The Emerald Research Register for this journal is available at
**www.emeraldinsight.com/researchregister**

The current issue and full text archive of this journal is available at
**www.emeraldinsight.com/0368-492X.htm**



# Second-order cybernetics and enactive perception

J.M. Bishop
*Goldsmiths College, University of London, London, UK, and*
J.S. Nasuto
*Department of Cybernetics, University of Reading, Whiteknights, Reading, UK*

## Abstract

**Purpose** – To present an account of cognition integrating second-order cybernetics (SOC) together with enactive perception and dynamic systems theory.

**Design/methodology/approach** – The paper presents a brief critique of classical models of cognition then outlines how integration of SOC, enactive perception and dynamic systems theory can overcome some weaknesses of the classical paradigm.

**Findings** – Presents the critique of evolutionary robotics showing how the issues of teleology and autonomy are left unresolved by this paradigm although their solution fits within the proposed framework.

**Research limitations/implications** – The paper highlights the importance of genuine autonomy in the development of artificial cognitive systems. It sets out a framework within which the robotic research of cognitive systems could succeed.

**Practical implications** – There are no immediate practical implications but see research implications.

**Originality/value** – It joins the discussion on the fundamental nature of cognitive systems and emphasise the importance of autonomy and embodiment.

**Keywords** Cybernetics, Cognition

**Paper type** Conceptual paper

## 1. Introduction

> It should be noted that from now on "the system" means not the nervous system but the whole complex of the organism and the environment. Thus, if it should be shown that "the system" has some property, it must not be assumed that this property is attributed to the nervous system: it belongs to the whole; and detailed examination may be necessary to ascertain the contributions of the separate parts (W. Ross Ashby, 1952).

An oft repeated aphorism is that the world is in a perpetual state of flux and hence that our universe is constantly changing. Thus, in order to behave intelligently within the natural environment any Cybernetic system, be it man, machine, or animal, faces the problem of perceiving invariant aspects of a world in which no two situations are ever exactly the same. Cartesian theories of perception can be broken down into what Chalmers (1996), calls the "easy" problem of perception; the classification and identification of sense stimuli and a corresponding "hard" problem, which is the realization of the associated phenomenal state. The difference between the "easy" and the "hard" problems and an apparent lack of the link between theories of the former and an account of the latter has been termed the "explanatory gap".

Many current theories of natural visual processes are grounded upon the idea that when we perceive, sense data are processed by the brain to form an internal

Kybernetes
Vol. 34 No. 9/10, 2005
pp. 1309-1320

0368-492X
DOI 10.1108/03684920510614696

representation of the world. The act of perception thus involves the activation of an appropriate representation. Thus, the easy problem reduces to forming a correct internal representation of the world and the hard problem reduces to answering how the activation of a representation gives rise to a sensory experience.

In machine perception progress in solving even the "easy" problem has so far been slow; typical bottom-up (or data driven) methodologies, involve the processing of raw sense data to extract a set of features; the binding of these features into groups then classifying each group by reference to a putative set of models. Conversely, in typical top down methods, a set of hypotheses of likely perceptions are generated, which are then compared to a set of features in a search for evidence to support each hypothesis. Historically, cybernetic approaches have favoured the former and computer science the latter; however, amalgams of the two have also been explored. To date the success of all approaches has at best been patchy and limited to a very small subset of the human perceptual gamut.

## 2. First-order cybernetics

First-order cybernetics (FOC) characterises agent-environment systems in terms of feedback loops whose operation can be interpreted by an observer (or engineer) in terms of teleological behaviour (i.e. moving towards a goal). Alternatively, an engineer may manipulate a system and include in it feedback loops in order to achieve behaviour consistent with proscribed purpose. An early example of such a behaviour is found in the work of W. Grey Walter. He demonstrated that apparent teleological behaviours such as following a light source (without approaching too closely) can be instantiated in a very simple FOC device (Plate 1). Observing their behaviour Walter remarked that, "despite being crude (the tortoises) conveyed the impression of having goals, independence and spontaneity". Yet, in our opinion, such a teleological interpretation of tortoise behaviour is unwarranted as this behaviour was explicit in the design.

