A Narrative Architecture for Functioning Minds: A Social Constructionist Approach

William Clocksin
Computer Laboratory, University of Cambridge
wfc@CL.cam.ac.uk

Abstract

We aim to build a conceptual framework for artificial intelligence (AI) that gives priority to social relationships as a key component of intelligent behaviour. It starts from Weizenbaum's premise that intelligence manifests itself only relative to specific social and cultural contexts. This is in contrast to one prevailing view, which sees intelligence as an abstract capability of the individual based on a mechanism for rational thought. The new approach is not based on the idea that the mind is a rational processor of symbolic information, nor does it require the idea that thought is a kind of abstract problem-solving with a semantics that can be understood independently of its embodiment. Instead, priority is given to affective and mimetic responses that serve to engage the whole organism in the life of the communities in which it participates. Intelligence is seen not as the deployment of capabilities for problem-solving, but as the continual, ever-changing and unfinished engagement with the environment in terms of narratives. The construction of the identity of the intelligent individual involves the appropriation or taking up of positions within the narratives in which it participates. Thus, the new approach argues that the functioning mind is shaped by the meaning ascribed to experience, by its situation in the social matrix, and by practices of self and of relationship into which its life is recruited. Classic AI models such as goal-directed problem solving then can be seen as special cases of narrative practices instead of as ontological foundations. There are implications for representation and performativity that have given us new confidence in the form of a 'strong AI' attitude.

1 Introduction

A symposium on designing a functioning mind in the year 2000 reminds me of the endeavours of aircraft inventors about a hundred years ago. Each inventor built his own flying machine and attempted to demonstrate it before witnesses. Every known idea for heavier-than-air flight was represented: machines with flapping wings, machines with a stack of six or more wings, giant rotating helical spirals; airplane/balloon combinations; with power by bicycle, steam and diesel engines. Test flights invariably ended in failure, and whatever achievements were demonstrated, sustained controllable flight was not one of them. We can look at newsreels of early demonstrations with amusement secure in the knowledge that modern air travel is safe and efficient. This is possible because we have only one thing the pioneers of flight lacked: a readily available theory of aerodynamics. The pioneers had sufficient technology and engineering knowledge at the time, but no foundation describing how engineering skills might make the best use of the available technology.

Designing a functioning mind is difficult partly because there are no ready-made theories, and partly because it is not clear that what we can theorise about is relevant to the concern. If there were a reasonably complete theory of the mind, we would be in a position to implement at least parts of it to an impressive degree. And yet this has not yet happened to the satisfaction of most researchers. Other fields have had an easier task because of the existence of ready-made theories developed within the last 150 years. For example, electromagnetism gives the relationship between an electric field and a magnetic field (Heaviside, Maxwell), and can be used to simulate, design and analyse electrical circuits as well as to develop better theories of electromagnetism. Aerodynamics gives a relationship between lift, thrust, drag and gravity. This provides a foundation for numerical simulation of aircraft, helps to design aircraft, helps to understand natural flight, and helps to develop better theories of aerodynamics. AI has a set of ready-made theories such as algorithms for problem solving, but these theories are not about the mind. The lack of a suitable theory of functioning minds is a problem.

Problems are accompanied by paradoxes. The paradox of AI lies in the distinction between mundane and expert tasks. People perform mundane tasks automatically, yet these seem to require complex reasoning. The principles of washing dishes or changing babies' diapers or hammering nails or installing automobile windshields

can be taught within minutes to people with widely varying intellectual abilities, creativity and insight. These tasks involve the handling of imprecise quantities of intractable materials such as soap suds, cloth, slippery dishes and glass sheets, all difficult to model mathematically. In these tasks no precise measurement is called for ('squirt about this much soap into the sink'), specifications are incomplete ('then jiggle it round until it fits'; 'then pin these two bits together') and circular ('you know it fits when it snaps into place'), yet dextrous perceptual-motor operations and sensitive judgements are employed. The fact that apparently easy tasks may conceal a large amount of unconscious sophisticated analysis and processing is not the point. Whether these tasks need 'insight' or 'representations' or 'knowledge' or 'problem solving' is also not the point. The point is that these tasks are taught by 'showing how' and learned 'by doing', relying on social and mimetic interaction, with performance parameters limited by embodiment.

