cycle (the period of time between refreshes) was measured. This was to enable interpretation of the per-cycle performance. The regularity of environmental dynamics were expected to coerce the agent into producing a routine activity, and then progressively refining this over subsequent cycles. For each parameter combination, there are 5 'profiles', giving the per cycle performances which are averaged to obtain an overall mean per cycle performance for each parameter combination. While providing useful profiling data, a further summary is produced to enable an overall assessment of performance for further investigations. Taking each of the 81 averaged profiles (one for each parameter combination) the overall mean drive/arousal and punishments are derived to represent a 'lifetime performance' statistic. This then enables a parameter surface to be explored. The 'lifetime' summary data and tables can be found in (Joyce, 2001a), reproduced in Volume II, therefore only the conclusions are presented here.

8.9 Trends and Experimental Conclusions

As stated, *no* parameter adaptation occured during each of the experiments; η and γ where set to one combination of the values stated above. The raw results, averaged over 5 runs, were analysed for the local network using the Hebbian interpretation of the Q-learning rule, and the MLP using the Bellman residual as the target vector. From the analyses, it was possible to establish the following empirical trends:

- For the MLP networks, the best performance was always with high η, suggesting the agents need to constantly adapt the responses to maintain goals. However, the role of γ seems to be that in very hard environments, this value should be higher than for moderate and easy environments. In latter cases, the variability of γ is high for easy environments and was more localised in the higher range (towards 1) for the moderate environment.

- The local networks marginally outperform the MLP networks, but this is most likely due to the closer-to-exhaustive mapping of the stimuli space, enabling greater flexibility in constructing appropriate maps from stimuli to desirable actions/outcomes.

- the reliance on high η suggests that *all* networks are constantly adjusting, necessarily, the response associated to any given internal state. This suggests that a more appropriate method might be to encourage exploration while reducing learning rates. This can be achieved by raising Boltzmann temperature while both η and γ are lowered in situations which are 'no win'.

- For the local networks, an important trend (opposite to the MLPs) is that the balance between γ and η appears more critical. For MLPs, it is almost always the case that small regions of the parameter space with high performance, suggest that both parameters should be high. However, for local networks, the higher values of γ affect the punishments for approximately similar overall goal attainment (i.e. the overall drive/arousal is similar). This was especially true for the moderate environment where punishment was clearly traded-off against performance as η and γ vary. This suggests that future rewards balanced against learning the current outcome are critical, and that too high γ skews this considerably. For example, in the moderate environment, the mean drive/arousal level varied from around 0.05 to around 0.5 for high constant values of η as γ was varied in the range $[0, 1]$.

While these results are not conclusive, they aided in deciding how to proceed in building a control architecture for agents. The above empirical conclusions are carried forward into this design process, because the theoretical proposition of (Doya, 1999a) remains underconstrained.

For MLPs the trend appears to be to set both parameters high. This suggests that the experience of novel situations is overwriting existing, learned patterns. Whereas for the local network, the relationship between η and γ is more complex, suggesting a requirement to balance the two parameters. The purpose of testing was to see if there is any empirical support for Doya's proposal that it might be possible to better understand learning, in situated agents, via cognitive and neuro-modulatory mechanisms. The following discussion elaborates on the architecture evolved, the connections to Doya's proposal and the other literature reviewed.

8.10 Control Architecture

8.10.1 Novelty

The agent will need to estimate the novelty of the state if it is to interact with the environment in a continuous way, in order to establish routine activities. A common method of converging machine learning tasks is to reduce learning rate parameters in a manner inversely proportional to time, for example a simulated annealing approach; see (Hertz, Krogh and Palmer, 1991).

The category nodes (as for the Gaussian ART) can be easily enumerated, because each one is representative of a 'total' internal state *including* perception. As a result, frequency of internal state selection corresponds to a kind of novelty. The MLP is more difficult since the internal state is coded for by all hidden nodes, rather than one winning node, requiring some associative memory which records the hidden unit signalling vector, and associates this with frequency of occurrence.

Three options (using the local net) were explored:

1. set $\eta \propto 1/t$

2. set $\eta \propto 1 - \frac{f_i}{\sum_{j \neq i} f_j}$ where f_i denotes the frequency of visiting internal state i.

3. set $\eta \propto 1/f_i$

The first option is objectionable as η must in part depend on the agent's relative perception of the novelty of its current situation. A dependence on overall elapsed time t ignores this. However, the annealing effect is desirable since it facilitates convergence. The second option resulted in the situation where if η is set to be proportional to 1 minus the relative frequency, then very quickly, all novelties are suppressed and η rises to unity. The third option proved most useful, because of its simplicity and its maintaining of the dependence on the agent's perceived novelty. (Humphrys, 1997) used a similar method, as did (Sun and Peterson, 1998). This has the effect of forcing convergence very quickly (at least based on internal state novelty). Only if an infrequent state occurs, does the value of η rise. Other sources of influence would be the success of the routine learned, so that once η has decayed, it can be 'reset' if further learning is required. This will occur when the environment is well-known to the agent, but something 'unseen' has changed and the routine behaviour fails. The agent must adapt, but cannot achieve this based on internal state novelty alone. This problem is taken up again in Chapter 9 where the agent-theoretic notion of Breakdown is revisited.

8.10.2 Cognitive Evidence: Novelty Qualities

Novelty assessment requires some justification. From (LeDoux and Fellous, 1995), activity in the locus coeruleus (LC) can be seen as analogous to the "unexpectedness" of a stimulus; see also (Ricart, 1991) and (Churchland, 1986) pp. 115-116. The LC is a small area close to the top of the brainstem (in vertebrates) and innervates other regions including the cerebellum, brainstem and spinal cord, cortical areas and thalamic regions. Its widespread affect suggests a broad regulatory role in cognition, most notably to provide a generalised arousal or vigilant state. Similarly, (Jackson, Marrocco and Posner, 1994) posited a regulatory role in a functional network model of visual-spatial attention. Their work suggests that the widespread use of norepinephrine (NE) as a neurotransmitter in connections from the LC to other regions, might enable further understanding of attention in cognitive processes. The size (radius or width) of receptive fields was grown and shrunk in response to the level of LC activity in Jackson *et al*'s model. This suggests a similar function and regulatory mechanism for vigilance parameters in this thesis and Grossberg's work. It was not seen as suitable to be explored in this thesis, see also (Peterson and Sun, 1998), since it complicated the setting of vigilance parameters. Further work including a fuller implementation would allow an investigation of the tuning of receptive fields.

The 'unexpectedness' of the internal state is coded for by the relative frequency of it being visited. The effects of novelty on learning parameters will be explored later. However, in terms of adjusting the activation dynamics of neurons, the experiments conducted by Usher, Cohen, Servan-Schreiber, Rajkowski and Aston-Jones (1999) are revisited. This work modelled the LC as a collection (pool or field) of massively interconnected leaky integrate-and-fire neurons. Their network simulated decision tasks in the presence of distracting stimuli, which is behaviourally very simple and involves no learning. However, the focus of the study was on behavioural properties of the spiking LC neurons, of which there was 250. Further inspiration from the work of Usher *et al* includes their demonstration of the widespread affect of the LC, as it not only biases the perceptual task, but also the response generated. A single response node is configured so that its activation function gain, e.g. how receptive the neuron is to build activation, is increased. This models the proposed effects of norepinephrine on target neurons. Namely, the tendency for them to be more receptive to inputs, which contributes to heightened activation. The response node is therefore more likely to fire as a consequence of the effects of LC activity. However, the most significant contribution here is as follows:

- LC neurons have phasic activity which correlates with RD activity (through its widespread innervation of many regions of the nervous system). Activity in RD neurons (using 5HT) affect the balance of DA/5HT in the NAC. This suggests that in highly novel situations, the gating effect of the NAC is increased causing *more random* action selection. This is supported by (Devor, 1995) who states that NE activity in the ascending brainstem can be seen as analgesic, causing the animal to 'ignore' pain in the service of exploration. The effect of the LC (and noradrenergic system) on action selection, is therefore indirect via the 5HT pathway into the NAC from the RD.

The neurodynamics of the LC/NE systems are poorly understood, especially with respect to their constraining models for connectionism, which typically focus on functional systems. In conclusion, Usher, Cohen, Servan-Schreiber, Rajkowski and Aston-Jones (1999) state: "It will also be important to determine the relation of the LC-NE neuromodulatory system to others, such as the dopamine system, that are thought to regulate behaviour based on expectations about future events." pp. 553.

8.11 Controlling Learning

The mechanism considered, to this point, is the LC connection to novelty. Consideration now turns to the roles of η and γ.

8.11.1 Global Effects

Doya's proposal that η be connected to the behaviour of the acetylcholinergic system transpired to be very difficult to validate in this work (partly due to a lack of implementation and cognitive/neuroscientific elaboration in Doya's paper). If learning is a function of modulation, of all varieties including first, second and diffuse messengers and pre- and post-synaptic activity, (in accordance to the variants of the Hebbian postulate) then the acetylcholinergic (ACh) basis of this phenomenon is partly justified. According to (Shepherd, 1994) chapter 7, the effects of ACh on neuromuscular junctions is significant, and has been known since the 1950s. ACh has the modulating effect of imbuing the post-synaptic membrane with a charge. This is later transmitted to the soma of the post-synaptic neuron thus depolarising the cell membrane. ACh belongs to a family of transmitters and modulators, which work to effectively depolarise the post-synaptic membrane, and therefore they appear to have a highly significant role in facilitating spike behaviour in neural networks; see (Destexhe, Mainen and Sejnowski, 1995). If they lead to increases in spike (signal behaviour) it is likewise reasonable to assume that this impacts on LTP (long-term potentiation of synapses). Shepherd summarises the role of ACh as a fast, excitatory, neurotransmitter. At the cellular level, the role of different transmitters (and consequently modulators) varies, depending on the function and location of the neuron in the nervous system. Again, one simplified correlation is insufficient to find the necessary neurophysiological evidence to support a model (cf. Doya's assertion that the root of the global learning rate is ACh).

Re-examining the Q-learning update rule (viewed as a modulated Hebbian rule – see Chapter 4) it is apparent that ACh and η serve a similar functional role. The discrete version of Q-learning with the appropriate weights from section 4.4.2 becomes:

$$w_{jk}(t) = w_{jk}(t-1) - \eta w_{jk}(t-1) + \eta r_t + \eta\gamma \max_{k'} w_{j'k'} \tag{8.6}$$

Note that η is distributed over the passive decay term, the joint reward signal r_t and the tentative future reward term $\max_{k'} w_{j'k'}$. Recall also, from Chapter 4, if updates are performed on state transitions, between action and the next perceptual category choice, then the information to evaluate the future reward term becomes available to the agent via the subsequent firing of the next action node for the new state, and so on.

While η affects the whole of the RHS of equation 8.6, the effect of γ is localised to future rewards, implicating serotonergic systems. It can be inferred that the complimentary effect of dopaminergic and serotonergic activity is distributed over the range of values of r_t. It may be inferred that if $r_t < 0$, DA activity is low, and by the principle that these two modulators maintain an equilibrium during 'normal' conditions, then 5HT is high.

In order to test this hypothesis, the global learning rate η was set to be equal to the novelty which implements the correlation described in (Davis, 1992), where the levels of ACh are correlated with NE and activity in the VTA. The agent with varying η as a function of internal state novelty was tested over a range of γ values. Agents with fixed learning parameters, over multiple trials, revealed that the relationship between the local network and *gamma* appeared more complex than for MLPs. Figure 8.8 shows the resulting performances.

However, with self-adjusting η, the value of γ suddenly becomes more significant. Any further adjustments to the agents learned behaviour must become the exclusive 'responsibility' of γ. In effect, the early convergence of η forces γ to be necessarily high so that *some* further

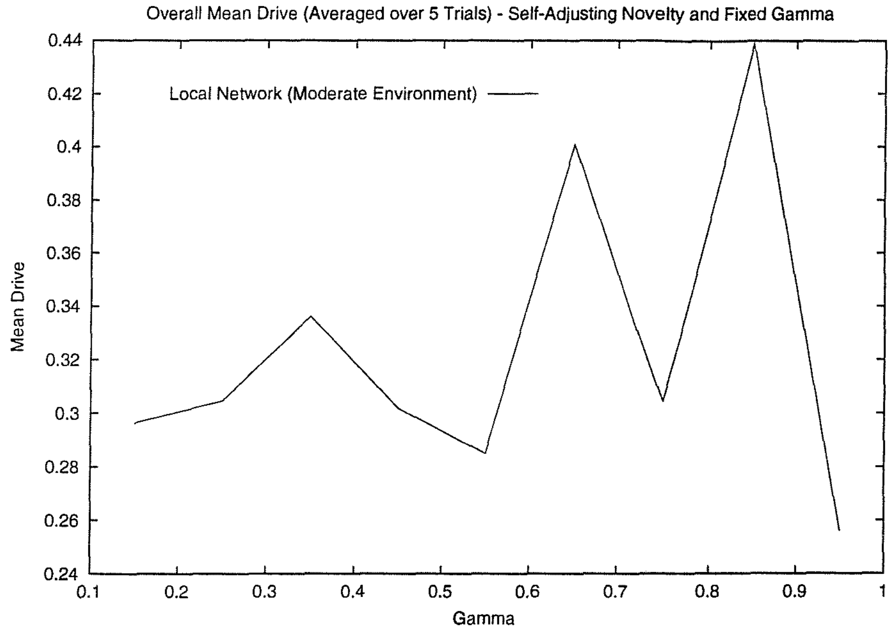

Figure 8.8: Self-Adjusting η and Constant Gamma (Local Network in Moderate Environment)

---

adaptation can occur.

Figure 8.9 shows the novelty detection mechanism incorporated into the schematic of Figure 8.2. As previously stated, the MLP admits no easy method for enumerating internal state, since this is a real valued vector over the hidden units, rather than a binary vector as for the ART / RBF networks used in the local network.

Figure 8.10 shows an individual agent (not averaged over a number of trials) performing in the environment, and the corresponding η, which is set to the current novelty. Notice how after the initial period of the agent constructively mapping the stimuli space, it eventually settles out and η reaches its asymptotic value of close to zero.

From the overall results, Figure 8.8 and 8.10, the cognitive/biological evidence was surveyed and the following points emerge:

- the LC responds *not only* to novelty in the perceived environment, but also to somatic sensing i.e. impinging punishments and homeostasis violations (such as low energy). This is not accounted for in the model of novelty implemented above.

- contrary to the tests of (Joyce, 2001a), γ now appears to be critically high for the agent to succeed.

In essence, the global learning variable η is critically useful in determining the success of the agent, and γ seems to require consistently high values for the agent to succeed. Doya's postulate that the neuromodulatory correlate of γ is 5HT is therefore open, given the kinds of agents modelled here and in the explicit use of reinforcement learning.

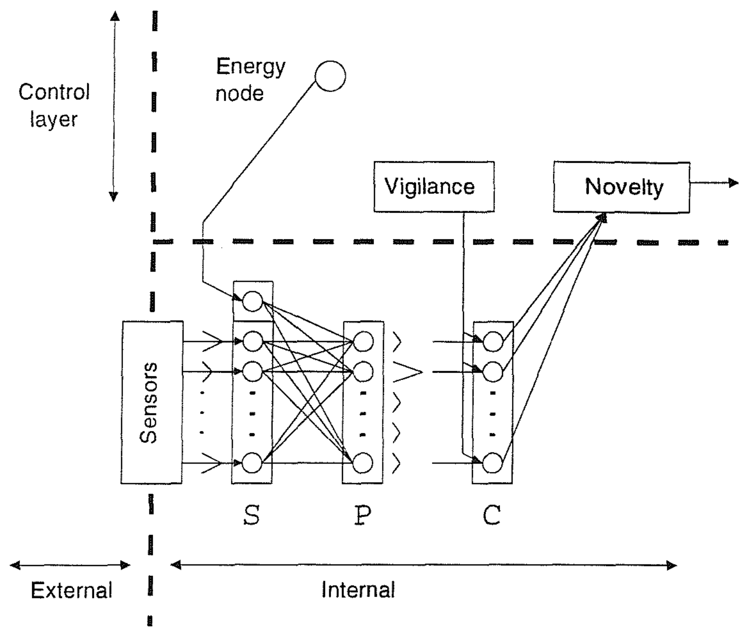

Figure 8.9: Elaboration of Control Layer

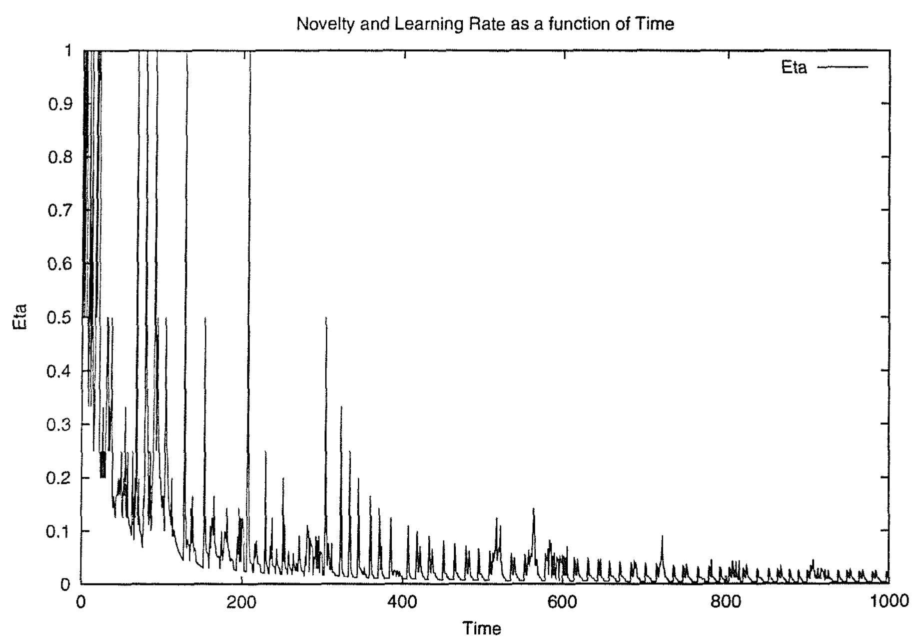

Figure 8.10: Individual Agent Performance with Novelty and $\gamma = 0.95$

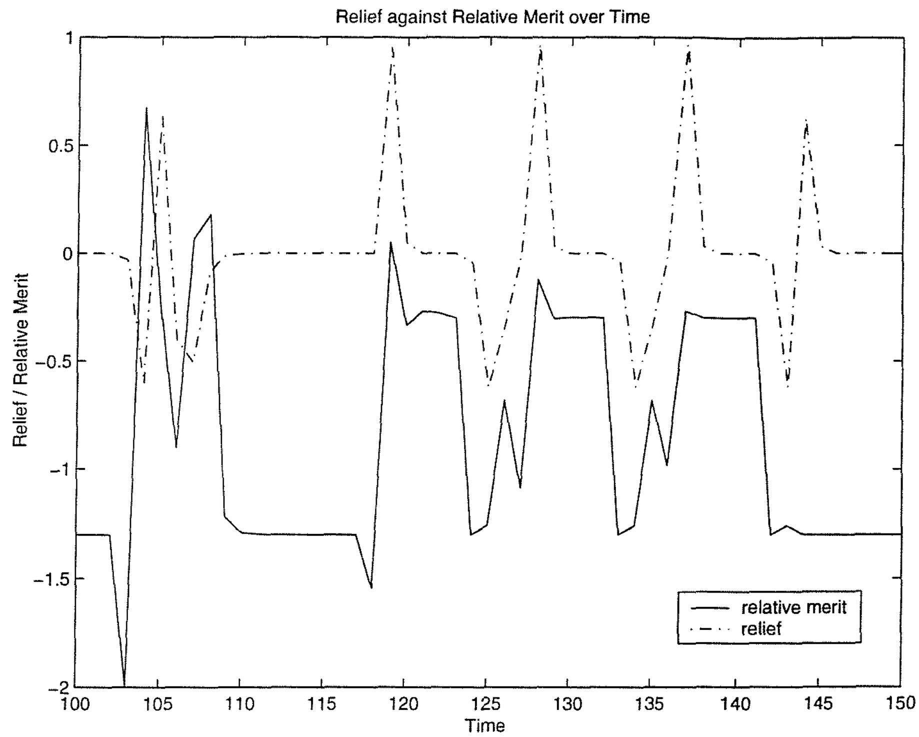

Figure 8.11: Graph showing trends in relief and relative merit signals

8.11.2 Combining Drives, Relative Merit and Reward

Evidence which now needs to be considered, is how 5HT correlates with DA, and how this can be modelled in γ and r_t relationships. Doya proposes that the relative merit is related to DA, but r_t and γ both feature in the relative merit calculation. Empirical evidence suggests that if η varies, then the local network requires a γ value close to unity. In the MLP implementation, both η and γ must be high. Given these results, it is unlikely that a varying γ will help agent performance.