Other examples of FOC systems include Gaia, Lovelock's cybernetic view of planetary feedback processes that maintain stable conditions suitable for life; or paradigmatic control system such as the Watt's governor, used to control the speed of a steam engine under varying loads. Cyberneticians, such as Wiener and McCulloch

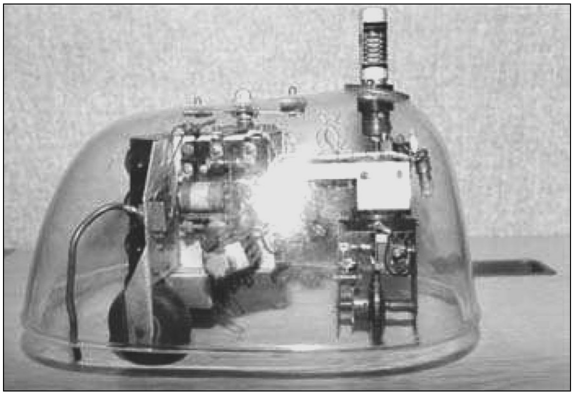

**Plate 1.**
Walter's tortoise (copyright Burden Neurological Institute, Bristol, UK)

investigated the operation of the nervous system from this perspective, which led to the development of bottom up, (connectionist), models of cognitive processes.

### 3. Second-order cybernetics

FOC is concerned with the observation and control of systems, (agent and environment) positing a distinguished role for an observer, as an entity decoupled and independent of the system. In contrast, second-order cybernetics (SOC) recognises the inseparability of the observer and the system. There is no observer outside the system; the agent is the cognitive observer of its environment, exposing a distinction between the FOC of observed systems and the SOC of observing systems (Von Foerster, 1974). This change of view explicitly focuses SOC onto the explanation of cognitive processes as determined by the agent-environment coupling dynamics.

In SOC the observer has no knowledge of how the world "really is" – there is no homunculus observing an internal model of the external world. Instead, SOC highlights the fundamental distinction between the physical world and our perception of it; inner "models" are not representations of an outer reality, but subjective dynamic constructions that, by complex feedback paths within the observer and environment, move the system towards its emergent goals.

In fact, SOC recognises a multitude of potential alternative reality constructs – our everyday life unfolds but one of them. This implies a move away from a concept of a common objective reality replacing it with individual observer relative constructs. Many illustrations of this come from the study of brain damage; one such example being motion blindness. A patient with this condition is unable to cross a street because the motion of cars is invisible to her: a car is up the street and then upon her, without ever seeming to occupy the intervening space. Analogous to someone watching a movie with a low frame rate, this patient perception of the world is of a series of still images.

The fact that we have evolved able to perceive the world in motion and at rest is a result of one possible path of evolution; alternative paths are feasible, for example, some simple vertebrates (e.g. frogs) are only able to perceive moving prey – they will ignore a nearby stationary fly even when hungry.

However, in practice the danger of complete relativism, where any perception of the world is as good as any other, is avoided by "coherence" and "invariance": coherence being a social process whereby phenomena become real by consensus; invariance being a fundamental property of the world entailed by physical laws whereby entities tend to maintain their properties over time.

At the heart of FOC there is an asymmetry in the closed loop feedback (circular causality), which posits the observer outside of the loop, which SOC deconstructs by including a human (observer) in the loop. This results in the different concepts of goal-directed system behaviour (teleology) championed by both theories. In contrast to FOC, in which the system's teleology is a result of manipulation or intepretation of the agent's behaviour by an external observer decoupled from that system, in SOC teleological properties are observer/agent relative emergent properties and not externally defined objective properties of the system.

To summarise, the deconstruction of the asymmetry inherent in FOC has at least three significant implications.

(1) The observer looses its distinguished position and can be treated as just another part of the whole system.
(2) SOC is inherently and explicitly concerned with explaining the observer's cognitive processes (including goal orientedness).
(3) SOC entails a constructivist epistemology (theory of knowledge) which starts from the assumption that, "the thinking subject has no alternative but to construct what he or she knows on the basis of his or her own experience" (Von Glasersfeld, 1995).

In the remainder of the paper, we will explore links between SOC, the enactive theory of perception (ETP) and dynamic systems (DS).

## 4. Dynamic systems theory of cognition (DSC)

The DSC (Port and Van Gelder, 1981; Van Gelder, 1997), offers another alternative framework to conventional computational theories of mind in which . . .