By contrast, expert tasks require specialised skill and training, for example medical diagnosis, playing expert chess, financial decision making. Paradoxically, mundane tasks are the harder to automate, and expert tasks are the easier to automate. Why is it hard to program computers to do things people find easy to do? And why is it easy to program computers to do things that people find hard to do? The answer is that implementing expert tasks does not require theories of the person. It requires only a theory of medical diagnosis or chess or financial markets or whatever. We can formulate such theories, and the implemented system can function with reference to and perform relative to such theories, and this is often sufficient to perform the task. On the other hand, implementing the mundane tasks involve knowing how the embodied person relates to things or other persons, and this is one of the theories we lack today.

I have argued elsewhere (Clocksin, 1995, 1998) that AI is limited not by processing power nor in handling complexity, but by taken-for-granted concepts that form an implicit normative load. Historically, AI research has been freighted by a 'command and control' paradigm (Edwards, 1996) in which AI is seen as the application of decision processes, represented using formal reasoning, to abstract well-specified tasks. Of the vast literature demonstrating this point, let two examples (Turner, 1984, Ramsay, 1988) suffice, or look at any large AI textbook. Indeed, the effort to make progress in AI through one style of knowledge representation and automated reasoning, logic programming and Prolog (e.g. Baral and Gelfond, 1994), culminating in Japan's Fifth Generation Programme of the 1980's, is probably the most ambitious and well supported manifestation of this paradigm. Autonomous agent and rational agent research is one of the more recent followers of the 'rationalist' paradigm.

It may be necessary to break away from the rationalist tradition in order to do justice to alternative models of minds, models based on persons, identity and social context, that may ultimately provide a workable foundation for AI achievements. There are signs that this 'reframing' of AI is happening. For example, Picard (1997) argues that AI has long ignored the importance of emotions, and that emotions are part of the essential foundation for intelligence. Although I am in sympathy with Picard's conclusion, I agree with Sloman (1999) that Picard and others misinterpret Damasio's (1994) observations concerning frontal lobe damage. The observed effects of such damage do not actually imply that emotions are essential for intelligence.

The work described in this paper is informed by two strands of thought emerging from social and developmental psychology. First, there has been an increasing concern with personhood: with persons, agency and action, rather than causes, behaviour and objects (Shotter and Gergen, 1989). Second, an emphasis on the self as a social construct, that persons are the result of interactions with significant others, and that the nature of these interactions is in turn shaped by the settings in which these interactions occur (Levine, 1992).

2 Reframing Rationality

While AI's emphasis on logical symbol-processing as the basis of intelligence has been criticised by Weizenbaum (1976) and by Dreyfus (1972), Winograd and Flores (1985) were probably the first to call for a re-examination of the rationalistic tradition in its influence on artificial intelligence research. According to Dreyfus and Dreyfus (1986),

> the failure of AI research will continue until intelligence ceases to be understood as abstract reason and computers cease to be used as reasoning machines.

Furthermore, McDermott (1987) argues that

> ...the skimpy progress observed so far is no accident, and in fact it is going to be very difficult to do much better in the future. The reason is that the unspoken premise..., that a lot of reasoning can be analysed as deductive or approximately deductive, is erroneous. (p 151)

Finally, the following observation from Weizenbaum (1976) has enormous significance:

> …intelligence manifests itself only relative to specific social and cultural contexts.