The relationship between relative merit, r_t and drive satiation (rel_G) is more important. Recall that r_t is a conjoint, secondary reinforcement signal. The preferred method of implementing reinforcement effects in Grossberg's work, is the gated dipole network of (Grossberg, 1972a). This proposed a tonic input which competes with the phasic stimulus in an opponent arrangement of interneurons. The tonic input was suggested by Grossberg to be serotonergic in origin. Similarly, if Doya's proposal is valid, then serotonergic analogs in reinforcement models would be constant. Hestenes reports phasic activity due to the complex interactions between the VTA, RD, substantia nigra and NAC. In the light of empirical results, this was not explored further.

In the relative merit calculation, and Schultz *et al*'s model, the temporal difference term $V^{\pi}_{t+1}(y) - V^{\pi}_{t}(x)$ is delivered to the VTA along with the conjoint reward signal r_t. If the final

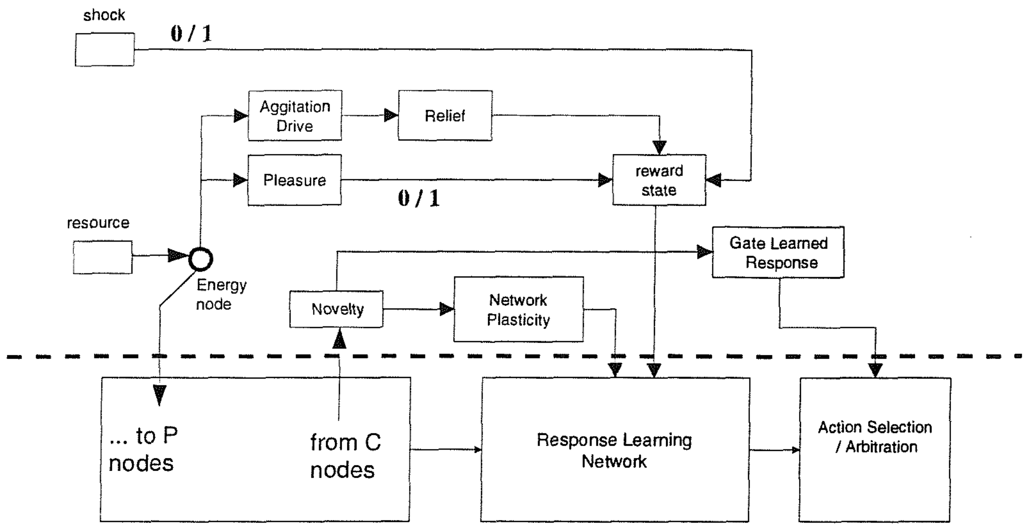

Figure 8.12: The Control Architecture

---

output $delta_t < 0$, then DA activity is low. Intuitively, if γ is close to unity, then a poor state transition from a high value state to a low value state will be heavily coded in the resulting weight update, which will tend to zero (if r_t is low) or be negative (if $r_t \leq 0$). However, the input r_t can be decomposed to reveal a more complex relationship. In the control architecture, rewards are modulated by the satiation of drive. This implies that an agent will not be rewarded if it develops a greedy policy. This is to encourage routines of activity with the environment, as when drive satiation is low, then $r_t \approx 0$, or $r_t < 0$ if a shock is delivered.

As a result, the relief measure is a more direct indicator of the agent's goal completion in an uncertain environment, which separates the contributing primary reinforcers that compose the secondary reinforcing signal r_t. While the two measures follow similar trends, they will differ when the agent is accepting punishment in the service of goal completion (whereas relative merit does not). Figure 8.11 shows an example of the trend, where to expose the correlation and differences between the two behaviours, relief is plotted as $-\frac{dA_G}{dt}$ (to show positive gradients as positive relief instead of negative).

This is implemented as an heuristic critic which combines the drive model of (Chang and Gaudiano, 1998) with RL. Figure 8.12 shows the configuration. The relief signal, indicating satiation of drive (a primary reinforcement), is combined with a 'shock' signal (indicating the arrival or absence of a punishment respectively). The shock signal is an unconditioned stimulus which serves to indicate a negative consequence. This is then used to produce r_t according to the following rule:

$$r_t = (-rel_G(t) \times I_t) - shock \tag{8.7}$$

where I_t is defined as:

$$I_t = \begin{cases} 1 & \text{iff } \frac{dE}{dt} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{8.8}$$

so that if the agent's energy increased, I_t and rel_G combine in such a way as to reinforce the action taken in the corresponding internal state. If the energy level is sufficiently high, implying low agitation/drive, then $rel_G \approx 0$ meaning that only the shock affects the agent, and serves to punish the internal state/action association.

The system described converges with the heuristic critic (Barto, Sutton and Anderson, 1983) and is considered by Grossberg to be analogous to the opponent part of the gated-dipole circuit, which similarly evaluates the contributions of fear/pleasure. In comparing the gated-dipole against methods based on differential Hebbian learning (for example, the temporal difference method extensions of Klopf's work), Grossberg and Levine (1987) note that the level of arousal is an important factor in determining the value (or magnitude of effect) of reinforcement. Similarly, in equation 8.7 above, the relief signal provides a gain for the value of the last resource collected by the agent. If the agent is significantly aroused, then the relief will be high, amplifying the effect of resource intake. Referring to (Chang and Gaudiano, 1998) and (Grossberg, 1972a; Grossberg, 1972b), it can also be seen that equation 8.7 is representative of Grossberg's conflicting drive nodes, which output motivational signals and where the combination is additive competition (see equation 8.7). Further elaboration on this architecture is given in the next chapter when routines are revisited.

8.12 Conclusion

The following relationships were modelled:

- Hullian exploratory drives were implemented as a function of activity, indicating novelty of internal states. That is, Boltzmann temperature is an analog of LC activity indirectly influencing NAC gating behaviour via the VTA/RD connectivity (this is expanded upon in the next chapter when breakdown of activity is considered).
- Noxious drives were generated by the unconditioned stimulus 'shock'.
- Homeostatic drives were modelled as the agitation and relief derived from internal energy.
- Neuromodulatory effects were the production of the reinforcement signal r_t, congruent with (Schultz, Dayan and Montague, 1997), and using (Usher, Cohen, Servan-Schreiber, Rajkowski and Aston-Jones, 1999) an analogue of LC/NE activity was implemented.

This chapter attempted to find suitable implementations of the proposed neuromodulatory basis of 'metalearning'. While some support could be found for the serotonergic / future discount relationship, empirical validation was complex as the MLP and local network required high γ to work well. The dissonance between levels of analysis (e.g. agent implementation with connectionist systems and neuroscientific and behavioural evidence) demonstrates that in part, neuromodulatory effects might best be studied at a temporal and spatial level of abstraction beneath those typical of connectionist architectures. In this case, the application to agent behaviour is premature since it is difficult to see how (for example) the models of (Destexhe, Mainen and Sejnowski, 1995) could be used in network implementations because of the intricate detail of the neuronal model.

The relationship between η and novelty, the locus coeruleus, arousal and drive were described and implemented. Also implemented was the phasic properties of the LC, affecting η. Another property of the LC relationship and learning is when an agent's routine fails. More detail is needed to understand how the action selection temperature might be usefully employed. Therefore, the next chapter returns to the philosophical treatment of agency and shows how the control architecture presented functions in relation to qualitative aspects of behaviour, routines and the dynamics of the environment.

Chapter 9

Routine Activity and Connectionist Agents

This chapter attempts to reconcile the proposed phenomenological agent theory with the empirical work on connectionist networks.

The following will be taken up:

1. models of Breakdown (and those achievable in connectionist models)

2. self-organisation and representation (functional, indexical and deictic representations)

The experiments presented in (Joyce, 2001a) – see Volume II of this thesis – explain quantitative performance as descriptive statistics, but do not enable the exploration of individual agent behaviour. This chapter explores the detailed qualitative aspects of the agent's behaviour.

9.1 Establishing Routines

Recall that for a routine activity to exist, there must be a regularity in the dynamics of the environment, which is both affected by and affects the agent (by the principle of structural coupling). In the simulation used here, the cyclic availability of resources in the environment, the 'danger', 'neutral' and 'safe' levels, are dynamic properties of the environment in that they are things which affect the agent. However, the agent's actions have unintended consequences. That is to say, acting in the environment causes these particular features to impinge on the agent sooner or later. For example, a greedy agent will encounter the consequences of the 'danger' level much quicker than a cautious agent. A particular dynamic (Agre, 1997) is the time-course of these interactions and the regularity established as a result. A routine activity, and an environmental dynamic, are components of a duality; mutually shaping and co-dependent on

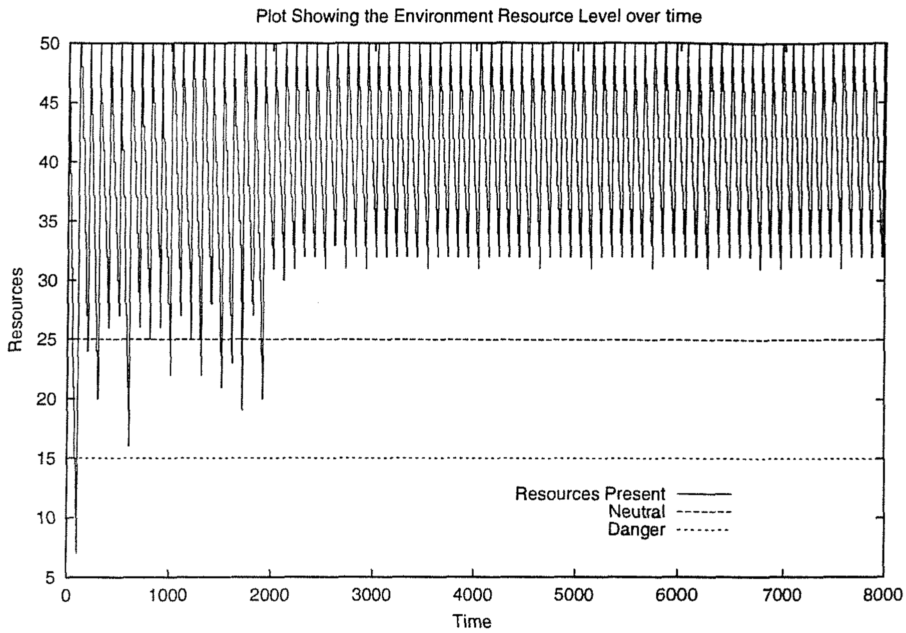

Figure 9.1: Environment Dynamics (Moderate Environment)

each other for existence. If the agent did not act in the environment, the resource level would remain constant, and the indicators (the light and button as equipment to the agent) would likewise remain static. A very different dynamic would occur, and the agent would establish a different routine, if the environment never punished the agent for its actions.

9.1.1 Environment Dynamics

Figure 9.1 shows the environment-side of one such dynamic/routine activity dyad. Note how initially, the agent explores the environment taking actions which sometimes deplete all resources, and eventually, by around 2000 iterations, the agent has decided to only collect resources when the light indicates 'safe'. The agent has, clearly, both influenced its environment and learned to cope within it.

9.1.2 Agent Routines

The complement to the environment-side dynamics, are the agent's goal attainments, which from Chapter 8, was to maintain a low level of drive via collecting resources to satiate the drive. Greedy behaviour was discouraged by the overall reward obtained being modulated by the amount by which drive is satiated. Figures 9.2 and 9.3 show two variables which functionally correlate with the agent's learning, and establishing of routinised activity, in the environment of Figure 9.1. Figure 9.2 shows which internal state node which is most activated at a given time therefore causing the agent to select an action based on the response network weights.

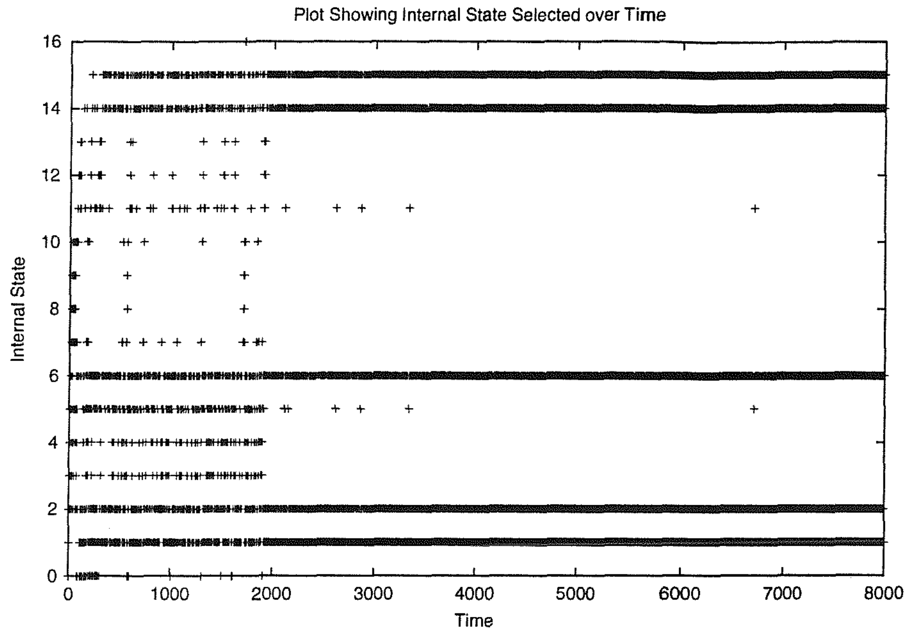

Figure 9.2: Agent Internal State (Moderate Environment)

Figure 9.3 shows the agent's internal energy variable, which affects the level of drive[1]. Notice how the agent settles into a pattern of collecting resources (evidenced by the environment dynamics) which then has a corresponding impact on the sustainability of the internal energy therefor minimising the drive. This in turn established a routine, which causes the agent to oscillate between a selected few internal states.

For a finer grained view of what is happening, we can examine the short-time scale interactions between the agent and the environment. Taking the examples above, it is possible to force the agent to be more conservative by extending the cycle duration. For the moderate environment, extending the refresh cycle to 200 iterations tends to force the agent to display, more overtly, the actions which exploit sequences of micro-routines. That is, we might expect the agent to undertake a short sequence of actions resulting in drive satiation or energy increase, and subsequently settle on NO-OP to preserve energy.

It is simple to conceive of a possible solution. For example, one might switch the button to the 'on' state, and wait until punishment occurs. When punishment does occur, exploration ensues until discovering that the button has state, and switching off results in cessation of punishment. It would then be reasonable to discover that selecting NO-OP results in a comfortable, goal attaining position for some time (since energy decays slowly and hence drive will be 0). However, such greedy behaviour is not available to the agent. It only receives high rewards when punishment is low and drive satiation is high. If the agent receives resources continuously, then the overall reward signal r_t drops to zero quickly, when the agent is not satiating a

[1]In both graphs, values are shown as points to avoid clutter caused by continuous lines

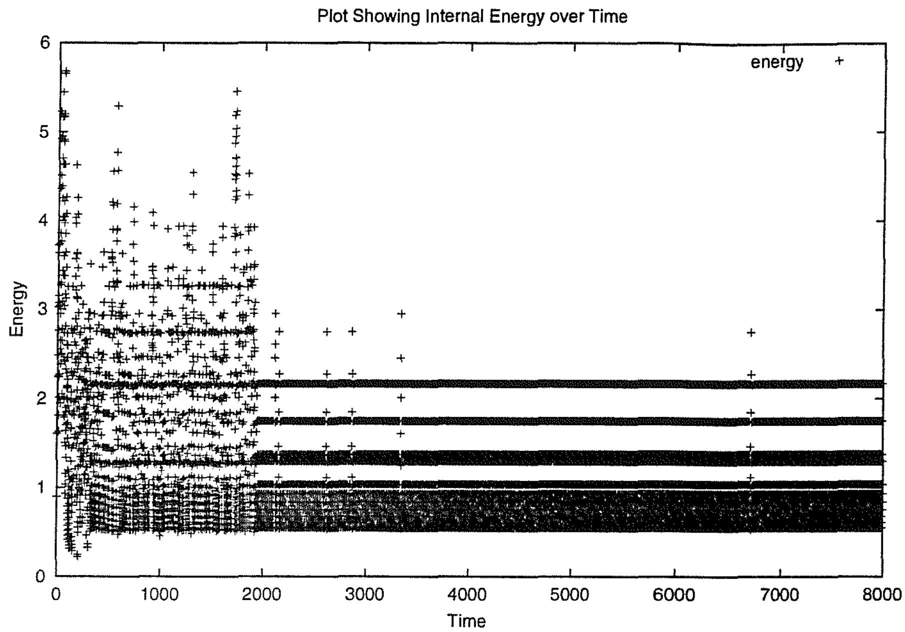

Figure 9.3: Agent Internal Energy (Moderate Environment)

drive.