> . . . cognitive systems are computers (digital, rule-governed, interpretable systems), with a modular internal structure; they interact with their environments in a cyclic process that begins with input transducers producing symbolic representations in response to the environment, continues with sequential internal computations over symbolic structures, and ends with output transducers affecting the environment in response to the symbolic specifications; the whole process can be considered independently of the body and the environment except insofar as they deliver occasional inputs and receive outputs (Van Gelder, 1997).

Thus, in computational systems cognition is equivalent to transformation of symbolic states representing knowledge or particular cognitive states. The transitions between states are instantaneous hence time does not play a role in their evolution; only the relative order of states does.

The symbolic nature of the state space implies that the magnitude of the changes or the time it takes to achieve them are undefined notions in computational theories of cognition. Moreover, the rules of evolution act locally, on a particular subset of representations and hence in this framework it is possible to consider cognitive acts in isolation from each other, the external environment and even the body.

In contrast, the dynamical approach treats cognitive systems as inherently dynamic which implies a profound change of perspective on their operation.

In the dynamical systems approach to cognition, the states are defined in terms of some numerical attributes and rules of state evolution are defined over those attributes and not over the knowledge representations. The latter can still be instantiated in, for example, system attractors, system trajectories, etc. However, the potential relations between such representations are not explicitly encoded in the system dynamics. The states are instantiated in the continuous state space and their changes take place in time, hence the latter can assume arbitrarily small values given correspondingly small intervals of observation. It is the rate of state change that is paramount to the dynamical system description. In contrast to the computational systems, it is much more natural within the dynamical framework to consider the changes of the total state of the system composed of mutually interacting or coupled parts with the ongoing modulation of states changes. This ultimately implies coupling of the cognitive system with the environment and the body.

However, although for Dynamicists the cognitive system is fundamentally embodied (being intimately coupled with its environment), supporters of DSC "conceptualise mental phenomena as state space evolution in certain forms of dynamical system" (Wheeler, 2002), hence firmly positing mental phenomena within the agent.

> Viewed under a post Cartesian perspective (rejecting the view of perception as the activation of appropriate a-temporal representations), the advantages of the dynamical account of cognition, which emphasises ongoing, real-time interaction of situated agents with a changing world, becomes clear (Van Gelder, 1997).

In viewing cognition as a continuous dynamic process, Dynamicists explicitly reject the notion of cognition as the computational manipulation of representations. The DSC outlines how to intelligently interact with the world, without the necessity of its explicit representation.

## 5. Enactive theory of perception

The enactive theory of perception (ETP) suggests an alternative paradigm for perception. Instead of considering that the operation of the nervous system leads to the creation of appropriate internal representations of the world, which somehow jump the "explanatory gap" to realise the relevant phenomenal states of experience, it considers the world as its own "representation" and perception as an embodied exploratory (enactive) process of the world mediated by sensorimotor contingencies. The theory championed by Varela and Thompson (Varela *et al.*, 1991) proposes three fundamental components which together give a full account of cognition. These include the low-level biological/neural processes (subject of third person accounts), the high-level phenomenological data (first person accounts) and formal dynamical theory as a bridge criss-crossing the explanatory gap between the two seemingly irreconcilable domains. ETP attributes the unity of apperception to large-scale neural synchrony. In its view the phenomenological states are emergent properties of the non-linear interactions of the body and the nervous system (upwards causation). It explains phenomenological casualty (downwards causation) by referring to the modulation of neuronal processes by global order parameters (phenomenological states). Although reference to the non-linear dynamics seems to bring ETP close to the DSC, the fundamental difference is the emphasis both theories place on the role of dynamics and embodiment. ETP, in contrast to DSC, stresses the importance of embodiment, constituting the physical substrate in which the cognitive processes evolve – "conscious experience occurs only at the level of the whole embodied and situated agent" (Varela *et al.*, 1991).

Varela and Thompson characterise three dimensions of embodiment which describe the relation between the embodied neural dynamics and phenomenology. These dimensions are intersubjective interactions in social behaviour; organismic regulation related to the operation of the autonomic nervous system, linking the fundamental physiological processes of the body to primal consciousness, or sentience – feeling of self; and finally, the sensorimotor coupling between an agent and the environment. A particular subset of ETP, which has recently attracted much attention, focuses predominantly on such sensorimotor coupling and purports to dissolve the explanatory gap and solve the problem of qualia by redefining this notion, "experience is not a thing that happens to people, but a thing that people do" (O'Regan, 2004).