Is man, as Aristotle put it, a rational animal? Rationality in logical terms is related to consistency: an argument is rational if it does not produce a contradiction. Beginning with Doyle (1979) and McCarthy (1980), much effort has devoted to methods for preventing contradictions in the presence of changing assumptions. And yet human performance in problem solving is marked by two characteristics that suggest that people sit rather more lightly to logical consistency: people can happily entertain contradictory views (sometimes without being aware of it), and when put to the test, human 'rationality' is frail and fallible. People do not rely on deduction in their everyday thinking. Human experience is marked by incompetence, blunders, and acts of misjudgment. The average AI programmer (but not your average psychologist) might be surprised to learn that for normal people, irrational behaviour is the *norm* rather than the exception. The more accessible background literature on the 'normality' of irrationality includes Sutherland, Manktelow and Over, Dixon, and Piattelli-Palmerini. Human 'reasoning' thrives on the basis of actions and beliefs that cannot be justified nor supported logically.

So is the human mind based on a capacity for rational thought? If so, then why is there a lack of connection between rationality and everyday behaviour? And, if a rational mind is based on a rational brain, then why should we have a rational brain at all, as it seems to find little employment as an engine of rationality? So which is the more likely hypothesis:

> Intelligence derives from a rational mind, operating according to rules of logic (with truth-maintenance and the rest), but is hardly ever used in the way 'intended';

OR

> The mind is not based on principles of rationality, but as people in society we have come to perform and to value a developed but imperfect rational discourse because it has tangible benefits.

3 Construction and the Mind

The social psychology literature tells us that it is useful to make a distinction between *constructi**vism*** and *construc**tionism***. Constructivism, a term coined probably by Nelson Goodman and also called cognitive constructivism, describes the work of Piaget and Neisser. The main idea is that reality is created through the operation of a variety of mental schemata, analysis and synthesis procedures. This is the main assumption behind cognitive psychology, and it is safe to say this is the prime assumption of AI research. AI is concerned with the design of data structures to implement mental schemata, and the design of algorithms to implement analysis and synthesis procedure.

By contrast, constructionism describes the work of Gergen, Potter and Harré. It has also been called discursive constructivism, which is probably the source of the confusion of terminology. The idea is that constructive processes are to be found in relationships, often discursive, between persons. These relationships embody situated reasoning: that is, a contextualised meaning-producing performance. Figure 1 shows cartoon versions of constructivism and constructionism.

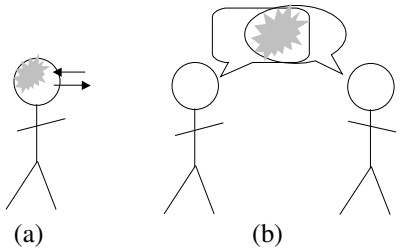

Figure 1. (a) the constructivist concern; (b) the constructionist concern.

The cartoon in Fig 1(a) is a familiar AI model: the problem-solving agent, operating a perception-processing-action loop. Meaning is constructed in the mind as a result of internal capability. By contrast, Fig 1(b) privileges social interaction as the meaning-constructing process. Here the mind is for social development, engagement with persons, and the institution and appropriation of personhood and identity. Figure 2 shows a pair of layers describing the functional dependencies posited by both approaches.

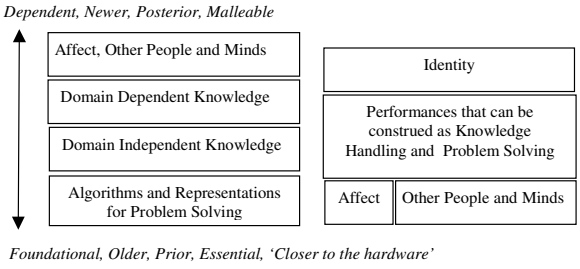

Figure 2. (a) Conventional layering; (b) Constructionist layering.

An example of way that affect is marginalised by the conventional understanding of the mind is illustrated by Mr Data, the android character on the popular television programme *Star Trek: The Next Generation*. Mr Data can function perfectly well without emotions, but has an optional plug-in 'emotion chip' that makes him more hu-

man. With the chip at his disposal, Mr Data is at least more versatile than Mr Spock of the original *Star Trek* series. Half-human, half-Vulcan, Spock is trained to abhor emotion. Given that this is after all science fiction, these characters will particularly appeal to adolescent males who themselves can find difficulty coming to terms with their emotions and the complexities of the social realities in which they are beginning to find themselves. How much more comforting to retreat to a simplified world by identifying with heroes who succeed by applying pure logic and calculation.