An interpretation is as follows. Figure 9.4 shows a much shorter time profile for an agent acting in the environment between two refreshes of the resources. Note how the agent at first collects a number of resources, then follows an incremental pattern of fetching resources and then waits (when the resources gradient is 0). To see what is happening more closely, Figure 9.5 shows the same information for the same agent/environment pairing over the interval [850 : 870]. Again, this amplifies short-range dynamics between the agent and environment. Figure 9.6 shows the agent's action sequence over the interval [850 : 870], where lines are shown to emphasise continuity of actions. Viewing the two figures together indicates the strategy the agent is employing. Returning to Figure 9.4, it can be seen that the agent collects resources gradually, but as resources approach the danger level, this routine of collecting and pausing begins to slow. As the agent approaches the next refresh time, it meters out resource-collection activity.

The actual routine for the specific instance of goal attainment and environment shown in Figure 9.6 appears to be:

1. NO-OP ($t = 850 \ldots 855$)

2. PRESS (switch to 'on' state $t = 856$)

3. NO-OP (wait for one iteration at $t = 857$)

4. RELEASE (after collecting two resources, repeatedly switch off at $t = 858 \ldots 861$, which

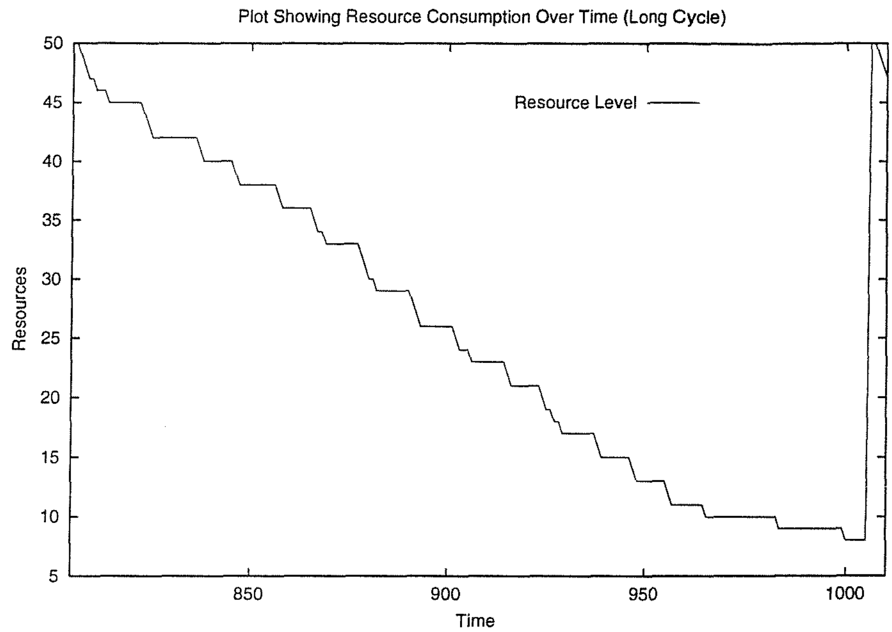

Figure 9.4: Environment Resources ($t \in [800, 1000]$)

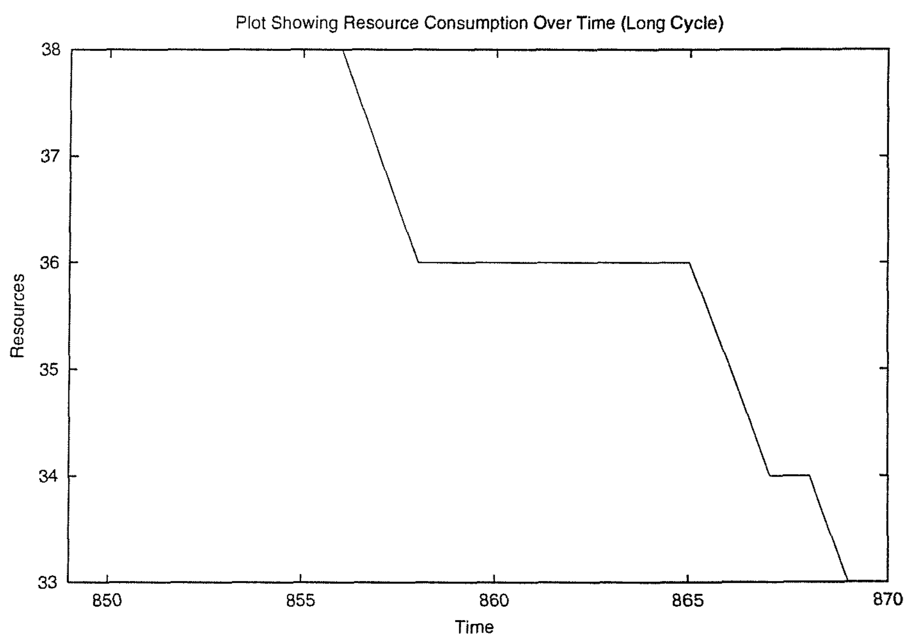

Figure 9.5: Environment Resources ($t \in [850, 870]$)

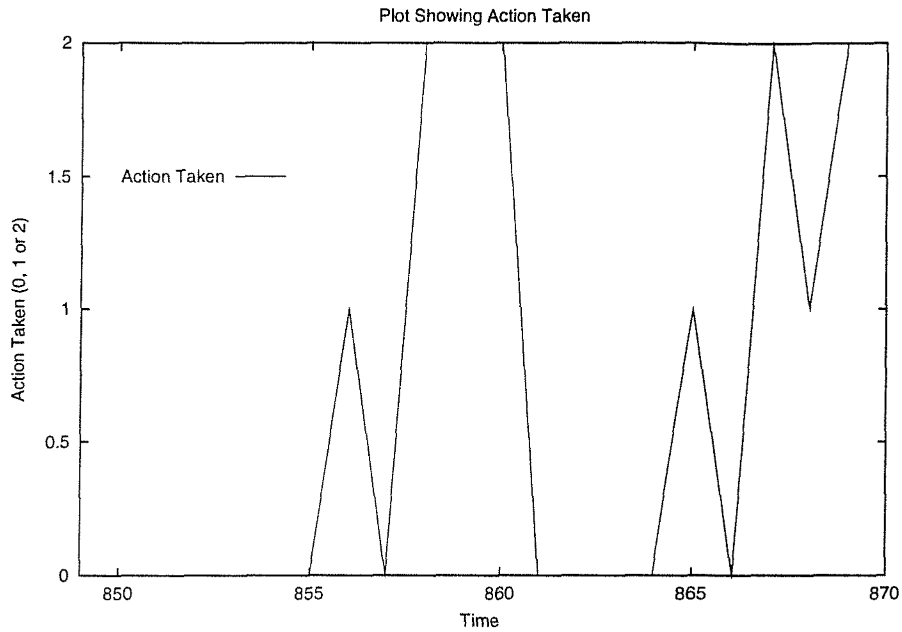

Figure 9.6: Agent Actions ($t \in [850, 870]$) enumerated NO-OP = 0, PRESS = 1 and RELEASE = 2

causes an unnecessary drop in energy through wasted action)

5. NO-OP (now do nothing over the period $t = 861 \ldots 864$)
6. PRESS (switch to 'on' state $t = 865$)
7. NO-OP (wait for one iteration at $t = 866$)
8. RELEASE (at time $t = 867$)
9. PRESS (switch back on at $t = 868$)
10. RELEASE (switch off at $t = 869$)

The agent is favouring a 'PRESS, wait, RELEASE, NO-OP' micro-routine, in order that it maintains goal attainment. The duration of waiting seems to reflect the fact that the agent has learned to 'test' the environment for resources and punishment, by repeating quick sequences which, after the first few cycles in the environment, reveal a strategy for coping.

9.2 Routines and Breakdown

In the foregoing discussion, routines where shown to exist in the local connectionist network-based agent. The behaviour is related to the notions of Heideggarian Breakdown and the establishing of intentional arcs. In the first few iterations of the simulation, the agent has little idea of the function of the equipment present (the button, its state properties and the indicating light). Over time, the agent must acquire skill in manipulating equipment. According to Heidegger, the light is equally a piece of equipment, as its use is defined by context and circumscription. According to (Dreyfus, 1996), such incremental acquisition is the establishing of intentional arcs. This theoretical construct describes the gradual acquisition of a skill (a routine) in dealing with an environment. In effect, the local network described implements an intentional arc (note that Merleau-Ponty gives the *descriptive* name intentional arc to the concept, not its realisation). It does this by gradually refining (through the adaptive mechanisms) the correct comportment, or 'dealing with' the equipment in a way that has utility to the agent. This might be called the *functional meaning* of the equipment with respect to the environment and the agent (see section 9.4.5 later). It is, at least according to (Searle, 1980), meaningless to argue about content semantics and intentionality for the agents studied here.

Breakdown, it is argued, is the principle of phenomenological agency which enables the forming of intentional arcs. Recall, that Merleau-Ponty's work dates from around the rise of cognitivism, whereas Heidegger's work pre-dates this. For example, a human agent has no requirement for a description of a hammer during circumscription (routine activity). If that equipment fails, this descriptive intentionality emerges as breakdown progresses through the 'levels' described in Chapter 5. Likewise, in artificial agents like those in this thesis, there is no such symbolic, sentential description of the button, its state properties, the light and their functional meanings to the agent. The agent infers this from indications of malfunction (the lowest 'order' of breakdown). Hence, during absorbed circumscription, the button, its state and the indicator are present in comportment (the mode of *dealing-with*) but the agent has no intentional arc with which to skillfully use the equipment.

Circumscription is then, initially, a series of random acts. These random acts cause or lead to breakdown because the agent will not be goal attaining. The agent doesn't receive a resource, receives punishment and so on. Failure to complete the assigned goal (maintain a steady level of sustainable energy) and the reciept of punishments, functionally indicate breakdown or more precisely, malfunction. Note that each of these aforementioned variables are *mutually inclusive* in implementing breakdown. Failure to attain the goal is because of the failing of comportment, that is the assignment of actions to deal with the current situation. The agent then experiences malfunction, and adjusts comportment to re-establish circumscription.

Simultaneously, the agent is (by iterations of manifested malfunction) acquiring skill in the activity. The intentional arc necessary to maintain circumscription is gradually shaped by the learning algorithms implemented in the agent, and the function of the equipment is grasped over a series of trials while the agent learns. Eventually (after around 2000 iterations in the examples above) the agent establishes an intentional arc where functionally, the indicator light and the button have some meaning, with respect to activity. The agent has established a routine activity.

It is certainly true that this is not *unique* to agency implemented via connectionism, but the adaptivity present helps ensure that intentional arcs are formed and coupled with the goal direction implicit in the homeostatic mechanism. It is ensured that these intentional arcs are established by breakdown and re-establishing of circumscription via the adapted comportment. It is proposed that the causal properties of the network, that is the activation function, connectivity and weight matrices, contribute an implementation of intentional arcs. Coupled with goal directed behaviour, the entire mechanism underwrites purposeful routine behaviour (Preston, 1993), viz Sloman's 'control state'.

A similar interpretation can be found for the MAVIS2 agents. If the MAVIS2 system utilises the agents, then (see Chapter 6) as they train, they are establishing similar routines of activity by building classifications. Eventually, when intentional arcs have been established by training, the agent is capable of the routine activity of reporting or suggesting classifications for given queries. Perhaps the main contentious issue with the MAVIS2 agents might be that they are not as situated or truly embedded in their environment as the principle of structural coupling suggests (such as the mutually affective nature of enaction evidenced by the simulations of Chapter 7). It was this that motivated the treatment of learning mechanisms based on reinforcement learning, rather than traditional supervised learning. Despite this critique, the principles of routine activity (it is proposed) still feature strongly in the MAVIS2 agents, if only for the utility of the speed of pattern matching against exhaustive searches of the multimedia database. A similar critique has been given by (Bersini, 1997) under the slogan 'breakdown' in connectionist implementations of plans, implemented using Q-learning, but he does not provide a fully worked out description of how this might be observed, or the implementation issues.

9.3 Breakdown to Deliberate Coping

If the agent exists in a stationary environment, then the preceding discussion describes the progress of Breakdown from circumscription to malfunction. However, if the environment's dynamics shift then the agent can no longer rely on its routine and malfunction to correct errors in circumscription. The agent must resort to higher levels of Breakdown such as *deliberate coping* and *deliberation*. See Chapter 5 for a description.

9.3.1 Non-stationary Resource Levels

As an example, consider Figure 9.7. At $t = 4000$ the environment dynamics shift so that the danger level is set at the neutral level. This causes the agent to experience a variety of different internal states. Because the internal state novelty has dropped, and hence $\eta \approx 0$, the agent the agent has encountered this internal state before *and* the learning rate is so low it cannot adapt comportment through malfunction. The agent therefore either fails to achieve its goals, or it suffices.

In Figure 9.7, at $t = 4000$, the agent's internal state transitions change their pattern of

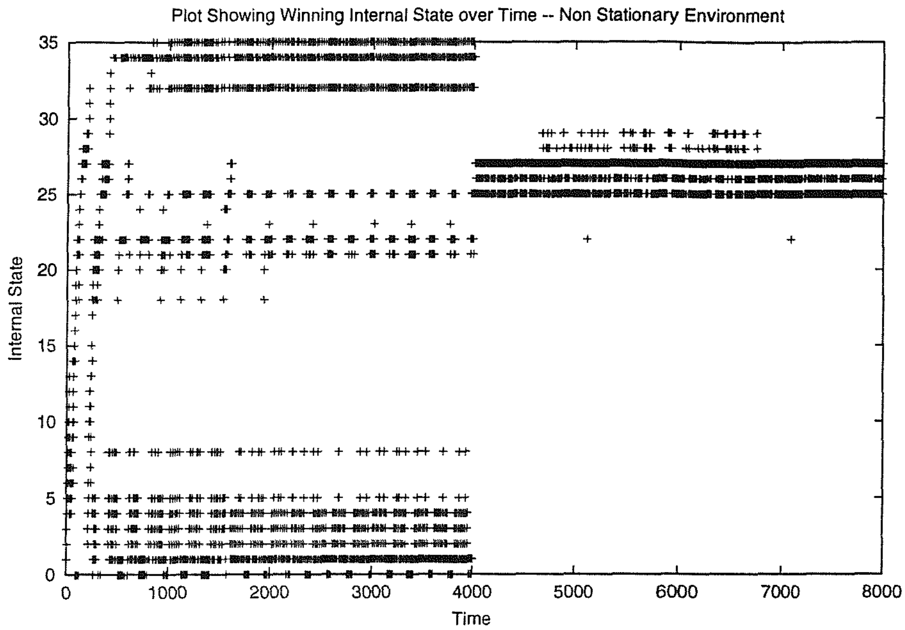

Figure 9.7: Agent Internal State where Environment Dynamics Change at $t = 4000$

activation, and the agent settles into a sufficing routine which attempts to match the dynamics of the environment given current learned behaviour. This example represents a case where the routine fails, but another routine is established in the absence of learning new intentional arcs.

9.3.2 Non-stationary Equipment/Environment Relationship

A more extreme form of non-stationarity, is when the environment and equipment begin to behave differently. As a further test, at $t = 4000$, the behaviour of the equipment with respect to resource delivery changed. The button state is inverted, so that PRESS and RELEASE still toggle the button state, but RELEASE now instigates resource delivery. In effect, the old routine of PRESS, NO-OP, RELEASE will no longer work. The agent must therefore learn the new state behaviour of the equipment/environment dyad, and adjust the routine so that RELEASE is substituted for PRESS and vice versa. NO-OP retains the same properties.

9.3.3 Architectural Considerations for Deliberate Coping

Using the control architecture developed, the inverting of equipment properties resulted in far more catastrophic breakdown. The agent simply fixed on a *single* internal state, and took nothing but NO-OP actions. This is breakdown to the level of deliberate coping, where a new intentional arc and routine must be established. The learning rate has decayed to zero, and no adaptation can occur. In autopoietic terms, this represents structural coupling decay, so that the class identity of the agent would change, as it no longer attained its goals, and its survival is compromised. At this juncture, the notion of the locus coeruleus, as respondent to both internal state novelty as well as somato-sensory information, needs expansion. The decaying of η aids intentional arc adaptation and learning. The dynamics of the routine emerge as a function of malfunction and circumscription. Now, the agent must estimate the value of the routine with respect to its 'knowledge' of the environment dynamics and comportment (dealing with the equipment).

What is required is a mechanism to enable adaptive comportment when established routines fail. The following where initially tried:

1. running averages of relief, drive and energy
2. a leaky-integrator model accumulating over only positive satiating relief levels
3. entropy measures of the 'degree of disorder' over the frequency histograms of internal states

The first method produced a barely detectable, insignificant shift over the internal energy, indicating to the agent the failure of its learned routine to maintain goal achievements. The slow convergence of the running average to the actual mean, meant that this was too unreliable. The second method worked by defining a simple real-time neuron, with very slow growth and decay of activation. The reasoning being that such slow rates enabled the agent to estimate the time averaged positive relief. However, the fluctuation of relief was so vast that even the smallest growth/decay rates resulted in a variable which oscillated too quickly over cycles of the environment. Similar to the running averages, the entropy measure of disorder in the internal state selection adapted too slowly, and the resulting trend was difficult to extract.

Finally, it was observed that the variables used to estimate the stationarity of the environment where those that the agent depends upon to establish routines within cycles. What is therefore needed is a measure which works at a much higher time scale, and evaluates the micro-routines performance per cycle.

Clearly, to simply tell the agent the cycle timing would tie it to one particular environment dynamic. Instead, a kind of deliberate coping strategy was devised which acknowledges that *some* domain knowledge is required, but kept minimal. Notably, the agent must know of the cyclical nature of its routine engagements with the environment. Such domain knowledge is the function of the application specific layer of the MAVIS2 agent architecture in Chapter 6, suggesting a need for a kind of control layer present in (Verschure and Voegtlin, 1998) (where higher control layers also receive sensory input).

To implement this, the architecture was augmented with the following:

- a device which also received the indicator light state, and computed the gradient of this stimulus. The agent now has *a priori* knowledge that the light changes state, and is in some way indicative of the cyclical nature of the environment

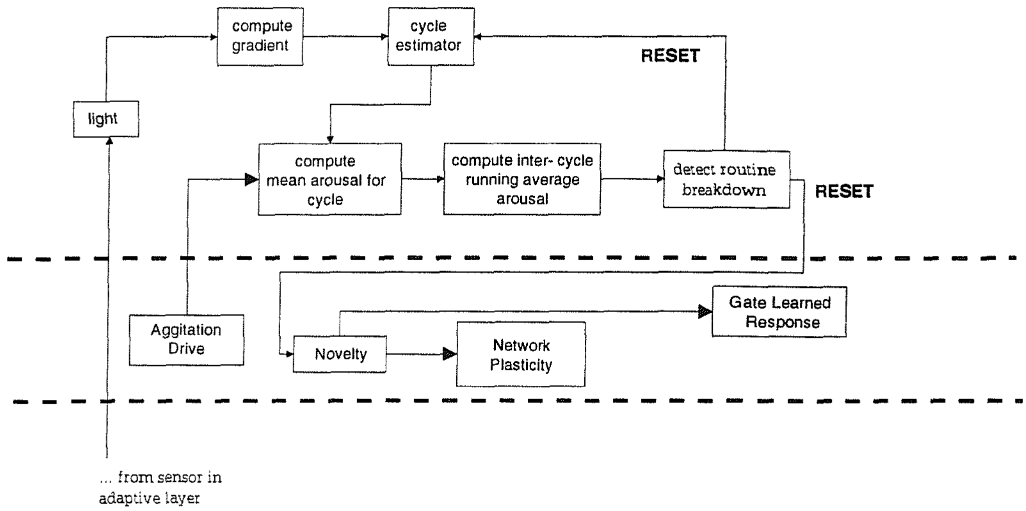

Figure 9.8: Incorporating Breakdown; the third layer (representing increasing specificity and orders of control) uses the control layer to estimate the reliability or routines established.