*5.1 Sensorimotor account of visual cognition*

Sensorimotor account of visual cognition (SMC) is an idea rooted in Ryle's (1949) description of a thimble defined by the different perspective views it imparts as it is moved around in the visual field and which recently has been successfully developed and extended by O'Regan and Noe (2001)[1].

In SMC first person experiences are not states they are simply activities; hence to speak of phenomenal states of the brain is an example of what Ryle's (1949), termed a "category mistake", as there are no such states; qualia are illusions – there exist only the different acts of experience. The first person feeling of perceiving, say, the ineffable pink of a rose, arises purely from the specific sensorimotor contingencies of interacting with a pink rose; as opposed to say, interacting with a green apple. Experience is something we do and its qualitative features are simply aspects of our interactions with the world. Damasio (1996) proposed a somewhat analogous interaction between cognitive processes related to decision making and physiological states of the entire body.

As SMC is a general framework for vision, evidence for it is not direct and does not test the theory in the conventional sense; rather SMC accounts for several puzzling observations that are difficult to reconcile with conventional theories of vision. Hence, as evidence for their "Sensorimotor Account" O'Reagan and Noe (2001) discuss several problems which appear to fade under the generic spotlight of SMC, which are discussed in the following sub-sections.

*5.1.1 The stability of visual perception independent of eye saccades*. For over a century, when viewed within the standard framework of model based vision, where the job of the visual system is to transform the image of the world that is projected onto the retina into an equivalent internal 3D representation of the scene, it has been difficult to understand why perturbations of the image projected onto the retina caused by eye saccades, do not cause similar perturbations in phenomenal perception.

The mechanism that has historically been suggested to correct for such disturbances is the existence of correcting "extraretinal" signal. However, experimental results from Martin (1986) suggest that the candidate signals are both too inaccurate and too sluggish to correct for such perturbations. Conversely, viewed in the context of SMC, what remains invariant as the world is perceived is just the knowledge of how the sensed image (i.e. the pattern projected onto the retina) will transform as the eye saccades across the perceived scene.

*5.1.2 The non-perception of the "blind spot" and the perception of smooth visual continuity despite the non homegeneity of spatial and colour sensors in the eye*. A second puzzle for classical accounts of visual perception is why we do not explicitly perceive the blind spot (i.e. the location on the retina where the optic nerve emerges and where there are no photoreceptors), or detect that the acuity of the eye, (and the distribution of photoreceptors across the retina), falls off steadily from the central foveal area to the edge. Classical theories postulate some kind of compensation or "filling in" mechanism to account for this[2]. Conversely, in SMC perceptual experience of the world is exercised only by the sensorimotor contingencies defining how we expect the sensed data from a scene to change as the eyes saccades across it; in this context the fact that photoreceptors are not uniformly distributed across the retina does not pose a particular problem.

*5.1.3 Change blindness*. Change blindness experiments in particular highlight specific problems with classical theories of vision, which are not present in Enactive theories. In change blindness a subject alternatively observes a specific scene A and a modified version of it, scene B, significantly changed from the original. The succession of images is not instantaneous, between presentation of scene A and scene B there is a short period of blank screen, giving the appearance to the subject of a blink, masking the detection of low level, transient change. It has been experimentally observed that large-scale changes can be made to a visual scene and yet not be perceived by the subject. In classical "bottom up" theories of vision, the large scale differences in the two alternating images shown in Plates 2 and 3, would cause very different patterns of activation in low level visual processes.

Conversely, in "top down" theories the hypothesis that the images are the same is trivially disproved by the large scale disparity between the two. However, the subjects report of "not noticing the change" is entirely consistent with SMC, where the world serves as its own "outside memory" and the subject perceives only what he/she is enactively attending too.

**Plate 2.**
Change blindness; scene A
O'Regan (2001)

**Plate 3.**
Change blindness; scene B
O'Regan (2001)

Like SOC, SMC is inherently constructivist as knowledge of the world is actively constructed by the perceiving agent through its interaction with the environment. As Scott (1996) observes, "the 'objects' that we experience are 'tokens' for the behaviours that give rise to those experiences".