By contrast, the constructionist finds such accounts incoherent. Not only are there glaring inconsistencies in the model – the Data and Spock characters actually behave in ways that use the affective skills they are reputed not to require – but also the 'optionality' of affect is associated with individuals one sees only in the human population as profoundly disordered. It is probably intuitions of this type that lead myself, Picard and others to reject the optionality of affect implied in Fig 1(a), and instead to adopt a view like Fig 2(b) which treats the world of affect as something prior to intelligence and problem-solving. Developmentally this reflects the early requirement for social relationship in the form of mother-newborn relationship.

4 The Unfinished Mind

Intelligent behaviour depends on the challenges that arise from being embodied and being expected to participate in social interactions. Every organism is socialised by dealing with -- engaging with, participating with -- the problems that confront it according to its capabilities. This activity has a spiraling effect, as capabilities bring about social experience which in turn modifies -- conditions, improves -- the capabilities for further social experience. The functioning mind's standpoint in this engagement is what brings about -- or institutes or constructs or establishes -- an identity. Because development occurs throughout the organism's life, its identity is never fixed or foreclosed. From the constructionist perspective, identity can never be understood as a prior given, even if the organism may understand it (*i.e.* endorses a metanarrative from a culture that names it) as such. AI research will make progress when this spiral process of social and cultural development, and its relationship to the ontogenesis of the individual's identity, can be understood in clear and unambiguous (*i.e.* computational) terms.

One way to describe the developmental process is the so-called 'hermeneutic circle' of critical theory, depicted in Figure 3.

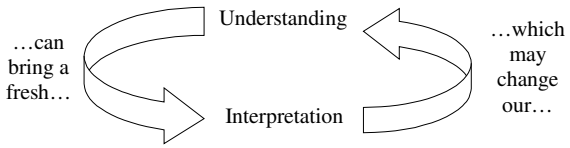

Figure 3. The provisional, contingent, unfinished nature of engagement illustrated by the hermeneutic circle.

Interpretation and understanding influence each other in a circular way. To understand a text, we need to interpret it. But understanding it may lead us to consider a fresh interpretation. When the text is read in the light of the new interpretation, it may change our understanding of it. This may call for a revised interpretation, and so on. This circle is again better described as a spiral, for one never visits the same locus in hermeneutic phase-space (as it were) twice. There is also no reason to restrict the number of trajectories through the space to one. AI research needs to be more attentive to the continual and developmental character of the hermeneutic circle. A system that is not committed to the hermeneutic circle cannot be said to understand. It is this capacity for continual co-evolution (in this example, of interpretation and understanding), or ontogenesis, that is characteristic of a sustaining system. Such a sustaining system is what I would call a functioning mind. The AI thoughtworld needs to reject models such as 'finding the answer to a problem' and 'maintaining consistency of beliefs'. By contrast, ontogenesis of the type advocated here may proceed for a system in which problems and answers are not neatly defined, and consistency in a logical sense is never achievable. Thus there needs to be an emphasis on the *provisionality* of thought, another hallmark of a social constructionist perspective.

5 A Narrative Architecture

The new approach to AI research outlined here is built upon the idea of a narrative, and is adapted from Clocksin (1998). The narrative or story provides a framework that facilitates the interpretation of experience, for it is through the narratives people have about their own lives and the lives of others that they make sense of their experience. Narratives are not restricted to the form of a story with its initial complication followed by its denouement, although stories are good examples of narratives. The idea of narrative is not restricted to written texts. Not only do narratives influence the meaning that people give to experience, they also influence which aspects of experience people select for articulation. Narratives provide not merely a reflection or mirror of life, but provide for the shaping and structure of life. Alasdair MacIntyre (1981) recognises several diverse uses of nar-

rative. He argues that human action is narrative in form, that human life has a fundamentally narrative shape, and that people are story-tellers who position their lives and arguments within narrative histories. Communities and traditions are invested with continuity through narrative histories, and epistemological progress is marked by the construction and reconstruction of more adequate narratives. For Jerome Bruner (1986), narrative is one of the modes of cognitive functioning, a way of knowing that provides a distinctive way of ordering experience and constructing reality. The other mode is logical argument, and Bruner notes that while narrative deals with 'the vicissitudes of human intentions' and is built upon concern for the human condition, logical arguments are either conclusive or inconclusive.