- a device which attempts to estimate the cycle period (by repeatedly sampling the light stimulus and spotting when this gradient becomes positive, so that if the light switches from indicating 'danger' to 'okay', the agent is alerted and the cycle time estimated
- a device which (given the cycle period estimate) computed the mean of the drive level for a cycle
- a device which computed an inter-cycle running average of the drive level
- a device which detected the failing of a routine. A routine being seen to have failed if the inter-cycle running average rises (e.g. its gradient > 0) and resets the cycle estimate while sending a reset signal to the control layer, which cancels the frequency counts of internal states. This has the effect of resetting the novelty to approximately unity as if the agent was starting over.

This is shown in Figure 9.8. The above mechanism endows the agent with knowledge that the environment has a cyclical dynamic, and that one stimulus modality can be used to estimate this cycle period. However, it is not tied to the specific instantiation of the cyclic dynamics. That is to say, the agent is not given the parameters of the cycle. Once the agent has settled on an estimate, it can measure the mean drive level over the period. It can also, at an even higher time scale, measure a running average of the period drive levels. If this running average begins to rise, then the routine is causing violation of homeostasis suggesting that the currently learned routine is proving stressful. The result is that the agent can recover, and adapt comportment when these internal violations occur.

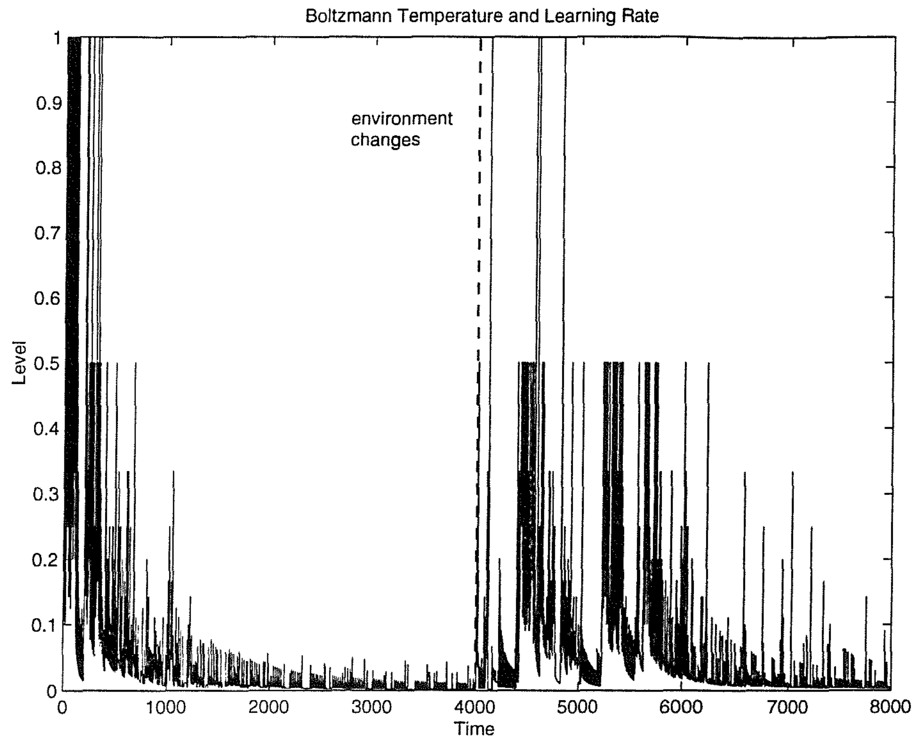

Figure 9.9: Functioning of η and β with the top layer of the agent architecture in place; at $t = 4000$ the functioning of the equipment is inverted

9.3.4 Metalearning Revisited

The actual mechanics of the above layer now merge with the qualitative principles of neuromodulatory effects. It has been previously stated that the NAC, in vertebrates, implements a gating function. Also, a rise in NE levels indirectly causes this gating, while according to (Davis, 1992), raising the base level of ACh in the nervous system.

When the top level of the architecture detects temporary breakdown, the reset signal, to the novelty unit in the control layer, causes resetting of the internal state frequencies. Hence, η rises again and the network is again plastic because the (learned) routine has broken down. Initial testing showed that this enabled a partial, but unreliable, recovery. It would appear that the agent was still 'trapped' in a different routine of internal states. An action gating connection was established, such that the Boltzmann temperature was $\beta = \eta$. This has the effect of perturbing the agent out of any poor internal states routines, while the response network is plastic.

This was tested on the scenario described in section 9.3.2 above. The agent can partially recover from breakdown when resource levels change, but is unable to recover when the equipment changes. The architectural modifications enabled the agent to continue functioning. This is shown in Figure 9.9. Note how breakdown (beyond malfunction to deliberate coping) occurs

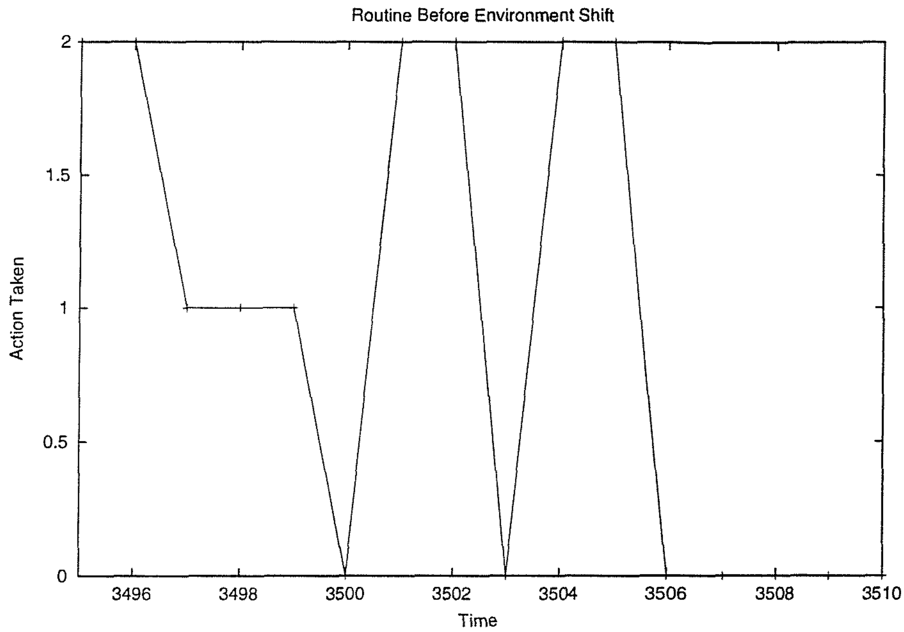

Figure 9.10: Routine Established Before $t = 4000$

shortly after $t = 4000$. The agent then enters a period of exploration and learns to adapt comportment towards the equipment and its changed functionality. The difference in established routines can be seen by examining a routine just before and some time after after $t = 4000$, when the value of η has decayed again.

Figures 9.10 and 9.11 show the routine of actions used before and after the shift in environment dynamics. Note that before the shift, the agent has learned a routine which relies on RELEASE as a non-punishment action, using NO-OP as a means of preserving energy. This is an imperfect routine, but one that suffices to preserve goal attainment. After the shift, comportment is re-established by the agent, as it learns that the roles of PRESS and RELEASE are inverted (i.e. the equipment has changed its functioning). The agent now uses the non-punishing action PRESS in place of RELEASE. An imperfect routine is established, at least from a heuristic perspective, it is easy to see that using a non-punishing action still costs the action/energy penalty. However this imperfect routine suffices given the shifted environment.

The top layer of the architecture uses two variables; the cycle mean arousal and the running average of cycle means. The value of the latter variable is that over time it decreases monotonically, provided that the cycle mean arousal stabilises (indicating a successful routine). If this collapses, then the gradient of the running average will rise. This is shown in Figure 9.12 – note that high η and β from Figure 9.9 correlates with rises in the running average.

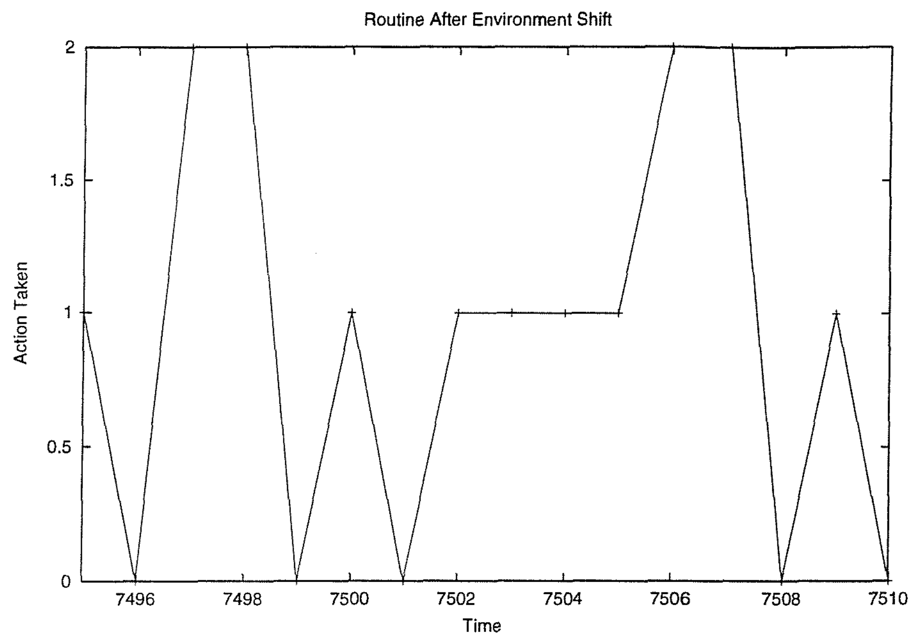

Figure 9.11: Resulting Routine After $t = 4000$

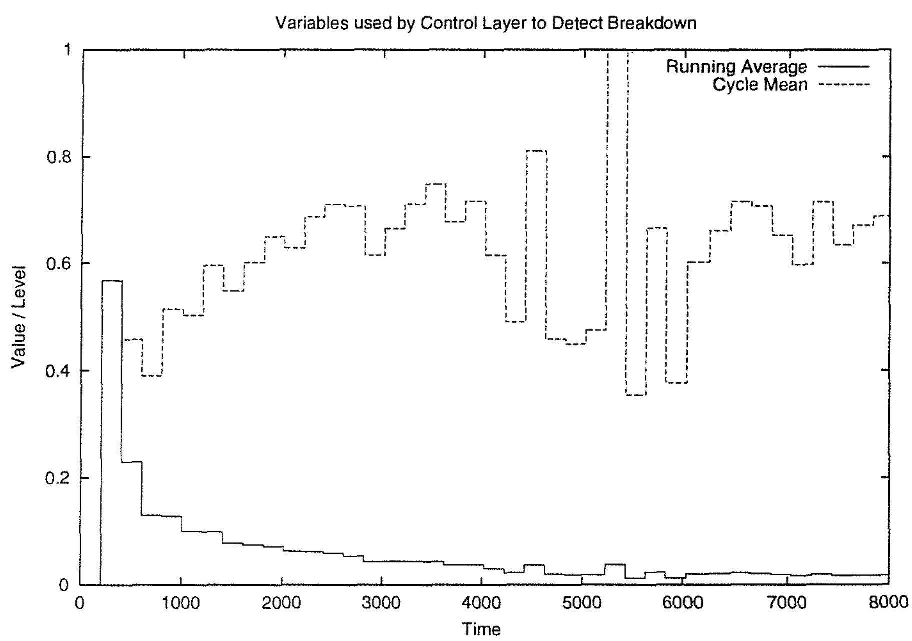

Figure 9.12: Top Layer Variables used to Detect Breakdown

9.3.5 Discussion

The above architectural changes show how a shifting between technical practice, experiments and the proposed theory of agency enable engineering of structurally coupled mechanisms. This enable more sophisticated agent behaviour in a non-stationary environment. The results are indicative, rather than conclusive, but show how the dynamics of learning can correlate with the theory of agency and intentional arcs instantiated in a connectionist substrate which enable comportment and hence, circumscription. The assignment of equipment to the task at hand is adapted during malfunction and deliberate coping; the latter being due to radical changes in the equipment/environment relationship and the agent.

All three layers of the architecture influence each other reciprocally and are simultaneously active. The top layer adds a kind of internal control state which represents the somatosensory component of LC activity in novel and stressful situations. As Merleau-Ponty and Preston observed, the goal directed behaviour emerges from interactions which are not explicit propositional forms. Rather they are transparent to the agent, except in as much as they affect actions and learning. The mechanism (intentional arcs) underpins (or subtends) behaviour and the agent, in accordance with Merleau-Ponty (1962) and McFarland (1996), and attempts to establish and maintain an internal equilibrium or homeostasis.

9.4 Representation and Intentionality

At this stage, it is necessary to unpack the representational theories drawn upon in the preceding chapters. Firstly, the reactive model proposed in (Joyce, 2001b)[2] introduced the notion that heterogenous disjunctions of individual internal states could form the basis of correlational (and thus ascriptional) meaning for internal state. Effectively, purely functional definitions of causal internal state become meaningful when they are correlated with external objects and events (a semiotic principle). Over time, the guarantee of strong epistemic correlations does not hold if the agent is adaptive, so the proposed modification (to accept shifting internal state and corresponding correlation) was referred to as the weak epistemic condition. From Chapter 5, correlational (or ascriptional) theory requires an observer, and hence semiotics was presented, alongside autopoiesis, to explain the necessary and sufficient conditions for internal state to be considered as representations (rather than just functional properties of the system). This is purely derived intentionality.

Secondly, Agre (1997) proposed indexical functional representation. This is internal state which is representation, but with a phenomenological bias. The indexical property is the linguistic principle of devices such as 'now', 'here' and 'I' that are all relative to the agent but which server to individuate temporal and spatial relationships. They are said to be efficient – to say 'here' in different contexts means different *absolute* places or referrents, but the same

[2]Reproduced in Volume II

linguistic token suffices to capture all of these. There is a difficulty in applying Agre's theory here, in that connectionism does not take concatenative syntactic formulae as fundamental to the representational repertoire.

Thirdly, connectionism has its own representational repertoire, and much work has focused on this. There have been many studies of the representational properties of connectionist models e.g. (Smolensky, 1988; Smolensky, 1990; Sharkey and Jackson, 1994; van Gelder, 1990; van Gelder, 1999). Most studies focus on the role of connectionist models in producing systematic phenomena, such as Smolensky's tensor-product filler binding, which demonstrates how connectionist models can encode syntactic structures, of the concatenative variety, and recover that structure from the encoded version. As (van Gelder, 1999) notes, the usual locus of representation is taken as the mapping of input to the representational resources (e.g. hidden nodes in an MLP). Usually, these models focus on explicit variable binding between syntactic structures and the activities of nodes which functionally code the structures. This coding can take a number of forms, see the summary of van Gelder (1999) and the examples of Sun (1992) and Niklasson and Bodén (1997) for discussion. In the agent described here, the localised nature of the basis functions suggests a coarse coding. The network is not encoding syntactic structures, but rather sensory input to the agent. The *P* layer of the network is a composition representing the totality of Schultz *et al*'s cortical modalities which descend to the response learning network (e.g. the striatum and basal ganglia region). The level of analysis is therefore quite different to the kinds of representation traditionally discussed. There is not a concatenatively structured sentence which is presented to the network, where the concatenative order is significant to the interpretation of the input. The presented concatenated vector has no serial ordering that is significant. In terms of dynamics, the absence or presence of a node in the *P* layer reconfigures the possible competition in the *C* layer. That is to say, one node p_i might win at some time, but lose later when a more appropriate node p_j is added or recruited. The chosen internal state is dependent on the *structural* configuration of nodes in the dynamics of competition, and the removing or addition of a node would change this structural relationship cf. (Page, 2000).

Sharkey and Jackson (1994) have proposed that weights form the context-independent features of representation, and activations the context-dependent parts. In their example, from (Smolensky, 1988), symbols such as 'COFFEE' and 'CUP' are used. Each being presented to the network as binary presence/absence indicators. The sentence, albeit a fairly degenerate one, 'COFFEE CUP' is coded $\langle 1, 1 \rangle$. The example can be expanded so that the micro-features of 'COFFEE' and 'CUP' become more complex (e.g. 'brown liquid', 'cylindrical receptacle', 'burnt odour'). The weights are then context independent codings in the service of the desired output (e.g. '1' for full cup of coffee and '0' otherwise). The activations are functions of weights, and so represent the input with respect to the presence of micro-features. The weights, however, do not depend on a specific input, in that they induce activations irrespective of the feature vector input. So, Sharkey and Jackson argue, the activations are context (input) dependent representation induced by the current state of the weights.

It has been proposed that intentional arcs can be implemented in connectionist substrates (Dreyfus, 1996) and this was attempted in the preceding chapters. The implementations presented still require a theory of representation because: i) they are not linguistic and hence peculiar in Agre's indexical-functional theory ii) internal state (if taken as representation) is situated, and hence could benefit from Agre's deictic theory. However, connectionist representation is usually studied from the perspective of structured representations and the recovery of structure from activity over internal or mediating nodes. Smolensky's theory, therefore lacks situatedness, or an explicit account of the deictic properties or representation. Agre's work assumes linguistic indexicality, whereas connectionist agents 'build' representations which are deictic and indexical but have no linguistic or propositional 'value'. It is then a question of sophistication as to whether or not they have the *linguistic* property of indexicality.

The questions a theory must address are therefore:

- semiotic and indexical: in what relation do the representations stand with respect to the agent, environment and observer?
- functionalism and causality: how do they influence action?
- persistence and deicticism: activities on nodes are ephemeral, whereas weights are persistent, what kinds of explanation connect both of these?

9.4.1 Indexicality

Agre focuses on the indexical properties of representation. According to a semiotic analysis (cf. Chapter 5) indexical signs have a multitude of meanings that the linguistic principle of indexicality subsume, and which will be of use in a 'sub-linguistic' representation theory. Linguistic use of indexicals implies something that can pick out similar things on different occasions of use. Agre's example of 'I am here' is (see Table 5.1) a formal indexical, or dicent sign. This is because it is a sentence or proposition which individuates the agent and a location, but not by absolute reference. For example, 'I am here' uttered in a library picks out the agent and location differently to when uttered in a restaurant.