## 6. An alternative model of cognition; the unification of DSC, ETP and SOC

At the heart of computational theory of cognition there is the notion of computation as calculation arising from Turing's work at Blechley Park in WWII. During this period Turing was concerned with abstracting the essential processes carried out by [human] "computers" in the course of code breaking. This led to the formalisation of computation via the universal Turing machine. Subsequently, this notion led to the widely held view of the so-called Church/Turing thesis, asserting the equivalence of all "effective procedures" to perform tasks [models of computation] with Turing machines.

Hence, the view that computations are central to cognition was born. In fact, many scientists, particularly in cognitive science, take this metaphor literally, claiming that cognition is computation, meaning that cognitive processes are serial computational operations on appropriate knowledge representations.

In fact, any claim on the serial computational nature of cognitive processes is valid only insofar as it is a convenient mental shortcut meaning that their operation can be described in such terms. The equivalence of cognition with serial computation can only be the case insofar as the external third person interpretation of the cognitive processes goes; a view similar to the observer/(cognitive) system asymmetry of FOC. Such asymmetry imposes meaning on the observed behaviour of the (cognitive) system; a meaning that may not be unique (as different observers may attribute different meanings to the system's behaviour) and hence not necessarily a truly intrinsic property of the (cognitive) system's operation. Viewed in this light, the explanation of the (purported cognitive) system's operation remains ultimately formal. Hence, the theory lacks the explanatory power required to differentiate between cognitive and non-cognitive systems; it equally ascribes teleological behaviour to both. Similarly it cannot account for a cognitive system's phenomenology simply because any phenomenology ascribed to cognition is in reality a reflection of the external phenomenology of the observer (the "external observer fallacy"); a view consistent with the Searlian idea of the observer relativity of computational processes (Searle, 1990). This observation is valid for both FOC and computational approaches.

In addition to classical model of cognition's inherent inability to explain teleological behaviour, serial computation has recently lost its unique position as the only mode of computation. Recently, alternative views of computation have emerged encompassing different modes of concurrency and mutual interaction between the sub-systems, system and real world. Indeed, various authors have proposed a paradigm shift to such "interactive models of computing" (Wegner, 1997; Stein, 1999), which enable the description of systems whose operation falls outside the Turing machine framework. Thus, this begs the question why one would equate cognition with Turing equivalent computation.

A successful theory of cognition must account for fundamental properties of cognitive systems. At the very least these would seem to involve many coupled components simultaneously affecting each other, embedded in the world and interacting with the environment in real time giving rise to teleological behaviour.

From the earlier descriptions of SOC, ETP and DSC there emerges a unified view of the three theories which could account for the above characteristics.

In SOC, the observer and the environment are considered as one interacting system, encompassing genuine cognitive processes and in particular leading to teleological behaviour.

However, SOC is a meta-theory that forces us to explicitly account for cognitive processes without specifying the mechanisms that give rise to them. ETP explicitly defines cognitive processes as arising from the coupling of the observer and the environment fits perfectly within the SOC framework. Hence, we can consider it as a particular instantiation of SOC emphasizing the role of embodiment as the specific neuro-psychological mechanism essential for cognition.

Further, the natural way to formally describe the embodied interactions between cognitive processes and the environment is in terms of DS, in which the observer/system evolution is modelled by a series of coupled differential equations. The subsets of differential equations correspond to the subsystems being modelled; the couplings between the corresponding equations reflect the couplings between the subsystems. However, in contrast to DSC, which claims cognitive processes are the instantiations of specific state space trajectories within the agent, in the unified view cognition emerges from agent – environment interactions, and hence is not solely situated within the agent.

Hence by combining SOC, ETP and DSC there is the possibility of closing the phenomenological gap as such a combination does not lead to the external observer fallacy described earlier, yet accounts for the fundamental attributes of cognitive processes.

## 7. Evolutionary robotics

Dynamical approaches have also found favour in evolutionary robotics, particularly in the work of Harvey (1996, 2002). After Varela *et al.* (1991), Harvey considers an agent (human, animal, robot) as a perturbed complex dynamic system tightly coupled to its environment. However, designing controllers with these properties is a very difficult task. Unlike conventional computational robotics, such approaches are not amenable to traditional "divide and conquer" methodologies as, "the design of any one small component depends on an understanding of how it interacts in real time with the other components, such interaction possibly being mediated by the environment" (Harvey, 2002). Hence, Harvey (1996) uses evolutionary algorithms in order to achieve the required dynamic behaviours of the robots.