The functioning mind is shaped by a play of images and allusions of the subtle and elusive kind that owes more to the imagination of the storyteller than to the rational operations of the logician. Consider first the question of emplotment. Every story has a plot that involves movement from an initial tension to the resolution of that tension, what Aristotle literally called binding and loosing (*Poetics* 18.1-3) respectively. More generally, a narrative may be described as any time-extended behaviour of a system that exhibits one or more sequential *tension• release* patterns, each of which may recursively embed other *tension•release* patterns within it. Each *tension• release* pattern may be termed an episode. We identify narratives at every scale of time-extended behaviour, from the charge•discharge electro-chemical activity of nerve cells to the impulse•relax rhythmic control of coordinated motor activity, to the expectation•emission of activity that makes a visual or acoustic impact, to the problem•solution of goal-directed behaviour, to the preparation•resolution of musical and poetic cadence, to the structure of songs and dances, building up all the way to people discussing the latest round of tax increases and where they're going to eat lunch. My idea of narrative is informed but not restricted by the ideas of the discursive psychology of Harré and Gillett (1994) who are mainly concerned with higher cognitive functions and who see the mind as

> ...embedded in historical, political, cultural, social, and interpersonal contexts. It is not definable in isolation. And to be a psychological being at all, one must be in possession of some minimal repertoire of the cluster of skills necessary to the management of the discourses into which one may from time to time enter. [p 25-26]

Thus discourse can be related to emplotment, and the other typical dimensions of narrative, namely setting and character, are related to environment (context) and identity (being), respectively.

The main difference between narratives and other structures such as knowledge representations (Way, 1991) is that narratives serve to engage the individual with the contingencies of identity and relationships with others. The individual develops or constructs identity by locating itself within the discourse -- the stories, narratives -- with which it engages. Thus, identity is not simply a matter of building up a store of knowledge or cognitive capabilities, or the self-assembly of numerous micro-personalities into a smaller number of personae (*pace* Jung), but the taking up of a position within a set of discourses as they are negotiated. Therefore, behaviour involves a certain performance and production of a self which is the effect of a discourse, and which has a circular quality. The fact that a developing self may claim to represent itself as a truth prior to the discourse must be seen as a phenomenon emerging from the engagement rather than as an ontological assumption.

This 'taking up of a position' within a narrative may be identified at all levels: from the steady-state behaviour of an electro-chemical process in tissues, to the habituation of a tissue in response to a pattern of conditioning stimuli, to the formation of a reflex. At the highest social level it may involve the articulation of or appropriation of or commitment to a particular policy, a system of beliefs and values, a family group, a football team or a gender role. The ways in which one plays social games is the way in which the identity becomes, to use Butler's (1991) list of words, 'established, instituted, circulated and confirmed' [p. 18]. It is important to remember the ontogenetic nature of the process: the taking up of a position within a narrative is not entirely determined by genetics, nor is it entirely determined by culture. All enabling capabilities participate in their own way with the development of the system.