However, no such signs can be found in connectionist agents. This is because with Agre's example, the semiotic devices are of the relational and (mainly) formal trichotomy. However, and this is the utility of Peirce's theory of semiotics, the material trichotomy provides an answer. Take, for example, the indicator light. Before the agent has any experience of this stimulus, it acts as a qualisign – only a property of some object that is incident on the agent; literally a stimulus value or intensity. It does not direct the agent's attention, nor does it have any functional meaning – it is unconditioned. Over time, the agents interactions and learning provide it with a gradual categorisation of the different illumination levels, by categorical features of its perceptual system. At this juncture in learning, the qualisign becomes a sinsign, since the agent's attention is drawn to its categorical nature (crudely, different category nodes are activated by different states of the light – an attentive feature of the agent's perception of the stimulus). Finally, as time progresses and interactions continue, the agent will acquire a 'rule' which *we* might interpret and report as "if the light is in a certain state, punishment ensues if certain actions are taken". By habitual use of the sign, the agent has learned the association between punishment and the stimulus. It has therefore become a legisign.

Of course, the complex stimulus presented to the agent is a combination of many stimuli, but the same principle applies. The key difference between Agre's linguistic notion and connectionist agents is that Agre's formulation assumes *a priori* the ontological detail of the material trichotomy, and debatably, the relational also. Whereas in adaptive connectionist networks, this ontological detail is 'worked out' by interaction with the environment at a level of detail beneath the sentences and propositions of the formal trichotomy. It is here Heidegger reconnects with Peirce. To reach the formal trichotomy, (thematised detached envisioning) there *must* be a substrate of circumscription. Peirce's material trichotomy provides a representational explanation of just that kind of circumscription and the role of signs.

The contrast between this and Smolenksy's approach is now also evident. If a disembodied, unsituated connectionist network is associating structured input to outputs (e.g. classification labels such as 'verb', 'noun' or phonemes from the text string input) then it is mapping between one set of symbols and another. This is because the agent is using the string representation as a codified symbol, absent from its reference, when learning the mapping. The designer (as interpretant) must then be present for any hermeneutic act to take place at the relational level. For example, a user might see the network's output classification 'CAR' and understand its reference. In MAVIS2, it was argued that the multimedia agents where moving from qualisigns to symbols, since the feature vector representation of the media object was more like a qualisign than a symbol (although this is perhaps more controversial).

This is not the case for a situated agent, where there is no classification output which is a symbol. The output is an action which has significance to the agent in respect of the environment. In terms of learning algorithms for connectionist networks, the unsupervised network implemented here is establishing legisigns from qualisigns. A supervised network (e.g. where a designer specifies the desired output) is mapping between symbols (relational-mediation) and qualisigns.

The locus of indexicality is then in the categories built and deployed as sinsigns by the agent's 'perceptual machinery' as these have the familiar indexical property of being directed at something in the world. Malfunction and circumscription shape this movement around the material trichotomy, and the connectionist model implements (functionally) the intentional arc.

9.4.2 The Dynamics of the Material Trichotomy

The above attempts a theory of representation during circumscription, where the agent self-organises what is perceived. A compatible theory was recently proposed by (Smith, 1996). While fundamentally semiotic, it brings elements of the dynamics that Hjemslev added in the planes of content and expression to Saussurian semiology. Smith's notion of 'deictic' is also pertinent to connectionism. Starting with the assumption of two intimately connected 'fields', the 's' region being the agent and 'o' region being the continuous flux of activity in the world, Smith proposes the following:

- that feature detectors (e.g. P layer nodes and their weights centred on regions of the stimuli space) be called *feature participators*. This is because "the edges of their own activity are aligned with, in virtue of being connected to – wired or plugged into or effectively coupled with – the edges of the target entities towards which the systems of which they are part will ultimately be said to be directed" – (Smith, 1996) pp. 218-219. In summary, as the combined input to the P layer nodes shifts in the stimuli space, a previously highly activated node will gradually decrease in activity as the shift occurs, until another node is more highly activated.

- a description of perception as a continuous phenomenon of "effective coupling" between two regions; the 'o-region' (the world) and the 's-region' (the agent's feature participators). Neither of these are classifications, but are mutual components of a duality, which is realised by the feature participators (the necessity for feature participators being one condition for robotic capacity as Harnad (1990) proposed)

- that continuous coupling establishes deictic *register* – the epistemic question of what an agent is 'perceiving' is obvious when the connection between 's' and 'o' regions is spatially and temporally continuous (that is, an object is continually present in an agents field of vision). For example, the agent's internal state will be directed toward (cf. intentionality) the light in a particular state when a corresponding P node is highly activated over a number of iterations.

- for cognition to be successful, the agent must be able to cope with *separation* – when the spatial and temporal coupling between the 's' and 'o' regions is lost (when the light changes state, an object disappears from view). The agent must 'remember', or form a category, to enable non-deictic registration that can be 'reinvoked' when needed in the future. In short, the movement from qualisign to sinsign *requires* that the 's' and 'o' region part so deictic registrater is made impossible. The agents presented here use category nodes to achieve this, which entails learning and adaptivity. Hence, the indexical phenomenon of Peirce's semiotics recurs because of Smith's requirement for separation and the consequences for the agent's perceptual machinery.

- *registration*, being the name given to the non-deictic, uncoupled representations formed in the service of cognitive activity. Registration is asymmetric, according to Smith (1996), because the 's' region (the agent) must maintain the registration, rather than the 'o' region.

Smith's theory is certainly semiotic, but it explains the dynamics in a way congruent with the apparent continuity of cognition that (under his theory) only becomes discretised by separation and the utility of indexicality of the 's' region with the 'o' region. It is a semiotic *and dynamic* theory of structural coupling. Recall that the Saussurian model can be augmented to include plane of content/expression shifts, but Peirce's model has no such expansion. For discussion of the dynamics of Saussurian semiology, see the catastrophe-theoretic metaphor given by (Andersen, 1994) which is similar to Smith's notion of separation.

Linguistic indexicality does not explain the movement between categories in the material trichotomy, which can be seen as the primary sensory underpinning of complex signs such as linguistic tokenings. It is proposed that agents which self-organise functional representations are manipulating the material properties, and the relational properties are parasitic upon these. Hence, connectionism operates, principally, at the material level.

Returning to the proposal of Sharkey and Jackson (1994), the weight and activation roles are not in contention, neither is Smolenksy's tensor product representations. However, the semiotic perspective offers an explanation of what goes on in routine *embedded* activity (with respect to sensori-motor loops and intentional arcs) in an unambiguous way. The features of the coffee cup example in Smolensky's work is (at the very least) parasitic on some robotic capacity which underpins this, and it has been proposed here that the dynamics of the material trichotomy explain this.

An agent experiences raw sensory and internal stimuli. Together they form a 'total' sensory experience. For example, the indicator light (before learning assigns a functional role to this stimulus) is paired with the button state and the internal variable representing goal attainment (energy). This can be understood using in the biological semiotic notion of interpretant (due to von Uexküll) by aligning the interpretant with the agent's disposition (homeostatic state) to interpret signs accordingly.

9.4.3 Deictic Codes and Connectionism

Haugeland (1991) suggests that the notion of what a representation is depends on what is being represented, and the level of analysis. It might therefore be asked if models such as Smolenksy's hold explanatory power, when the recovery of input structure from the representation is not necessary. Smolensky's tensor product theory makes no reference to the source of input, the functional relationships and also does not discuss situatedness. That is to say, Smolenksy's approach would not distinguish between issues of reference as discussed in the semiotic theory of Chapter 5, and in the models of, say, linguistic ambiguity tasks (Rumelhart

Abstraction	Temporal Scale	Primitive	Example
Cognitive	2-3 sec	Unit Task	Dialling a phone number
Embodiment	0.3 sec	Physical Act	Eye movement
Attentive	50 msec	Deliberate Act	Noticing a stimulus
Neural	10 msec	Neural Circuit	Lateral Inhibition
Neural	1 msec	Neuron Spike	Basic Signal

Table 9.1: Temporal Abstractions of Ballard, Hayhoe, Pook and Rao (1997)

---

and McLelland, 1986) and the agents presented here.

As this thesis argues, the principle of deictic representation is necessary for embodied, situated agents – even if the embodiment is a peculiar form of 'virtual' world more likely for software agents. However, Agre's treatment is at the level of linguistic indexicality. Semiotics has been proposed as a means of reconnecting phenomenology with representation discourse. The cognitive basis must be explored with respect to connectionism and the *kinds* of deictic codes it implements.

Ballard, Hayhoe, Pook and Rao (1997) have argued, similarly to Allan Newell, that there are temporal abstractions which constrain processing in human cognitive acts. Their specific contribution was a novel level, that of embodiment, which describes (for humans cognitive ability at least) the temporal granularity as one-third of a second. The argument being that the mechanics of the human form in addition to the specifics of the nervous system, combine in such a way that the kinds of codings (read, representations) used are constrained by that form. For example, they site examples of saccadic eye movements. They argue that the speed of an eye movement is mutually constraining and constrained by the cognitive processing in the visual cortex and motor areas. The theory presented here fits comfortably as follows. Table 9.1 summarises Ballard *et al*'s descriptive levels.

It is now possible to relate the levels to those of the agents implemented and the agent theory proposed as follows:

1. qualisign 'processing' : neural (10 msec)
2. sinsign semiosis : attentive (50 msec)
3. legisign : formation of intentional arcs; a component of embodiment
4. internal state, control architecture influence and response selection : remaining components of intentional arc formation; a further component of embodiment

As previously stated, Ballard (1997) proposed that the functioning of chemical reward indicators was at the level of 300 msec. This would suggest that embodiment (of the artificial agent) is critically dependent on the design decisions made with respect to the functional proposals of (Doya, 1999a; Doya, 1999b; Hestenes, 1991) and (Schultz, Dayan and Montague, 1997). Finally, it is necessary to interpret the other major import from Ballard *et al*'s work, that of deictic codes. It is proposed that semiosis in the agent results in the formation of sinsigns, which corresponds to an implementation of separation given by Smith (1996). The activation of a particular internal state acts as a representation of the world and internal agent state at any moment, and is therefore a *marker* in the discourse of (Agre and Chapman, 1987; Agre, 1997) and (Ballard, Hayhoe, Pook and Rao, 1997). A marker is a deictic code for cognitive purposes.

9.4.4 Functionalism and Causality

It is apparent that the activations (dependent on weight values) influence what action is chosen. This is uncontroversial, so the causal properties also establish the representations (both weights and activations) as functional state in that they are present because they are necessary.

9.4.5 Functionalism and Content

What is not so clear is whether there are any content or semantic issues. Both Smith (1996) and Dennett (1987) debunk the notion that intentionality is either derivative or original cf. (Searle, 1980) and especially (Searle, 1999) pp. 94. In this context, it is to say that the micro-featural representation of the input is derived by some agent (the designer) who has original intentionality, and the associated 'mindfulness' to interpret them and give meaning. The hermeneutic dispute, is whether or not an artificial agent has original or derived intentionality as a result. Smith and Dennett both argue that some artificial agents might possess their own original intentionality; although, as Nagel explained, this is unimaginable to other agents. Searle (1980) maintains this is implausible, and further, any intentionality present in an artifact of human production, possesses derivative intentionality at best.

In terms of the semiotic principles here, an autopoietic interpretation would place the original intentionality (e.g. the interpretant in the triadic semiotic model) with the observer and is therefore causally inert for the agent. Any semiotic theory must likewise concur with this, since signification requires ostensive definition (e.g. individuation of the sign with respect to an observer). The semiotic perspective is controversial, however, because it fails to ground the qualisigns, sinsigns and legisigns without referencing the signification and observer. Therefore in the artificial agent, we still have derivative intentionality.

Functional content semantics are one proposal. The semantics of representations are nothing more than the functional role they undertake in the production of behaviour; see for example (Block, 1980a). The 'meaning' is derivative from the context of use (e.g. other functional states). The 'other minds' approach would be to say that the agent has a 'special' intentionality only accessible to itself, and potentially similar to other agents of that kind. Whether this is a vacuous conception of intentionality is not of concern here, but Searle and Harnad (1990) would seem to argue it is.

Finally then, the semiotic perspective (including Peirce's classificatory scheme and Smith's explanatory dynamics of the 's' and "o' regions) offers an explanation of the kinds of deictic and separated representations agents form, and especially with reference to the construction of any deictic ontology, which is self-organised in a connectionist system.

Two points remain, both being controversial:

- agent intentionality : according to the discussion of this thesis, any intentionality (and therefore the content-carrying burden of representation) is one of the following: (i) a vacuous notion since we cannot ever know what the agent's phenomenal world is like or (ii) we adhere and accept the strong AI hypothesis, but see (Searle, 1980) (iii) purely derived intentionality, reducing the value of internal state to merely functional (iv) 'as-if' intentionality, see the work of (Dennett, 1987), which grants original intentionality to an agent if another agent must use the intentional stance.

- phenomenological and deictic ontology: self-organisation and adaptivity provide a means of placing the burden of building deictic representation with the agent and its environment mutually. This point, with the suitable embeddedness properties of reinforcement learning, and the agent's coupling with the environment, is why Dreyfus (1996) argues that intentional arcs might be implemented in a connectionist substrate

Clark (1989) defends functionalism in a specific way. Searle, he claims, only examined one principle of AI, the physical symbol system hypothesis. This is "coarse grained" functionalism and intentionality cannot reside at this level because of its coarseness (deliberately so, since Newell and Simon (1976) demanded systematic interpretability of the symbols and their standing in relation to the world). The causal properties Searle alludes to *might* be available because the "vast structural variation" of connectionist models enable different, finer grained 'micro-functionalist' formal properties to be captured (which are not available to a PSS system). Searle is, according to Clark, wrong to dismiss *any* formal model *just because* it is not the 'real thing' (Clark, 1989) pp. 32. Later, Clark uses folk-psychology as ascription in a semiotic way (see (Clark, 1989) pp. 58) and then argues that the meaning of micro-functional representation is 'speaker dependent' (read, requires an interpretant).

In a similar way, Dretske (1985) proposed the fact that a system can adapt and learn confers upon it certain valid representational properties akin to intentionality. Adopting a completely functional stance, he says the only meaning *is* that which is causally relevant in producing behaviour. Learning enables unique and definite types of internal state to become functionally relevant to behaviour. Both Fodor and (Dennett, 1987) pp. 305 oppose this, arguing that their must be an arbitrary line between 'learning' and 'performance'. Whether Dretske and Clark combined can recover a purely functional view of meaning has yet to be resolved, but the switch between learning and performance (which both Fodor and Dennett object to) is *absent* in an agent employing the principles espoused in this thesis (situatedness under phenomenological agent theory, including breakdown and connectionism). When routines fail, internal state is suitably adjusted, that is, functional meanings shift.

Recently, O'Brien and Opie (1999) described a 'vehicle theory' of phenomenal experience, which posited a further representational role for connectionism. They argue that at any moment, a connectionist network implements a set of all possible unconscious states, by exactly the kind of superpositional weight codes employed in weighted-sum activation rules. Only when a cycle of activity results in stable activations as outputs is there a particular realisation of a conscious state. So, the potential codes (weights) are the unconscious component, and the stable activations the consciously available component of phenomenal experience. They do not, however, explore the relationship between these states and the concept of intentionality, for example, how these stable activations and weights are *directed towards* the world of the agent, unless we are to assume that their theory makes the concept a moot point. Whether or not the logical implication of their thesis – being that if connectionism can explain phenomenal experience, then connectionist machines can have phenomenal experience – is applicable will depend on how they expand the argument with respect to grounding and situated cognition, necessarily requiring recourse to intentionality.

9.5 Conclusion

This chapter has shown how the phenomenological theory of agency proposed in Chapter 5 has been realised in a connectionist-based agent architecture. Notions of breakdown, intentional arcs and routines of activity have been demonstrated and the implications for architecture design described.

Examples of routines were presented from the simulation environment using the local net implementation of the agent. During these analyses, it became clear that certain non-stationary environmental conditions result in a failure to resume circumscription. Adapting comportment requires further recourse to the 'metalearning' idea. The final level of an architecture was then presented and with qualitative results shown.

Finally, a theory of representation was proposed and connected with recent work on deictic representation in both AI and cognitive science. It was suggested that semiotics offers a level of representation beneath the usual level of indexical-functional representations at the linguistic level. Also, Clark's recovery of functionalism was shown to be fundamentally parasitic on semiotic theories. It was argued semiotic theories can explain multiple levels or representation, *including* the indexical-functional theory of (Agre, 1997), while providing more explanatory power than the kinds of description available using only Smolensky's scheme.

Chapter 10

Conclusion and Future Work

10.1 Summary

This thesis attempted to examine the notion of computational agency in the light of recent work in connectionism and behaviour-based agency. By reconstructing the notion of a computational agent from phenomenology and automata-theoretic models, connectionism has been shown to be a plausible substrate for an embodied, situated agent. Recent advances in agent theory (that situatedness and rationality are key factors) were surveyed. Connectionism was shown to contain aspects of internal state (roughly, memory) which is adaptive – allowing for a self-organisation of functional internal state. Behaviour-functional models do not account for this kind of adaptive internal state and require some revision.

The theory of agency presented attempts to move the issues of representation and content to be both a function of the agent and the environment. Representation was treated from an autopoietic and semiotic perspective. The core of the agent theory presented was phenomenological. Utilising the Heideggarian tradition and the more cognitive psychological notions of intentional arcs, the common-place activities which constitute a majority of agent behaviour were argued to be of the form of routines.

Experiments in utilising such an agent theory took two forms. First, steps toward an implementation utilised associative learning to implement reinforcement effects in the MAVIS2 system. The design practice was informed by the theory presented, as well as empirical work on pattern recognition and requirements dictated by the MAVIS2 multimedia integration project.

The MAVIS2 system provides an exemplar, but is too complex to enable exploration of the situated properties. It is also questionable as to whether the connectionist networks are truly engaged in structural coupling, given the absence of sensori-motor loops. In order to explore situatedness, a simple simulation experiment was devised which enabled further experiments and elaboration of an architecture from the theory. Connectionism befits philosophical, biological and cognitive constraints. The agent theory proposed provides the philosophical backdrop where metalearning and connectionism acts as the cognitive and biological constraints.

Connectionism provides a means of implementing intentional arcs, but with qualification. The agent must be embedded, closely coupled to the environment, and not rely on external construction of target output vectors (at least, this is *at least* what (Dreyfus, 1996) proposed). The usual genera of representational theories make little comment on the principles of agency, or 'low level' deictic sensori-motor codes, which underwrite behaviour.

The penultimate chapter explored the representational issues, provided qualitative analysis of kinds of routine behaviours observable in the simulation, and showed how they emerge in interactions with the environment. The theory of representation was elaborated upon and connected to some unresolved controversies.