> *In this work a genetic encoding is set up such that an artificial genotype, typically a string of 0s and 1s, specifies a control system for a robot. This is visualised and implemented as a dynamical system acting in real time; different genotypes will specify different control systems. A genotype may additionally specify characteristics of the robot "body" and sensorimotor coupling with its environment. When we have settled on some particular encoding scheme, and we have some means of evaluating robots at the required task, we can apply artificial evolution to a population of genotypes over successive generations.*
>
> *Typically the initial population consists of a number of randomly generated genotypes, corresponding to randomly designed control systems. These are instantiated in a real robot one at a time, and the robot behaviour that results when placed in a test environment is observed and evaluated. [. . .]*

*The cycle of instantiation, evaluation, selection and reproduction then continues repeatedly, each time from a new population which should have improved over the average performance of its ancestors.*

Harvey's approach appears to be very much in line with the postulates of the SOC and ETP. This is because he explicitly embeds controllers in real robots and evolves them in response to real environmental pressures. This emphasises the embedding and coupling of the robot and the environment; the hallmarks of the integrated theory outlined above.

Harvey also asserts that his approach, which explicitly does not equate cognition with computation, renders arguments against machine intelligence based on Gödel's incompleteness theorems (Lucas, 1961; Penrose, 2002), "irrelevant". Hence, he claims that in principle his methodology can lead to the evolution of genuinely cognitive robots.

In order to address this position we deconstruct Harvey's position into a "strong" and a "weak" version. The strong position maintains that this methodology in principle can evolve any cognitive/conscious behaviour in a robot. The weaker claim is simply that the methodology can evolve at least some genuine cognitive/conscious behaviours.

We address the strong claim by reference to Penrose/Gödelian arguments against machine understanding. As Penrose has illustrated some aspects of cognition, (e.g. the aperiodic tiling decision problem), involve non-computational processes. However, Harvey has acknowledged (private conversation) that for convenience he often employs "cheats" by using computers plus essential clocking as the underlying DS. In this case all that is being achieved is the evolution of a formal/computational description of behaviour which is open to attack by Penrose style arguments, i.e. ". . . the powers of human reason could not be limited to any accepted preassigned system of formalised rules. What Gödel showed was how to transcend any such system of rules, so long as those rules could themselves be trusted" (Penrose, 1994).

Harvey's response to the above critique might be the weaker claim that he could achieve at least some genuine cognitive states within his robots, just not the full range of the cognitive powers of humans. However, it is apparent that in any artificial evolutionary system the critic that performs the evaluation function necessary to maintain the "selective pressure' is explicitly defined by the external observer/engineer; hence any teleological behaviour arises as a result of the external observer's teleology plus the built pre-ordained optimisation characteristics of evolutionary algorithms. Harvey might retort that he did not have to use his "cheat"; he could have used a genuine dynamical system. Nevertheless, our comments apply to any evolutionary algorithm irrespective of the underlying nature (computational emulation of or genuine dynamic) of the artificially evolved systems. All that is achieved by changing the nature of the systems from computational to real dynamic is a move between the computational and the FOC frameworks: as discussed earlier both are subject to the external observer fallacy.

Thus, although an interesting approach, which on the surface is generally in line with the integrated theory (SOC/ETP/DS), closer inspection reveals subtle differences that mean it does not fully conform to the unified theory.

## 8. Conclusions

We have argued herein that all computational (and FOC) approaches to cognition share the same fate: the external observer fallacy. We suggested that the best conceptual framework to avoid this fallacy and encompass the fundamental characteristics of cognitive systems is by integrating the ETP and DS in the general framework of SOC. This approach opens the possibility of giving an explanation of cognition that bridges the "explanatory gap" defined by Chalmers (1996).

Further, we critically review the evolutionary robotics paradigm, and conclude that, although it is an excellent way to build interesting robots, it is not fully consistent with the proposed integrated theory and hence does not escape the outlined critiques raised against FOC and computationalism.