Let me outline some broad implications of understanding the functioning mind in terms of narrative and the narrative generation process:

a) Narratives can be repeated and imitated. Aristotle defines man as the political and rational animal, but we should also remember his observation that 'man differs from other animals in his greater aptitude for imitation [mimesis]' (*Poetics* 4). Mimesis is seen as the fundamental operator over narratives, implicated in the processes of memory, consciousness and performance described below. Mimesis is not a process of loss-free copying: rather, it a capability that culminates in 're-storying', or re-telling or re-authoring of a story.

b) Narratives, being indeterminate and having hierarchically decomposable emplotment, are a convenient form for concealing and revealing subnarratives. A narrative is extensible both along its length and its hierarchy, for it is always possible to apply questions such as, 'and *then* what happened?', or 'where did he come from?', or 'how did she do that?', or 'why did he say that?', or 'why didn't she do that?' as a means of closing or opening new subnarratives and filling in or creating gaps.

c) Narrative serves as a memory, not only in a collective and institutional sense of telling stories to each other (memorial), but also as the foundation for the individual's memory (Middleton and Edwards, 1990). Memory as narrative provides an alternative to the conventional idea of memory as a storehouse of knowledge and experience, and of remembering as a process of retrieval from the storehouse. Instead, memory is understood not as the storage-and-retrieval manipulation of a network of entities, but rather as a particular practice within the context of a discourse. Memory is an active process in which experience is emplotted (*i.e.* the lived experience is 'recruited' by a prior narrative) and personified (*i.e.* attributed the status of a 'character' in a prior narrative) in ways constrained by the usages and contexts of prior narratives. This 'circular' definition is committed to the idea of memory not as storage and retrieval of facts, but as the continual re-membering of our memories of our memories *(sic)*. Memories cannot be decoupled from the narrative structure of remembering, for outside a narrative they have no meaning (Shotter, 1990).

d) Narratives can be edited, and this is particularly relevant to the question of memory and learning. As Platinga puts it, 'our memories are not inert, but undergo a process of editing, whereby they are regularized, rendered more retainable and reshaped with an eye to subsequent circumstances and events' [p. 45].

e) Learning and therapy are understood as means of editing or rewriting narratives. These both begin with the assumption that people experience problems when the stories of their lives, as they or others have constructed them, do not sufficiently represent their lived experience. Therapy and learning then become processes of storying and re-storying the lives and experiences of people. In this way narrative comes to play a central role in therapy and education, but for the purposes of this paper point to models of 'learning' and 'belief revision' in the functioning mind. There is a significant literature on what might be described in my terms as narrative-theoretic formulations of memorial, mimesis, ritual, education and therapy, for example Elliott (1995), Girard (1977), Buckland (1995) and Epston and White (1992).

f) A narrative is a suitable representation for self-reflection. Whether consciousness is considered as a capacity for, or a result of, self-reflection, consciousness is about making explicit in narrative form (*i.e.* telling a story about) the person's interpretation of the narratives with which it believes itself to be engaged. The nature and extent of the belief is also conditioned by the person's appropriation of narratives from the culture, giving rise to 'cultural conditioning'. Consciousness is the story we tell about where (context) we (character) are and what (plot) we are doing. The existence of what Dennett (1991) calls the 'Cartesian theatre' as a setting for this performance therefore should not be surprising, and certainly does not imply endorsement of Cartesian dualism. Consciousness as an 'account of ourselves' therefore proves necessary because we are animals who partly construct our accounts through interaction with a social world: we need accounts of ourselves because we are accountable 'as selves' to others, and this exchange of accountability again has a spiral instituting character. The nature and purpose of consciousness therefore cannot be understood in relation only to the individual.

g) Narratives can be performed. Performance (whether it has an inward or outward expression) is the means by which narratives are rendered to self and others, and depends on the capacity for mimesis. Performance is more general than reciting a story on a stage. It is the actual means of behaviour. Yet performance, like remembering, does not render a narrative intact. This understanding concurs with Butler (1993, p 187) in opposing the conventional notion that performativity is the efficacious expression of the human will. An act of performance is always ambiguous and open to interpretation. Each performance is a unique act of creation because it opens a new gap between intention and realisation, between presentation and representation.

h) Because mimesis (the mechanism for re-storying, performance, memory) generates non-identical narratives that are open to interpretation, it is a wellspring of perpetual conflict. Conflict leads to rivalry and violence, but also to dialectic and symbolic interaction. Thus the provision of narrative-building tension together with a motivation for social interaction are built in at a fundamental level.