10.2 Evaluation

The goals of the work (from Chapter 1) are evaluated:

- with respect to determining the implications of connectionism for agent theory, it proved to be advantageous to explore automata theory and phenomenology. Such an approach enabled issues of adaptivity to be dealt with in a suitably abstract framework that spans physical and virtual agency cf. discussion about the merits of vertical *versus* horizontal architectures

- a phenomenological agent theory is neither virtual or physically committed. It also precedes the usual ontology of agent theory which, for example, argues for game, market or logic/deliberative theoretic approaches. For an example, see the comparitive work of (Ygge and Akkermans, 1999) on centralised control theory versus multi-agent market economies. This is achieved at the expense of specificity by exploring what an agent is, without committing to a computational theory first. Hence, automata models provide enough generality while still containing intuitive aspects of a cognitive agent (e.g. internal state and transition dynamics which causally relate input to output)

- while the agent theory presented was general, examples were given which attempted to ground the theory in architectures and experiments. These were ostensibly focused on sensori-motor activity in virtual agents. The semiotic and phenomenological approach (it is believed) is migratable across implementation boundaries, because the strongest requirements are situatedness and multiple levels of activity, beginning at circumscription (loosely the reactive level) and moving to detached thematised envisioning (loosely, deliberative agency).

- no overall conclusion can be drawn about the most appropriate adaptive architectures for such agents. Learning based on associative and reinforcement learning mechanisms were implemented. Local coarse-coded and distributed networks were used. In terms of architectural design, local networks proved to more straight-forward when integrating into a control architecture.

- it is proposed that the issue of supervision in learning has been dealt with, within the confines of the experimental settings. For example, the use of Q-learning derivatives might not scale well to agents with large numbers of actions and may require much more sophisticated control architectures. However, the goal of this thesis was to explore this design space using connectionism and its methodological imports.

While it might appear that the proposal for a phenomenological theory of agency and its subsequent implementation in connectionism appeals to the eliminativist 'neurophilosophy' of (Churchland, 1986), the role of intentional stance has been integrated through semiotic theories.

10.3 Contribution

It is proposed that the contributions of this thesis are as follows:

- a proposal for agent theory from an automata-theoretic approach, but reconstructed using phenomenological method. This extends Agre's Heideggarian approach in two ways; firstly, representation was treated as sub-linguistic (importing semiotic principles) and secondly, an operational, vertical architecture derived from notions of circumscription through breakdown was described and implemented

- a "strong" theory of computational agency – that is to say, one that accounts for individual learning and adaptive behaviour. Other work focuses on the *reproduction* of behaviour over time (e.g. Gidden's structuration theory in the social sciences and autopoiesis in systems science), but devote less emphasis to the *original production* of that behaviour. It is proposed that such a theory is necessary for computational agency to advance beyond metaphor (for example, in simulations) and for social systems of agents to be realised.

- a demonstration that the hermeneutic tradition is useful and compatible with both AI and cognitive scientific perspectives on phenomenology – the substrate for this demonstration was showing Peircean semiotics to be capable of modelling deictic representations and the subsequent co-incidence with the reinforcement learning mechanisms and the formation of functional representations. Key to this is the realisation that self-organising (loosely, learning) systems are capable of a kind of semiosis, that these internal states are functionally useful to the production of behaviour, and that they can be grounded and implemented in connectionist models. Despite this, the actual arrangement, processing and role of these internal states is determined by the dynamics of the agent-environment interactions (cf. the phenomenological method espoused in Chapter 5) albeit within the constraints of the agent's embodiment (e.g. the perceptual and actuator faculties provided by the designer).

- an implementation of agents with connectionist internal mechanisms which respect and preserve the notion of situatedness by 'internalising' learning, instead of relying on explicit teaching vectors.

- The implemented connectionist networks (for the Skinner box analogy simulation) try to unify the cognitive and philosophical constraints of motor behaviour and Heideggarian phenomenology respectively. In this respect, the thesis contributes a methodology and exemplar implementation of a naturalisation of phenomenology effort. A similar contraining methodology is present in the "constrained connectionism" of (Regier, 1996). While Doya's meta-learning hypothesis could not be empirically validated in its entirety (for methodological reasons; the levels of analysis are incongruent), an attempt was made to use the neuropsychological principles of motor behaviour generation.

- the implemented agent architecture and the connectionist models used to provide circumscriptive activity implement a form of rationality akin to the sufficing principle of Dennett (1996). That is to say, in breaking from the tradition of global minimisation of error (e.g. through backpropagation over all weights) the localist network attempts to correct local perception/action pathways which in turn implies the agent is not rational in the sense of optimising the return for each action taken. Rationality, in agents using the theory and architecture developed here, attempt to find stable and sufficient routines of behaviour to maintain internal homeostatic state.

- concrete examples in the form of MAVIS2 agents for multimedia integration, and a simulated environment which enabled exploration of agent/environment dynamics

10.4 Future Work

10.4.1 Theory

The most obvious direction of advancement on the agent theory presented in this thesis, is to consider Heidegger's notion of Background. This includes the totality of cultural and social practices which shape the agents dispositions. Essentially, this thesis focused on agent-centric theory. In social theory, structuralism proposed that the relationships between entities is the important factor in shaping the order of the world. Saussure (1986) is widely held to be the instigator of this approach from his work in synchronic linguistics. However, anthropology and sociology now consider this to be ignorant of the historical and dynamic factors which shape these relations (cf. diachronic analyses) and the contribution of agency is decayed to the point of individuals becoming passive automata. The apparently dichotomous position (agency *versus* structure) is to a certain extent fielded by autopoiesis, which posits both structure and agency as mutually productive and reproductive in observable systems. Such systematic theories can easily be connected to more contemporary formal social theory, most notably (Giddens, 1986) in that Giddens' 'structuration theory' has strong connections with autopoiesis (Mingers, 1996), although sometimes only metaphorically, as a general systematic exposition of the ordering of social systems. However, structuration theory lacks a *strong* theory of agency, meaning that Giddens' construction of the agent is secondary to his formal theory, and cannot account for learning in everyday routine activity. His theory of time-space distantiation roughly corresponds to the autopoietic notion of structural coupling, and the recursive organisation of a system over both time and space, both being affected by and affecting the rules of a system's organisation [1]. However, he elaborates only two principle components of agency, *practical consciousness* and *discursive consciousness* – roughly corresponding to circumscription and detached thematising. So, to further the work presented here as *computational practice*, multi-agent aspects are essential. It would appear most beneficial to consider how the operational organisation of breakdown affects system behaviour in a multi-agent context. It is proposed that it might be most beneficial as a model of recovery for agents collaborating in a multi-agent environment. Such notions of recovery have implications for agent rationality; where breakdown requires something other than pure reactivity. With respect to a general 'agent science', the utility of the agent theory presented must be assessed with respect to a variety of applications in modelling and simulation as well as technological enterprises. Goldspink (2000) has made some progress on this enterprise from an autopoietic stance; his objective ontology dividing the world into two classes of agents; those that are the focus of investigation and those that are not. With regards to operationalising this meta-model, this thesis would contribute a strong model of agency with which to implement Goldspink's 'primary agent population' cf. Giddens' model of practical and discursive consciousness.

The self-organising properties of connectionism, and the utility of deictic ontology require elaboration. While congruent with the philosophical underpinning, the utility as an engineering method has only been partially demonstrated in a pattern recognition context (MAVIS2). Further, a more complete theory of computational semiotics (Andersen, 1997) might explore the role of connectionist models as implementations of catastrophe-theoretic semiosis. For example, the break in deictic register (Smith, 1996) is analogous to a distinct phase shift in Andersen's theory of semiosis. The settling of a network (such as Grossberg's ART) into an energy minima (after the presentation of a stimulus) has strong connections to (Andersen, 1994), where he proposed minima in non-linear continuous dynamical systems as points of

[1] Giddens defines a technical use of the term 'system' which will not be elaborated upon here, since the metaphorical connection is sufficient

semiotic stability cf. (O'Brien and Opie, 1999).

10.4.2 Implementation

The tractability of connectionist agent implementation is questionable – in that the diverse literatures which inform practice in the domain of neural networks imposes significant overheads. A possible revision to this approach would be to take the *principles* of neural networks which connect with agent theory, and implement them using a generalised model. For example, lateral inhibition forces contrast enhancement, decision making, stability and information compression between distinct populations of neurons. Using Arbib's schema theory, these qualitative aspects could be captured as spatio-temporal abstractions, such as a 'compress' schema operating at a certain temporal level (taking inputs and expressing outputs – see (Arbib, 1989) chapter 5 for examples) cf. (Ballard, Hayhoe, Pook and Rao, 1997). The same architectural and dynamic principle of lateral inhibition would have a different schema (i.e. at a different level of abstraction) for decision making, where the multiple schema provide input and the decision schema operate on a larger time frame, outputting decisions based on the lower timescale 'compression' schema. Hestenes (1991) lists seven criteria of neural network and brain theory, including; modularity, co-operative/competitive relationships between modules, active representation (by stable-states of dynamical systems), competitive selection, associative learning, opponent processing and adaptive resonance for generating codings. Schemata built from these principles might enable a functionally identical implementation of the phenomenological agent theory presented, but in a way which enables future engineering experiments that purely connectionist methods might prohibit because of the complexity of the models.

Doya's theory of metalearning requires further work. Notably, it was not possible to validate some of his claims about the reciprocal relationship between serotonergic and dopaminergic systems, that is, regions of the basal ganglia which are topologically segregated but mutually influential. This might be achieved by keeping simulation environments as parsimonious as those presented here, but keeping other agent-model factors constant, for example, ignoring the effects of learning cf. (Usher, Cohen, Servan-Schreiber, Rajkowski and Aston-Jones, 1999). However, it is proposed that at the temporal and spatial level of reinforcement learning, the relationships modelled in this thesis are perhaps all that can be achieved. This may be said to be due to the heterogeneity of neuronal functions over the nervous system, and the multitude of complex and simple synaptic effects which are not yet understood. To progress, it is suggested that (as was attempted here) heuristics are developed by switching between design practice, experimentation, agent theory and biological and cognitive evidence.

Appendix A

Glossary of Terms

This glossary is provided for the terminology arising from the discussions in the literature reviews and the proposals for a theory of agency.

autopoiesis: the recursive reproduction of observable behaviour of a system – a second-order cybernetic systems theory which posits the agent and environment (unity and medium respectively) as mutually affective structurally coupled entities

Background: Heidegger's notion of the sum of all contributing parts (e.g. cognitive, social and physical) that shape or influence the agent's activity in the world

bounded rationality:
an agent's actions are considered rational (see definition below) within hard limits on the computational (i.e. time and space) resources allocated to a decision – for example, the agent must make a decision in some time t, and the rationality of that decision is with respect to a function measuring the optimality of the decision and the time limit t

breakdown: Heiddeger's term for the way the world is explicitly revealed to the agent when circumscription fails – by analogy in AI, the gradual introduction of symbolic descriptions as existing routines no longer apply

circumscription:
the "mode" of everyday activity which makes no recourse to explicit representations of the world – e.g. circumscriptive activity is routine, and does not require explicit planning, deliberation or awareness of the use of equipment

cognitive domain:
a term used to distinguish between the way two structurally different agents perceive phenomena; if an agent A has different embodiment from agent B, then the cognitive domains (loosely, what and how they will perceive phenomena) of each agent are different

conceptualisation:
one of the many possible formal description of an ontology

deictic representation:
a representation that describes a temporal and spatial relationship between objects that is sufficient for the moment it is used – e.g. it may not be true, or useful a moment later

deliberate coping:
when malfunction fails to re-establish circumscriptive (e.g. routine) activity, the agent enters the mode of deliberate coping (temporary breakdown). In this thesis, this mode is characterised by adaptation which results in new behaviours being produced

deliberation: when deliberate coping fails, the agent resorts to "if then" contigencies (e.g. a weak kind of planning behaviour) to find a way to resume circumscription

distinction: the autopoietic notion of identifiying the unity (agent) from the medium (environment) – the proposal being that the cognitive act of distinction serves to articulate the agent/environment dyad, but is causally inert (e.g. the observer who makes the distinction is *interpreting* a system, but the distinction itself has no effect on the behaviour of the system)

embodiment: an agent that is embodied is one where the environment and agent mutually affect each other

equipment: the term Heidegger uses to define objects that participate in routine activity (i.e. the term *equipment* serves to place objects in the context of use, but this does not necessarily imply that all objects are handled by the agent, more that the agent encounters them)

functional-indexical representation:
a kind of representation that possesses the qualities of a) being necessary to enable activity (functional) and b) being able to individuate different objects with the same "token" (indexical). For example, the linguistic notion of indexical is captured by the use of terms such as "you" which when uttered on different occasions means the same thing (a person being addressed), but potentially individuates different objects on those different occasions of use (e.g. the person indicated on different occasions are not the same person).

GOFAI: a perjorative term used to refer to "good old fashioned artificial intelligence" and its tradition of planning, deliberative reasoning and symbolic formalism

hermeneutic: the study of the methodological principles of interpretation (from Mirrian Webster)

intentional arc: Merleau-Ponty's concept which describes the totality of experience, current internal state and context of the agent as underpinning behaviour – similar to the role of Peirce's interpretant in disposing the agent to perceive a sign but with the emphasis on the agent taking action (rather than Peirce's hermeneutic emphasis)

intentionality: when the "state" of something refers to something else – specifically, when mental state is about a certain thing in the world, then intentionality describes this state as *being about* the thing in the world

interactionism: a theory of agency which describes concepts of representation, perception and action as mutually dependent on the environment and the agent

interpretant: the entity which connects the sign and the referrant in Peircean semiotics. In this thesis, this is taken to be the agent and its current disposition (e.g. internal state) to perceive/interpret a sign.

malfunction: when circumscriptive routine activity fails, the agent experience malfunction, and attempts to re-establish circumscriptive activity by short-range adaptations of behaviour (e.g. by lessening the emphasis on a certain action in a routine in order that the undesirable effects which disturbed circumscription do not occur again)

manipulating: the use of equipment in a task

naturalisation: an attempt to ground a particular theory or perspective in terms familiar to the natural sciences (sometimes, to reduce the theory to a proposed set of natural scientific components)

new-AI: often, "nouvelle-AI", meaning approaches to the production of "intelligent" artefacts which breaks from the traditions of GOFAI

ontology: as a noun, an account of what there is – for example the natural sciences hold to a particular set of ontological commitments about what there is in the world. Such an account might be discussed as a taxonomy or in AI terms, as a conceptualisation

phenomenology:
: the study of subjective experience (e.g. those experiences particular to a given agent)

positive understanding:
: a term used by Heidegger analogous to the notion of symbolic theft

primordial understanding:
: the kind of understanding of objects (equipment) derived from pure sensori-motor experience of them – anologous to the cognitive scientific notion of sensori-motor toil (Cangelosi, Greco and Harnad, 2000).

rationality:
: an agent is said to be rational if each decision or action taken optimises its return according to some objective function. Polemically, this is a troubling definition of rationality, since satisfactory actions might be more advantageous (see (Dennett, 1996) for an evolutionary perspective on rationality and optimisation)

referrant:
: the object that a sign refers to in Peircean semiotics

semiotics / semiology:
: a perspective on the relational nature of knowledge and meaning which divides the object (referrent) from its representation / expression (the sign)

sensori-motor toil:
: see primordial understanding and cf. symbolic theft

sign:
: a concept from the hermeneutic tradition of semiotics and semiology which is an entity that subsumes the signifier and the signified (in Saussurian semiology) and is the third component of the object/referrent, interpretant and sign model of Peirce's semiotics

signified:
: the part of the Saussurian sign which is the actual object expressed by a signifier

signifier:
: the part of the Saussurian sign which is something that expresses (loosely, represents) the signified (the actual object); such as an iconic drawing of a car (the signifier) expressing the signified "a car"

situated:
: when an agent is constantly engaged with it's environment, in contrast to (say) an expert system, which is divorced from the environment and context of its deliberations

structural coupling:
: the autopoietic notion that the organisation of an agent and the environment are mutually affective over time – similar to embodiment

sub-symbolic:

Smolensky's term for a representational token which exists below the level of an atomic symbol (such as "apple") and for which any denotation relies on a composition of many other tokens (typically, sub-symbolic representation implies the tokens are activations of artificial neurons in a connectionist network)

symbolic theft:

the kind of knowledge about some object gained without direct sensori-motor experience of the object – for example, the knowledge one might receive through linguistic communication about a hammer, without having ever used or seen one.

References

Abelson, H. and G. J. Sussman (1985). *Structure and Interpretation of Computer Programs.* The MIT Press.

Ackley, D. H. and M. L. Littman (1991a). Generalisation and scaling in reinforcement learning. In *Advances in Neural Information Processing Systems*, Volume 2, pp. 550–557.

Ackley, D. H. and M. L. Littman (1991b). Interactions between learning and evolution. In *Artificial life II*, pp. 487–509. Addison Wesley.

Agre, P. (1995). The soul gained and lost: Artificial intelligence as a philosophical project. *Stanford Humanities Review 4*(2), 1–19.

Agre, P. (1997). *Computation and Human Experience.* Cambridge University Press.

Agre, P. E. (1988). *The Dynamic Structure of Everyday Life.* Ph.D. thesis, Massachusetts Institute of Technology, Deparment of Electrical Engineering and Computer Science.

Agre, P. E. and D. Chapman (1987). Pengi : An implementation of a theory of activity. In *Sixth National Conference on Artificial Intelligence*, pp. 268–272. Morgan Kaufmann:San Mateo, CA.

Albus, J. S. (1991). Outline for a theory of intelligence. *IEEE Transactions on Systems, Man and Cybernetics 21*(3), 473–509.

Alexander, G. E. (1995). Basal ganglia. See Arbib (1995a), pp. 139–144.

Andersen, P. B. (1994). The semiotics of autopoiesis: A catastrophe-theoretic approach. *Cybernetics and Human Knowing 2*(4).

Andersen, P. B. (1997). *A Theory of Computer Semiotics.* Cambridge University Press.

Arbib, M. A. (1989). *The Mataphorical Brain 2: Neural Networks and Beyond.* John Wiley and Son.

Arbib, M. A. (Ed.) (1995a). *The Handbook of Brain Theory and Neural Networks.* The MIT Press, Cambridge, MA.

Arbib, M. A. (1995b). Levels and styles of analysis. See Arbib (1995a), pp. 11–17.

Argyle, M. and A. Kendon (1967). The experimental analysis of social performance. In *Advances in Experimental Social Psychology*. Academic Press.

Arkin, R. C. (1998). *Behaviour-Based Robotics*. The MIT Press.

Balkenius, C. (1995). *Natural Intelligence in Artificial Creatures*. Ph.D. thesis, Lund University, Cognitive Sciences.

Balkenius, C. and J. Morén (2000). A computational model of emotional conditioning in the brain. In *In Jean-Arcady Meyer, Alain Berthoz, Dario Floreano, Herbert L. Roitblat, Stewart W. Wilson, (Eds) From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behaviour (in press)*. Cambridge, Mass., The MIT Press.

Ballard, D. H. (1997). *An Introduction to Natural Computation*. The MIT Press: Cambridge, MA.