Although we leave open the possibility that evolutionary robotics might 1 day help narrow the "explanatory gap", at present it would seem that, at best, such an approach is evolving "Computational Zombies".

### Notes

1. It is interesting to note that, in the domain of machine perception, a similar approach has also been explored in the field of robotics in the development of active vision systems (Blake and Yuille, 1992; Ballard *et al.*, 1997).
2. Although there is some evidence that there are brain processes that could perform something like "filling in", this does not mean that the brain actually does use them to "fill in" its putative internal representation of the world.

### References

Ballard, D., Hayhoe, M., Pook, P. and Rao, R. (1997), "Deictic codes for the embodiment of cognition", *Behavioural and Brain Sciences*, Vol. 20, pp. 723-67.

Blake, A. and Yuille, A. (1992), *Active Vision*, MIT Press, Cambridge, MA.

Chalmers, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, Oxford.

Damasio, A.R. (1996), "The somatic marker hypothesis and the possible functions of the prefrontal cortex", *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, Vol. 351 No. 1346, pp. 1413-20.

Harvey, I. (1996), "Untimed and misrepresented: connectionism and the computer metaphor", *AISB Quarterly*, Vol. 96, pp. 20-7.

Harvey, I. (2002), "Evolving robot consciousness: the easy problems and the rest I", in Fetzer, J.H. (Ed.), *Evolving Consciousness: Advances in Consciousness Research Series*, John Benjamins, Amsterdam.

Lucas, J.R. (1961), "Mind, machines and Gödel", *Philosophy*, Vol. 36, pp. 120-4.

Martin, L. (1986), "Visual localization and eye movements", in Boff, K.R., Kaufman, L. and Thomas, J.P. (Eds), *Handbook of Perception and Human Performance*, Wiley, New York, NY, Vol. 1, pp. 20.21-45.

O'Regan, J.K. (2001), "Experience is not something we feel but something we do: a principled way of explaining sensory phenomenology, with change blindness and other empirical consequences", talk given at Bressanone, 24 January.

O'Reagan, K. (2004), personal communication.

O'Reagan, J.K. and Noe, A. (2001), "A sensorimotor account of vision and visual consciousness", *Behavioural and Brain Sciences*, Vol. 24 No. 5, pp. 883-917.

Penrose, R. (1994), *Shadows of the Mind*, OUP, Oxford.

Penrose, R. (2002), "Consciousness, computation", in Preston, J. and Bishop, J.M. (Eds), *Views into the Chinese Room; New Essays on Searle and Artificial Intelligence*, Clarendon Press, Oxford.

Port, R.F. and Van Gelder, T. (1998), *Mind as Motion: Explorations in the Dynamics of Cognition*, MIT Press, Cambridge, MA.

Ryle, G. (1949), *The Concept of Mind*, Penguin, London.

Scott, B. (1996), "Second order cybernetics as cognitive methodology", *Systems Research*, Vol. 13 No. 3, pp. 393-406.

Searle, J.R. (1990), "Is the brain a digital computer?", *Proceedings of the American Philosophical Association*, Vol. 64, pp. 21-37.

Stein, L.A. (1999), "Challenging the computational metaphor; implications for how we think", *Cybernetics and Systems*, Vol. 30 No. 6, pp. 473-507.

Van Gelder, T. (1997), "Dynamics and cognition", *Mind Design II*, in Haugeland, J. (Ed) MIT Press, Cambridge, MA.

Varela, F., Thompson, E. and Rosch, E. (1991), *The Embodied Mind*, MIT Press, Cambridge, MA.

Von Foerster, H. (1974), "Notes pour un epistemologie des objets Vivants", in Morin, E. and Piatelli-Palmerini, M. (Eds), *L'Unité de L'Homme*, Editions du Seuil, Paris, pp. 401-17.

Von Glasersfeld, E. (1995), *Radical Constructivism: A Way of Knowing and Learning*, The Falmer Press, London and Washington.

Wegner, P. (1997), "Why interaction is more powerful than algorithms", *Communications of the ACM*, Vol. 40 No. 5, pp. 80-91.

Wheeler, M. (2002), "Change in the rules; computers, dynamical systems and searle", in Preston, J. and Bishop, J.M. (Eds), *Views into the Chinese Room; New Essays on Searle and Artificial Intelligence*, Clarendon Press, Oxford.