Narratives that emerge at the layer of surface behaviour are particularly important, for all performance is understood within the context of the social group. At a fundamental level, narratives are performed as rituals and myths, both of which also have the instituting effect of maintaining a corporate memory. Myths supply a framework of images, in narrative form, in which the needs, values and fears of the group – in short an articulation of affect within a network of social consciousness – can be expressed (Clocksin, 1995). Rituals are a demonstration of the way that sensations and actions are integrated as a result of their co-development, and are expressed by the articulation of a policy by members of the group. Recalling Weizenbaum's observation that 'intelligence manifests itself only relative to specific social and cultural contexts', I suggest that ritual and myth as the performative means of the social and cultural context are the prior and necessary components for the manifestation of intelligent behaviour. This, like the layerings of Figure 2, implies a view of the mind that is an inversion of the conventional view taken by artificial intelligence and cognitive science researchers. The conventional view assumes that behaviour arises from basic capabilities for cognition and problem solving. The emphasis seems to be on the individual mind that generates solutions to problems. The questions of emotion and interaction with others are considered as additional – possibly even unnecessary or irrelevant – complications or side-issues. By contrast, the 'upside down' view sees problem solving as a practice or policy that has been constructed as a *result* of a basic capacity for social interaction within a cultural context, which itself is based on fundamental reciprocalities of need and desire. Therefore, further progress in artificial intelligence research needs to be attentive to issues arising from the articulation of practices and policies by a group. An example of this is given in Clocksin (1998).

6 Conclusion

I have put forward a view that the functioning mind should not be seen as a kind of abstract puzzle solving applied by an individual to arbitrarily defined problems presented by an alien environment. It is true that certain isolated episodes of intelligent behaviour can be given an interpretation as puzzle-solving, maximising of utility, and so forth. However, we need to accept that there is a close ontogenetic connection between the nervous system and the environment. It is possible to accept that this connection has the effect of individualising precisely which capabilities and information sources are employed by the mind as it develops, without necessarily endorsing a structuralist ontological commitment as to the precise delineations of capability and information source.

The AI system of the future will be based upon what Darwin called 'the social instincts'. The functioning mind cannot be considered in isolation from an embodied existence that is committed to an ontogenetic engagement with the environment. The body matters, and in this way the functions of the mind might be said to resemble physical functions such as digestion which can't take place without a body. But intelligence is of a different order because at its core is the generation and use of images and narratives that serve to circulate within the gap between presentation and representation that is opened up by performance.

I repeat Weizenbaum's dictum: Intelligence manifests itself only relative to a matrix of social and cultural contexts. There is a propinquity of participation in communities that are defined by historically individualised mutualities and reciprocalities of need and desire. The ways we participate with this matrix, and construct identities that take up positions within this matrix, are through policy, contract, and institution. This participation and construction, which makes explicit the political imperative of intelligence, is made possible by a variety of means such as memorial, mimesis, ritual, education and therapy, all of which serve to expose and rewrite narratives. The functioning mind is thus seen as an historical contingency constituted in part by policies articulated by the social group, rather than by the underlying cognitive nature of the individual brain. This perspective on the mind is what Rorty would call 'ironic' and 'political' in the sense that it makes explicit its theoretical commitment to policy, community and institution.

An implication for future research is the idea of 'narrative architectures', built upon a foundation which is attentive to the way that signals and symbols influence (and are influenced by) the way we as communities of individuals make sense of experience, construct our identities, and produce meaning in the world. This makes me optimistic about the prospect for what John Searle called 'strong AI'. The characteristic of an AI system that really can be said to have a functioning mind is that it can perform an account of the identity it has constructed using narratives it has appropriated from the culture and society that it shares with us. Such a system negotiates its positions to become instituted among the group, participating in the social consciousness it constructs with members of the group.

Acknowledgements

I thank John Cornwell, John Searle, Pam Clocksin and Aaron Sloman for discussions.