Ballard, D. H., M. M. Hayhoe, P. Pook, and R. P. N. Rao (1997). Deictic codes for the embodiment of of cognition. *Behavioural and Brain Sciences 20*, 723–767.

Barto, A. G. and R. S. Sutton (1981). Landmark learning: an illustration of associative search. *Biological Cybernetics 42*, 1–8.

Barto, A. G., R. S. Sutton, and C. W. Anderson (1983). Neuronlike adaptive elements that can solve difficult learning problems. *IEEE Transactions on Systems, Man and Cybernetics 13*(5), 834–846.

Berns, G. S. and T. J. Sejnowski (1996). How the basal ganglia make decisions. In *The Neurobiology of Decision Making*, pp. 101–113. Springer-Verlag: Berlin. Also available from `http://info.pitt.edu/ berns/papers/damasio/wladamas.htm` .

Bersini, H. (1997). Towards a connectionist model of action sequences, active vision and breakdowns. See Dorffner (1997), pp. 39–65.

Bi, G.-q. and M.-m. Poo (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength and postsynaptic cell type. *The Journal of Neuroscience 18*(24), 10464–10472.

Bigus, J. P. and J. Bigus (1998). *Constructing Intelligent Agents with Java*. John Wiley and Sons, Inc., New York.

Billard, A. and G. Hayes (1999). DRAMA, a connectionist architecture for control and learning in autonomous robots. *Adaptive Behaviour 7*(1), 35–63.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition.* Oxford University Press.

Blake, C. and C. Merz (1998). UCI repository of machine learning databases.

Block, N. (1980a). Introduction: What is functionalism? See Block (1980b), pp. 171–184.

Block, N. (Ed.) (1980b). *Readings in philosophy of psychology.* Harvard University Press.

Block, N. (1995). The mind as the software of the brain. In D. Osherson, L. Gleitman, S. Kosslyn, E. Smith, and S. Sternberg (Eds.), *An Invitation to Cognitive Science.* The MIT Press.

Block, N. (1997). Anti-reductionism slaps back. *Philosophical Perspectives 11*, 107–132.

Bolles, R. C. (1967). *Theory of Motivation.* Harper and Row, New York.

Braitenberg, V. (1984). *Vehicles:Experiments in Synthetic Psychology.* The MIT Press, Cambridge, MA.

Bratko, I. (1990). *PROLOG-Programming for Artificial Intelligence.* Addison-Wesley Publishing Company, Inc.

Bratman, M. E. (1987). *Intention, Plans and Practical Reason.* Harvard University Press, Cambridge, MA.

Bridle, J. S. (1990). Probabilistic interpretation of feedforward classification network outputs with relationships to statistical pattern recognition. In F. Fogelman Soulié and J. Herault (Eds.), *Neurocomputing.* Springer Verlag, Berlin.

Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation 2*(1), 14–23.

Brooks, R. A. (1991a). Intelligence without reason. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, Sydney, Australia*, pp. 569–595.

Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence 47*, 139–159.

Brooks, R. A. (1997). From earwigs to humans. *Robotics and Autonomous Systems 20*, 291–304.

Brooks, R. A. (1999). *Cambrian Intelligence.* The MIT Press.

Brown, T. H. and S. Chatterji (1995). Hebbian synaptic plasticity. See Arbib (1995a), pp. 454–458.

Browne, A. (Ed.) (1997). *Neural Network Perspectives on Cognition and Adaptive Robotics.* Institute of Physics Publishing, Bristol.

Bryson, J. (1999). Hierarchy and sequence vs. full parallelism in action selection. In *Proceedings of the Second Workshop on Virtual Agents*, pp. 113–125. ISSN 1467-2154.

Cangelosi, A., A. Greco, and S. Harnad (2000). From robotic toil to symbolic theft: Grounding transfer from entry-level to higher-level categories. *Connection Science 12*(2), 1–20.

Carpenter, G. A. and S. Grossberg (1987). A massively parallel architecture for a self organising neural pattern recognition machine. *Computer Vision, Graphics and Image Processing 37*, 54–115.

Carpenter, G. A., S. Grossberg, and J. Reynolds (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self organising neural network. *Neural Networks 4*, 565–588.

Carpenter, G. A., S. Grossberg, and D. B. Rosen (1991). ART-2A: An adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks 4*, 493–504.

Catania, A. C. (1992). *Learning.* Prentice-Hall International Editions, Inc.

Chang, C. and P. Gaudiano (1998). Application of biological learning theories to mobile robot avoidance and approach behaviours. *Journal of Complex Systems 1*, 79–114.

Channon, A. and R. I. Damper (1997). The artificial evolution of real intelligence by natural selection. In *Proceedings of the Fourth European Conference on Artificial Life.*

Chapman, D. (1987). Planning for conjunctive goals. *Artificial Intelligence 32*(3), 333–377.

Charniak, E. and D. McDermott (1985). *Introduction to Artificial Intelligence.* Addison-Wesley Publishing Company.

Churchland, P. M. (1979). *Scientific Realism and the Plasticity of Mind.* Cambridge University Press.

Churchland, P. M. (1996). Learning and conceptual change: The view from the neurons. See Clark and Millican (1996), pp. 7–43.

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the Mind/Brain.* The MIT Press.

Clancey, W. (1997). *Situated Cognition.* Cambridge University Press.

Clark, A. (1989). *Microcognition.* The MIT Press.

Clark, A. and P. Millican (Eds.) (1996). *Connectionism, Concepts and Folk Psychology.* Oxford University Press.

Clark, A. and C. Thornton (1997). Trading spaces: Computation, representation and the limits of uniformed learning. *Behavioural and Brain Sciences 20*, 57–90.

Cliff, D., I. Harvey, and P. Husbands (1992). Incremental evolution of neural network architectures for adaptive behaviour – CSRP-256. Technical report, School of Cognitive and Computer science, University of Sussex.

Cliff, D. and J. Noble (1997). Knowledge-based vision and simple visual machines. *Philosophical Transactions of the Royal Society: Biological Sciences 352*(1358), 1165–1175.

Cohen, P. R. (1995). *Empirical Methods for Artificial Intelligence.* MIT Press.

Compiani, M. (1996). Remarks on the paradigms of connectionism. See Clark and Millican (1996), pp. 45–66.

Corbit, L. H. and B. W. Balleine (2000). The role of the hippocampus in instrumental conditioning. *Journal of Neuroscience 20*(11), 4233–4239.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems 2*, 303–314.

Dale, J. (1997). *A Mobile Agent Architecture for Distributed Information Management.* Ph.D. thesis, Department of Electronics and Computer Science, University of Southampton.

Damper, R. I. (1997). Parity is not a generalisation problem. reply to Clark. , A. and Thornton, C. Trading Spaces: Computation representation and the limits of uninformed learning. *Behavioural and Brain Sciences 20*, 69.

Damper, R. I., R. L. B. French, and T. W. Scutt (2000). Arbib: an autonomous robot based on inspirations from biology. *Robots and Autonomous Systems 31*(4), 247–274.

Dautenhahn, K. (1997). I could be you – the phenomenological dimension of social understanding. *Cybernetics and Systems Journal 28*(5), 417–453.

Dautenhahn, K. (Ed.) (2000). *Human Cognition and Social Agent Technology.* John Benjamins Publishing Company.

Davis, M. (1992). The role of the amygdala in fear-potentiated startle: implications for animal models of anxiety. *Trends in Pharmacological Science 13*, 35–42.

Dennett, D. C. (1978). *Brainstorms.* Bradford Books.

Dennett, D. C. (1987). *The Intentional Stance.* The MIT Press.

Dennett, D. C. (1996). *Kinds of Minds.* London: Phoenix/Orion Books.

Destexhe, A., Z. F. Mainen, and T. Sejnowski (1995). Synaptic currents, neuromodulation and kinetic models. See Arbib (1995a), pp. 956–959.

Devor, M. (1995). Pain networks. See Arbib (1995a), pp. 698–702.

D'Inverno, M., M. Fisher, A. Lomuscio, M. Luck, M. De Rijke, M. Ryan, and M. Wooldridge (1997). Formalisms for multi-agent systems. *Knowledge Engineering Review 12*(3), 315–321.

Dobie, M., R. Tansley, D. Joyce, M. Weal, P. Lewis, and W. Hall (1999). A flexible architecture for content and concept based multimedia information exploration. In *Proceedings of the Challenge of Information Retrieval*, pp. 1–12. Also available from `www.bib.ecs.soton.ac.uk/`.

Dorffner, G. (Ed.) (1997). *Neural Networks and a New Artificial Intelligence.* International Thomson Computer Press.

Doya, K. (1999a). Metalearning, neuromodulation and emotion. In *Abstracts of the 13th Toyota Conference on Affective Minds*, pp. 46–47. Also available from `http://www.erato.atr.co.jp/ doya/`.

Doya, K. (1999b). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks 12*(7-8), 961–974.

Dretske, F. (1985). Machines and the mental. *Proceedings and Addresses of the APA 59*, 23–33.

Dretske, F. (1995). *Naturalizing the Mind.* The MIT Press.

Dreyfus, H. L. (1991). *Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I.* The MIT Press, Cambridge: MA.

Dreyfus, H. L. (1992). *What Computers Still Can't Do.* The MIT Press, Cambridge, MA.

Dreyfus, H. L. (1996). The current relevance of Merleau-Ponty's phenomenology of embodiment. *Electronic Journal of Analytic Philosophy 4.* `www.phil.indiana.edu/ejap/ejap.html`.

Duch, W. and N. Jankowski (1999). Survey of neural transfer functions. *Neural Computing Surveys 2*, 163–212.

Etzioni, O. (1993). Intelligence without Robots (A reply to Brooks). *Artificial Intelligence Magazine 14*(4), 7–13.

Ewert, J.-P. (1995). Command neurons and command systems. See Arbib (1995a), pp. 215–219.

Ferguson, I. A. (1991). Towards an architecture for adaptive, rational, mobile agents. In E. Werner and Y. Demazeau (Eds.), *Decentralized AI3: Proceedings of the third European Workshop on Modelling Autonomous Agents and Mulit-Agent Worlds (MAAMAW-91)*. Elsevier.

Fickbohm, D. J. and P. S. Katz (2000). Paradoxical actions of the serotonin precursor 5-hydroxytryptophan on the activity of identified serotonergic neurons in a simple motor circuit. *Journal of Neuroscience 20*(4), 1622–1634.

Fikes, R. E. and N. J. Nilsson (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence 5*(2), 189–208.

Fodor, J. (1979). *The Language of Thought.* Harvard University Press.

Fodor, J. (1998). *Concepts.* Oxford University Press.

Fodor, J. A. (1980). Special sciences. See Block (1980b), pp. 97–115.

Fodor, J. A. and Z. W. Pylyshyn (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition 28*, 3–71.

Foley, D. H. (1972). Considerations of sample and feature size. *IEEE Transactions on Information Theory 18*(5), 618–626.

Gelder, B. d. (1996). Modularity and logical cognitivism. See Clark and Millican (1996), pp. 147–168.

Genesereth, M. R. and S. P. Ketchpel (1994). Software agents. *Communications of the ACM 37*(7), 48–53.

Gerstner, W. (1999). Spiking neurons. See Maass and Bishop (1999), pp. 3–53.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Giddens, A. (1986). *The Constitution of Society.* Polity Press.

Goldspink, C. (2000). Modelling social systems as complex: Towards a social simulation meta-model. *Journal of Artificial Societies and Social Simulation 3*(2). `www.soc.surrey.ac.uk/JASSS/3/2/1.html`.

Grossberg, S. (1972a). A neural theory of punishment and avoidance. i: Qualitative theory. *Mathematical Biosciences 15*, 39–67.

Grossberg, S. (1972b). A neural theory of punishment and avoidance. ii: Quantitative theory. *Mathematical Biosciences 15*, 253–285.

Grossberg, S. (1976). On the development of feature detectors in the visual cortex with applications to learning and reaction-diffusion systems. *Biological Cybernetics 21*, 145–149.

Grossberg, S. and D. S. Levine (1987). Neural dynamics of attentionally modulated pavlovian conditioning: Blocking, interstimulus interval, and attentive feedback. *Applied Optics 26*, 5015–5030.

Guarino, N. and P. Giaretta (1995). Ontologies and knowledge bases. In N. J. I. Mars (Ed.), *Towards Very Large Knowledge Bases*. IOS Press, Amsterdam.

Guiraud, P. (1975). *Semiology*. Routledge and Kegan Paul. Translated by George Gross.

Harel, D. (1992). *Algorithmics* (2 ed.). Addison Wesley Publishing Company.

Harnad, S. (Ed.) (1987). *Categorical Perception: The Groundwork of Cognition*. Cambridge University Press.

Harnad, S. (1990). The symbol grounding problem. *Physica D 42*, 335–346.

Harnad, S. (1992). Connecting object to symbol in modelling cognition. In A. Clark and R. Lutz (Eds.), *Connectionism in Context*, pp. 75–90. Springer-Verlag.

Harnad, S., S. J. Hanson, and J. Lubin (1994). Learned categorical perception in neural nets: Implications for symbol grounding. See Honavar and Uhr (1994), pp. 191–206.

Harvey, I. (1992). Untimed and misrepresented – CSRP-245. Technical report, School of Cognitive and Computer science, University of Sussex.

Haugeland, J. (1991). Representational genera. In W. Ramsey and S. P. Stich (Eds.), *Philosophy and Connectionist Theory*, pp. 61–89. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hayes, P. J. (1992). The naive physics manifesto. In M. A. Boden (Ed.), *Philosophy of Artificial Intelligence*, pp. 171–205. Oxford University Press.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall Inc.

Hebb, D. (1949). *The Organisation of Behaviour*. John Wiley and Sons, New York.

Heidegger, M. (1927). *Sein und Zeit*. Harper-Row, New York. Published in Translation as J. Macquarrie and E. Robinson "Being and Time" (1962).

Hertz, J., A. Krogh, and R. Palmer (1991). *Introduction to the Theory of Neural Computing*. Addison-Wesley, Reading, MA.

Hestenes, D. (1991). A neural network theory of manic-depressive illness. See Levine and Leven (1991), pp. 209–257.

Hilgard, E. R. and G. H. Bower (1966). *Theories of Learning (3rd Edition).* Meredith Publishing Company.

Hirata, K., S. Mukherjea, Y. Okamura, W.-S. Li, and Y. Hara (1997). Object-based navigation : An intuitive navigation style for content-oriented integration environment. In *Proceedings of ACM Hypertext 1997*, pp. 75–86.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems.* The MIT Press.

Holland, J. H. (1995). *Hidden Order.* Addison-Wesley Publishing Company.

Honavar, V. and L. Uhr (Eds.) (1994). *Artificial Intelligence and Neural Networks: Steps Toward Principled Integration.* Academic Press.

Hubel, D. H. and T. N. Wiesel (1965). Receptive fields and functional architecture in two nonstriate visual areas(18 and 19) of the cat. *Journal of Neurophysiology 28*, 229–289.

Huhns, M. N. and M. P. Singh (1997). Ontologies for agents. *IEEE Internet Computing*, 81–83.

Huhns, M. N. and M. P. Singh (Eds.) (1998). *Readings in Agents.* Morgan Kaufman Publishers, San Francisco.

Hull, C. L. (1943). *Principles of Behaviour.* Appleton-Centruy-Crofts, New York.

Humphrys, M. (1997). *Action Selection methods using Reinforcement Learning.* Ph.D. thesis, University of Cambridge, Computer Laboratory.

Hutchins, E. (1993). Learning to navigate. In S. Chaiklin and S. Lave (Eds.), *Understanding Practice: Perspectives on Activity and Context*, pp. 35–63. Cambridge University Press.

Hutchins, E. (1995). *Cognition in the Wild.* The MIT Press.

Jackson, S. R., R. Marrocco, and M. I. Posner (1994). Networks of anatomical areas controlling visuospatial attention. *Neural Networks 7*, 925–944.

Jennings, N. R., K. Sycara, and M. Wooldridge (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems 1*, 275–306.

Jordan, M. I. and C. M. Bishop (1996). Neural networks – Centre for Biological and Computational Learning Memo No. 131. Technical report, MIT.

Joyce, D. W. (2001a). Experiments with connectionist models of reinforcement learning for software agents (ECSTR-IAM005). Technical report, IAM Research Group, Department of Electronics and Computer Science, University of Southampton.

Joyce, D. W. (2001b). Modifying situated automata theory for connectionist agents (ECSTR-IAM004). Technical report, IAM Research Group, Department of Electronics and Computer Science, University of Southampton.

Joyce, D. W. and P. H. Lewis (1999a). Connectionist models for learning choice behaviour in reactive software agents. In *Proc. of the IASTED International Conf. on Artificial Intelligence and Soft Computing*, pp. 437–442.

Joyce, D. W. and P. H. Lewis (1999b). Modelling learning for intelligent software agents: A connectionist approach. In *Proceedings of the Second Workshop on Intelligent Virtual Agents*, pp. 143–146.

Joyce, D. W., P. H. Lewis, R. H. Tansley, M. R. Dobie, and W. Hall (2000). Semiotics and agents for integrating and navigating through media representations of concepts. In *Storage and Retrieval for Media Databases 2000, Proceedings of SPIE, Number 3972*, pp. 120–131.

Kaebling, L. P., M. L. Littman, and A. W. Moore (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research 4*, 237–285.

Kaski, S. and T. Kohonen (1994). Winner-take-all networks for physiological models of competitive learning. *Neural Networks 7*(6), 973–984.

Kim, J. (1998). *Philosophy of Mind.* Westview Press, Oxford.

Klopf, A. H. (1988). A neuronal model of classical conditioning. *Psychobiology 16*(2), 85–125.

Kohonen, T. (1997). *Self Organizing Maps* (2 ed.). Springer-Verlag, Berlin, Heidelberg, New York.

Kosko, B. (1992). *Neural Networks and Fuzzy Systems.* Prentice-Hall, Inc. N. J.

Krell, D. F. (1996). *Martin Heidegger: Basic Writings from Being and Time (1927) to The Task Of Thinking (1964).* Routledge, London.

Kripke, S. (1963). Semantical analysis of modal logic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 9*, 67–96.

Landauer, C. and K. L. Bellman (2000). Computational embodiment: Agents as constructed complex systems. See Dautenhahn (2000), pp. 301–322.

Lave, J. (1988). *Cognition in Practice.* Cambridge University Press.

LeDoux, J. E. and J.-M. Fellous (1995). Emotion and computational neuroscience. See Arbib (1995a), pp. 356–359.

Lee, M. D. (1997). The connectionist construction of psychological spaces. *Connection Science 9*(4), 323–351.

Lenat, D. B. (1995). CYC: A largescale investment in knowledge infrastructure. *Communications of the ACM 38*(11), 32–38.

Levine, D. S. (1991). *Introduction to Neural and Cognitive Modelling.* Lawrence-Erlbaum Associates.

Levine, D. S. and S. J. Leven (Eds.) (1991). *Motivation, Emotion, and Goal Direction in Neural Networks.* Lawrence-Erlbaum Associates.

Lewis, P. H., H. C. Davis, M. R. Dobie, and W. Hall (1997). Towards multimedia thesaurus support for media-based navigation. In A. W. M. Smeulders and R. Jain (Eds.), *Image Databases and Multimedia Search*, Volume 8, pp. 111–118. World Scientific.

Liebrand, W. B. G., A. Nowak, and R. Hegselmann (Eds.) (1998). *Computer Modelling of Social Processes.* Sage Publications, London.

Lippmann, R. P. (1987). Introduction to computing with neural nets. *IEEE ASSP Magazine 4*, 4–22.

Luce, R. (1959). *Individual Choice Behaviour : A theoretical analysis.* John Wiley and Sons Inc.

Maass, W. (1999). Computing with spiking neurons. See Maass and Bishop (1999), pp. 55–85.

Maass, W. and C. M. Bishop (Eds.) (1999). *Pulsed Neural Networks.* The MIT Press.

MacLennan, B. J. and G. M. Burghardt (1994). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behaviour 2*(2), 161–188.

Maes, P. (1989). How to do the right thing. *Connection Science 1*(3), 291–323.

Maes, P. (1991). Situated agents can have goals. In P. Maes (Ed.), *Designing Autonomous Agents – Theory and Practice; from Biology to Engineering and Back*, pp. 49–70. The MIT Press.

Marr, D. (1982). *Vision – A Computational investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman and Co, San Francisco.

Maturana, H. R. (1980). Biology of cognition. See Maturana and Varela (1980a), pp. 5–58.

Maturana, H. R. and F. J. Varela (Eds.) (1980a). *Autopoiesis and Cognition. Boston Studies in the philosophy of science.* D Reidel Publishing Company.

Maturana, H. R. and F. J. Varela (1980b). Autopoiesis: The organisation of the living. See Maturana and Varela (1980a), pp. 59–140.

Maturana, H. R. and F. J. Varela (1998). *El Arbol del Cononcimiento. Published in Translation as The Tree of Knowledge.* Shambhala Publications, Inc., Boston, MA.

McCarthy, J. (1978). Ascribing mental qualities to machines. Technical report, Stanford University AI Laboratory.

McCulloch, W. S. and W. Pitts (1943). A logic calculus of the ideas iminent in nervous activity. *Bulletin of Mathematical Biophysics 5*, 115–133.

McFarland, D. J. (1996). Animals as cost-based robots. In M. A. Boden (Ed.), *Philosophy of Artificial Life*, pp. 179–205. Oxford University Press.

Meehl, P. E. and W. Sellars (1956). The concept of emergence. In *In Feigl, H. and Scriven, M. (eds.) Minnesota studies in the Philosophy of Science. Volume One: The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, pp. 239–252. University of Minnesota Press.

Mehta, M. R., M. C. Quirk, and M. A. Wilson (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron 25*, 707–715.

Mendelson, E. (1987). *Introduction to Mathematical Logic* (3 ed.). Wadsworth and Brooks/Cole, California.

Merleau-Ponty, M. (1962). *Phenomenology of Perception (Translated by Colin Smith).* Routledge and Kegan-Paul.

Messer, K. and J. Kittler (1997). Choosing an optimal neural network size to aid a search through a large image database. In *Proceedings of the British Machine Vision Conference*, pp. 235–244.

Meyer, J.-A. (1997). From natural to artificial life: Biomimetic mechanisms in animat designs. *Robotics and Autonomous Systems 22*, 3–21.

Michalski, R. S., J. G. Carbonell, and T. M. Mitchell (1983). *Machine Learning – An artificial intelligence approach.* Tioga Publishing Company, Palo Alto, California.

Mingers, J. (1996). A comparison of Maturana's autopoietic social theory and Giddens' theory of structuration. *Systems Research 13*(4), 469–482.

Minka, T. P. and R. W. Picard (1997). Interactive learning with a society of models. *Pattern Recognition 30*(4), 565–581.

Minsky, M. (1986). *The Society of Mind.* The MIT Press, Cambridge, MA.

Mitchell, M. (1996). *An Introduction to Genetic Algorithms.* The MIT Press, Cambridge, MA.

Moody, J. and C. J. Darken (1989). Fast learning in networks of locally timed processing units. *Neural Computation 1*(2), 281–294.

Moore, A. W. (1994). The parti-game algorithm for variable resolution reinforcement learning in multi-dimensional state-spaces. In *Advances in Neural Information Processing Systems, Volume 6*, pp. 711–718. Morgan Kaufman.

Moravec, H. (1990). The stanford cart and the CMU rover. In *I.J. Cox and G.T. Wifong (eds.), Autonomous Robust Vehicles*, pp. 407–419. Springer Verlag.

Müller, J. P. (1996). *The Design of Intelligent Agents: A Layered Approach (Volume 1177, Lecture Notes in AI.* Springer-Verlag, Berlin.

Müller, J. P., M. Pishel, and M. Thiel (1995). Modelling reactive behaviour in vertically layered agent architectures. In *Wooldridge, M. and Jennings, N. (eds.), Intelligent Agents, Lecture Notes in Artificial Intelligence 890*, pp. 261–276. Springer Verlag.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review LXXXIII*(4), 435–50.

Narendra, K. S. and M. A. L. Thathachar (1974). Learning automata – a survey. *IEEE Transactions on Systems, Man and Cybernetics 4*(4), 323–334.

Nehmzow, U. (1995). Flexible control of mobile robots through autonomous competence acquisition. *Measurement and Control 28*(2). Also available from `www.cs.man.ac.uk/robotics/nehmzow.html`.

Nehmzow, U. and B. McConigle (1994). Achieving rapid adaptations in robots by means of external tuition. In *From Animals to Animats 3: Proceedings of the Third Conference on Simulation of Adaptive Behaviour*, pp. 301–308. MIT Press.

Newell, A. and H. A. Simon (1972). *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.

Newell, A. and H. A. Simon (1976). Computer science as empirical enquiry: symbols and search. *Communications of the Association for Computing Machinery 19*, 113–126.

Nigrin, A. (1994a). *Neural Networks for Pattern Recognition.* The MIT Press.

Nigrin, A. (1994b). Précis of neural networks for pattern recognition. *Psycoloquy 5*(2). Available from `http://www.cogsci.soton.ac.uk/psyc-bin/newpsy?5.2,`.

Niklasson, L. and M. Bodén (1997). Representing structure and structured representations in connectionist networks. See Browne (1997), pp. 20–50.

Nolfi, S., J. L. Elman, and D. Parisi (1994). Learning and evolution in neural networks. *Adaptive Behaviour 3*(1), 5–28.

Nolfi, S. and D. Parisi (1993). Auto teaching networks that develop their own teaching inputs. In *Deneubourg, J. L. and Bersini, H. and Goss, S. and Nicolis, G. and Dagonnier, R (eds) Proceedings of the second European Conference on Artificial Life* .

Nwana, H. S. and D. T. Ndumu (1996). An introduction to agent technology. *BT Technology Journal 14*(4), 55–67.

O'Brien, G. and J. Opie (1999). A connectionist theory of phenomenal experience. *Behavioural and Brain Sciences 22*, 127–196.

Ogden, C. K. and I. A. Richards (1949). *The Meaning of Meaning* (2 ed.). Routledge and Keagan Paul, London.

O'Keefe, J. and J. Dostrovsky (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research 34*, 171–175.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioural and Brain Sciences 23*(4), 443–512.

Peirce, C. S. (1868). On a new list of categories. *Proceedings of the American Academy of Arts and Sciences* (7), 287–298.

Peterson, T. and R. Sun (1998). An RBF network alternative for a hybrid architecture. In *Proceedings of the International Conference on Neural Information Processing (ICONIP-98)*.

Petitot, J., F. J. Varela, B. Pachoud, and J.-M. Roy (1999). *Naturalising Phenomenology.* Standford University Press.

Pfeifer, R. and P. Verschure (1997). Complete autonomous systems: a research strategy for cognitive science. See Dorffner (1997), pp. 67–92.

Preston, B. (1993). Heidegger and AI. *Philosophy and Phenomenological Research 53*(1), 43–69.

Pribram, K. (1991). Familiarity and novelty: The contributions of the limbic forebrain to valuation and the processing of relevance. See Levine and Leven (1991), pp. 338–365.

Quick, T., K. Dautenhahn, C. L. Nehaniv, and G. Roberts (1999). The essence of embodiment: A framework for understanding and exploiting structural coupling between system and environment. In *International Conference on Computing Anticipatory Systems.*

Ramsey, W. (1997). Do connectionist representations earn their keep? *Mind and Language 12*(1), 34–66.

Rao, A. S. and M. P. Georgeff (1991). Modelling rational agents within a BDI-architecture. In *Proceedings of 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pp. 473–484.

Rao, A. S. and M. P. Georgeff (1995). BDI-agents: From theory to practice. In *Proceedings of 1st International Conference on Multi-Agent Systems (ICMAS-95)*, pp. 312–319.

Regier, T. (1996). *The Human Semantic Potential.* The MIT Press.

Remez, R. E. (1987). Neural models of speech perception: A case history. See Harnad (1987), pp. 199–225.

Ricart, R. (1991). Neuromodulatory mechanisms in neural networks and their influence on inter-stimulus interval effects in pavlovian conditioning. See Levine and Leven (1991), pp. 115–165.

Rolls, E. T. (2000). Précis of 'the brain and emotion'. *Behavioural and Brain Sciences 23*, 177–234.

Rosenschein, S. J. and L. P. Kaelbling (1995). A situated view of representation and control. *Artificial Intelligence 73*(1-2), 149–173.

Roy, J.-M. (1999). Saving intentional phenomena: Intentionality, representation, and symbol. See Petitot, Varela, Pachoud and Roy (1999), pp. 111–147.

Roy, J.-M., J. Petitot, B. Pachoud, and F. J. Varela (1999). Beyond the gap: An introduction to naturalising phenomenology. See Petitot, Varela, Pachoud and Roy (1999), pp. 1–89.

Ruck, D. W., S. K. Rogers, M. Kabrisky, M. E. Oxley, and B. W. Suter (1990). The multilayer perceptron as an approximation to a Bayes optimal discriminant function. *IEEE Transactions on Neural Networks 1*(4), 296–298.

Rumelhart, D. E. and J. L. McClelland (1986). *Parallel Distributed Processing* (Volume 1 ed.). The MIT Press.

Rumelhart, D. E. and J. L. McLelland (1986). *Parallel Distributed Processing* (Volume 2 ed.). The MIT Press.

Russell, S. J. and D. Subramanian (1995). Provably bounded-optimal agents. *Journal of Artificial Intelligence Research 2*, 575–609.

Salomon, R. (1997). The evolution of different neuronal control structures for autonomous agents. *Robotics and Autonomous Systems 22*, 199–213.

Santini, S. and R. Jain (1999). Interfaces for emergent semantics in multimedia databases. In *Proceedings of the IS&T / SPIE Conference on Storage and Retrieval for Image and Video Databases*, pp. 167–175.

Saussure, F. d. (1986). *Course in General Linguistics.* Open Court Press. Translation of the original 'Cours de Linguistic Générale', 1972, Editions Payot, Paris.

Scheier, C. and R. Pfeifer (1995). Classification as sensory-motor coordination. In *Proceedings of the 3rd European Conference on Artificial Life*, pp. 656–667.

Scheier, C., R. Pfeifer, and Y. Kunyioshi (1998). Embedded neural networks: Exploiting constraints. *Neural Networks 11*, 1551–1569.

Schmajuk, N. (1997). *Animal Learning and Cognition.* Cambridge University Press, Cambridge.

Schultz, W., P. Dayan, and P. R. Montague (1997). A neural substract of predication and reward. *Science 275*, 1593–1599.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff and J. Lave (Eds.), *Everyday Cognition*, pp. 9–40. Harvard University Press.

Scutt, T. W. and R. I. Damper (1997). Designing a nervous system for an adaptive mobile robot. See Browne (1997), pp. 220–248.

Searle, J. (1980). Minds, brains and programs. *Behavioural and Brain Sciences 3*, 417–457.

Searle, J. (1999). *Mind, Language and Society: Philosophy in the Real World.* Orion Books, London.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In *Mechanisms of Thought Processes*, pp. 511–531. H.M. Stationary Office, U.K.

Shannon, C. and W. Weaver (1949). *The Mathematical Theory of Communication.* University of Illinois Press.

Sharkey, N. E. (1997). Artificial neural networks for coordination and control: The portability of experiential representations. *Robotics and Autonomous Systems 22*, 345–359.

Sharkey, N. E. and J. N. H. Heemskerk (1997). The neural mind and the robot. See Browne (1997), pp. 169–194.

Sharkey, N. E. and S. A. Jackson (1994). Three horns of the representational trilemma. See Honavar and Uhr (1994), pp. 155–188.

Shepard, R. N. (1987). Towards a universal law of generalisation for psychological science. *Science 237*, 1317–1323.

Shepherd, G. M. (1994). *Neurobiology*. Oxford University Press.

Shoham, Y. (1993). Agent oriented programming. *Artificial Intelligence Magazine 60*(1), 51–92.

Simon, H. A. (1947). *Administrative Behaviour*. New York: Macmillan.

Simon, H. A. (1985). *The Sciences of the Artificial* (2nd ed.). The MIT Press.

Skinner, B. F. (1938). *The Behaviour of Organisms*. New york, Appleton.

Sloman, A. (1996). What sort of architecture is required for a human-like agent? In F. Kessel, P. Cole, and D. Johnson (Eds.), *Foundations of Rational Agency*, pp. 27–39. Kluwer Academic Publishers.

Sloman, A. (2000). Architecture requirements for human-like agents both natural and artificial. See Dautenhahn (2000), pp. 163–195.

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review 68*, 141–156.

Smith, B. C. (1996). *On the Origin of Objects*. The MIT Press.

Smith, D. W. (1999). Intentionality naturalised? See Petitot, Varela, Pachoud and Roy (1999), pp. 83–110.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioural and Brain Sciences 11*, 1–74.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence 46*, 159–216.

Smoliar, S. W., J. D. Baker, T. Nakayama, and L. Wilcox (1996). Multimedia search: An authoring perspective. In *Proceedings of Image Databases and Multi-Media Search*, pp. 1–8.

Sowa, J. F. (1984). *Conceptual structures: Information Processing in Mind and Machine*. Addison-Wesley Publishing Company.

Sowa, J. F. (2000). *Knowledge Representation: Logical, Philosophical and Computational Foundations.* Brooks/Cole Publishers, CA.

Spier, E. and D. McFarland (1998). Learning to do without cognition. In *From Animals to Animats 5: The Fifth International Conference on Simulation of Adaptive Behaviour (SAB98)*, pp. 38–47. The MIT Press.

Sun, R. (1992). On variable binding in connectionist networks. *Connection Science 4*(2), 93–124.

Sun, R. and T. Peterson (1998). Autonomous learning of sequential tasks: Experiments and analyses. *IEEE Transactions on Neural Networks 9*(6), 1217–1234.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning 3*(1), 9–44.

Sutton, R. S. and A. G. Barto (1998). *Reinforcement Learning: An Introduction.* MIT Press.

Tansley, R. H. (2000). *The Multimedia Thesaurus: Adding a Semantic Layer to Multimedia Information.* Ph.D. thesis, Department of Electronics and Computer Science, University of Southampton.

Taylor, J. G. (1994). Goals, drives and consciousness. *Neural Networks 7*(6), 1181–1190.

Tesauro, G. (1992). Practical issues in temporal difference learning. *Machine Learning 8*, 257–277.

Tesauro, G. (1995). Temporal difference learning and TD-gammon. *Communications of the ACM 38*(5), 58–68.

Thornton, C. (1996). Why concept learning is a good idea. See Clark and Millican (1996), pp. 181–194.

Thornton, C. (2000). *Truth From Trash: How Learning Makes Sense.* The MIT Press.

Tinbergen, N. (1951). *The Study of Instinct.* Clarendon Press, Oxford.

Tyrell, T. (1993). *Computational Mechanisms for Action Selection.* Ph.D. thesis, University of Edinburgh.

Usher, M., J. D. Cohen, D. Servan-Schreiber, J. Rajkowski, and G. Aston-Jones (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science 283*, 549–554.

van Gelder, T. (1990). Compositionality: A connectionist variation on a classical theme. *Cognitive Science 14*, 355–384.

van Gelder, T. (1999). Distributed versus local representation. In R. Wilson and F. Keil (Eds.), *The MIT Encyclopedia of Cognitive Sciences*, pp. 235–237. The MIT Press.

Varela, F. J., E. Thompson, and E. Rosch (1991). *The Embodied Mind.* The MIT Press.

Verschure, P. F. M. J. and T. Voegtlin (1998). A bottom up approach towards the acquisition and expression of sequential representations applied to a behaving real-world device : Distributed adaptive control III. *Neural Networks 11*, 1531–1549.

Watt, S. N. K. (1996). Artificial societies and psychological agents. *BT Technology Journal 14*(4), 89–97.

Widrow, B. and M. E. Hoff (1960). Adaptive switching circuits. In *IRE WESCON Convention Record*, pp. 96–104.

Williamson, J. R. (1996). Gaussian ARTMAP: A neural network for fast incremental learning of noisy multi-dimensional maps. *Neural Networks 9*(5), 881–897.

Wilson, S. W. (1991). The animat path to AI. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behavior.* MIT Press/Bradford Books.

Winograd, T. (1973). A procedural model of language understanding. In R. Schank and K. Colby (Eds.), *Computer Models of Thought and Language.* Freeman.

Winograd, T. and F. Flores (1986). *Understanding Computers and Cognition.* Addison-Wesley Publishing Company, Reading, MA.

Wooldridge, M. (1999). Intelligent agents. In G. Weiss (Ed.), *Multiagent Systems.* MIT Press.

Wooldridge, M. and N. R. Jennings (1995). Intelligent agents: Theory and practice. *Knowledge Engineering Review 10*(2), 115–152.

Wooldridge, M. and N. R. Jennings (1996). Software agents. *IEE Review*, 17–20.

Yang, J., P. Pai, V. Honavar, and L. Miller (1998). Mobile intelligent agents for document classification and retrieval: A machine learning approach. In *Proceedings of the 14th European Meeting on Cybernetics and Systems Research*, pp. 707–712.

Ygge, F. and H. Akkermans (1999). Decentralised markets versus central control: A comparative study. *Journal of Artificial Intelligence Research 11*, 301–333.

Yuille, A. L. and D. Geiger (1995). Winner-take-all mechanisms. See Arbib (1995a), pp. 1056–1060.

Yuille, A. L. and N. M. Grzywacz (1989). A winner-take-all mechanism based on presynaptic inhibition. *Neural Computation 1*, 334–347.

Ziemke, T. and N. E. Sharkey (2000). A stroll through the worlds of robots and animals: Applying Jakob von Uexküll's theory of meaning to adaptive robots and artifical life. *Semiotica*.