References

Baral, C. and M. Gelfond, 1994. Logic programming and knowledge representation. *Journal of Logic Programming* **19**, 73-148.

Bruner, J., 1986. *Actual Minds, Possible Worlds*. Harvard University Press.

Buckland, S., 1995. Ritual, bodies and 'cultural memory''. *Concilium* 1995/3, 49-56

Butler, J., 1991. Imitation and gender subordination. *Inside/Out*, ed D. Fuss et al. Routledge. 13-31.

Butler, J., 1993. *Bodies that Matter*. Routledge.

W.F. Clocksin, 1995. Knowledge Representation and Myth. Chapter 12 of *Nature's Imagination: The Frontiers of Scientific Vision* (J. Cornwell, ed). Oxford University Press.

W.F. Clocksin, 1998. Artificial Intelligence and Human Identity. Chapter 6 of *Consciousness and Human Identity* (J. Cornwell, ed.) Oxford University Press.

Damasio, A.R., 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. Grosset/Putnam Books.

Dennett, D.C., 1991. *Consciousness Explained*. Penguin.

Dixon, N., 1976. *On the Psychology of Military Incompetence*. Futura.

Dixon, N., 1987. *Our Own Worst Enemy*. Futura.

Doyle, J., 1979. A truth maintenance system. *Artificial Intelligence* **12**, 231-272.

Dreyfus, H.L., 1972. *What Computers Can't Do*. MIT Press.

Dreyfus, H.L. and Dreyfus, S.E., 1986. *Mind over Machine*. Macmillan.

Edwards, P.N., 1996. *The Closed World: Computers and the Politics of Discourse in Cold War America*, MIT Press.

Elliott, C.M., 1995. *Memory and Salvation*, SPCK.

Epston, D. and White, M, 1992. *Experience, Contradiction, Narrative & Imagination*. Dulwich Centre Publications, Adelaide, South Australia.

Girard, R., ET:1977. *Violence and the Sacred*. Baltimore: Johns Hopkins University Press.

Harré, R. and Gillett, G., 1994. *The Discursive Mind.* London: Sage Publications.

Levine, G., 1992. *Constructions of the Self*. Rutgers University Press.

MacIntyre, A., 1981. *After Virtue: A study in moral theory*. Duckworth.

Manktelow and Over, 1990. *Inference and Understanding*. Routledge.

McCarthy, J., 1980. Circumscription – a form of non-monotonic reasoning. *Artificial Intelligence* **13**, 27-39.

McDermott, D., 1987. A critique of pure reason. *Computational Intelligence* **3**, 151-160.

Middleton, D. and D. Edwards, 1990. *Collective Remembering*. Sage.

Piattelli-Palmarini, M., 1994. *Inevitable Illusions: How mistakes of reason rule our minds*. MIT Press.

Picard, R., 1998. *Affective Computing*, MIT Press.

Platinga, T., 1992. *How Memory Shapes Narratives*. Lampeter:Mellen.

Ramsay, A., 1988. *Formal Methods in Artificial Intelligence*. Cambridge University Press.

Rorty, R., 1989. *Contingency, Irony, and Solidarity*. Cambridge University Press.

Shotter, J., 1990. The social construction of remembering and forgetting. In Middleton and Edwards (1990).

Shotter, J. and K.J. Gergen (1989). *Texts of Identity*. Sage.

Sloman, A., 1999. Review of Affective Computing, *AI Magazine* **20**(1), 127-133.

Sutherland, S., 1992. *Irrationality: The enemy within*. Penguin.

Turner, R., 1984. *Logics for Artificial Intelligence*. Ellis-Horwood.

Turner, R., 1990. *Truth and Modality for Knowledge Representation*. Ellis-Horwood.

Way, E.C., 1991. *Knowledge Representation and Metaphor*. Kluwer.

Weizenbaum, J., 1976. *Computer Power and Human Reason*. Penguin.

Winograd, T. and Flores, F., 1985. *Understanding Computers and Cognition: A new foundation for design*. Addison-Wesley.