Active Perception in Meaningful Worlds*

Michael Wheeler
School of Cognitive and Computing Sciences,
University of Sussex, Brighton BN1 9QH, U.K.
`Telephone:+44 273 678524`
`Fax:+44 273 671320`
`E-Mail:michaelw@cogs.susx.ac.uk`

Abstract

This paper sketches a conceptual framework for thinking about perception. By identifying certain significant aspects of key empirical studies, I endorse and defend the increasingly influential view that perception should be conceptualized an an activity performed by an autonomous agent in the context of some adaptive behaviour. I then draw upon insights from ecological, dynamical and enactive views on the relationship between perceiving agents and their environments to develop a philosophical account of active perception. Finally I suggest how the mechanisms identified by the previously mentioned empirical work fit into the conceptual framework.

1 Introduction

An adequate account of perception is essential both to understanding the situated action achieved by certain naturally-occurring systems, and to recreating such behaviour in human-made machines. Moreover, how we conceptualize perception partly determines the overall theoretical framework in which we conceive and develop our empirical research into adaptive behaviour in animals and artefacts. In such a context, the goal of this paper is to present the beginnings of a philosophical theory of perception which is informed by, and continuous with, progress in the simulation of adaptive behaviour. There are similarities and connections between what I have to say and what others — from both inside and outside the adaptive behaviour community — have already said. I shall do my best to highlight the links as the argument unfolds.

2 Perception in Artificial Intelligence

As a way of opening up the relevant theoretical space, this section identifies assumptions made by the majority of recent attempts to investigate perception by way of artefacts. I shall keep my remarks brief, because the key points have been well-rehearsed elsewhere.[1]

In general, not only researchers in classical A.I., but also (most) connectionists have treated perceptual capacities as a series of courier systems. According to this view, an agent's perception-module constructs a representational description of the external world. This world-model is then delivered to a central system made up of sub-modules for specialized sub-problems such as reasoning and planning. These sub-modules manipulate the representations in accordance with certain computational algorithms, and then output a

---

* This paper is to be presented in poster form at 'Artificial Life IV,' Boston, Massachusetts, July 1994.

[1] See, for example, [3, 7, 15, 31].

further representational description (this time of the desired actions) to which the action-mechanisms then respond. Brooks refers to these principles as *functional decomposition* or the *sense-model-plan-act* framework [7, 8]. In mainstream A.I., perception became essentially a process of transduction followed by inference.

As a key example of a theory of perception embedded in the traditional sense-model-plan-act framework, consider Marr's account of vision [21]. In broad terms, Marr thought that the function of the visual system is to construct three-dimensional internal representations of objects in the visual field from informational clues available in the two-dimensional retinal image. The staged process moves via an intermediate viewer-centred representation of distance and orientation to an object-centred model of three dimensional shape which is independent of the viewer's observational perspective. This output from the vision-module is then delivered as input to the central cognitive modules which carry out the task of object-categorization.

So, by the account on offer from the functional decomposition camp, the cognitive role of perception is to recover, in the form of internal representations, the properties and relations of an essentially pregiven external environment [31]. Perception is veridical just when the representational description delivered to 'cognition-central' by the perception module accurately captures those properties and relations. On this view, perception and action can be analysed and studied independently of each other, because they are held to be temporally distinct and conceptually separate functions of an intelligent control-system [7]. But however influential such a set of explanatory principles has been, there are good reasons to think that they do not present the most profitable way to proceed. In short, it is a confusion to think that particular perceptual processes can be studied independently of the specific adaptive activities in which those processes are embedded.

3 Active Perception

Let's start by taking a look at visual perception. The view adopted in mainstream A.I. has resulted in most computational vision research being orientated towards static scene analysis, rather than ongoing activity in a world. Even those researchers who have employed more 'dynamic,' ecological concepts, such as optic flow, have done so with the aim of constructing more useful world-models for cognition-central, rather than in the context of specific environmentally embedded activities. But now consider John Haugeland's fridge [15].

> [What's] remarkable about our refrigerator aptitudes is not just, or even mainly, that we can visually identify what's there, but rather that we can, easily and reliably, reach around the milk and over the baked beans to lift out the orange juice — without spilling any of them (p.11).

This everyday example from human affairs provides an initial indication that it is a mistake to conceptualize perception as a process of recognition and identification. Moreover, it suggests that we should not be prepared to divorce perception from action, in the way implied by the imposed theoretical wedge of a central reasoning system, receiving representations from the perceptual systems and sending representations to the action systems. This rejection of what is, in essence, the mainstream-A.I. view should not be confused with the claim that there is never any interesting sense in which states of the nervous system intervene between sensory stimuli and bodily movement during situated activities. That is not the point. The claim I wish to endorse is that the way in which an organism engages in vision-based real-time interaction with an environment is not by taking what are, in effect, a series of sophisticated photographic snapshots, which can then be used by central modules in inferential, search-based processes of classification and categorization.[2] (If you find this characterization too crude, notice that the construction of elaborate data structures containing all the information necessary to recreate a visual scene is, in all respects relevant to the point at issue, equivalent to the snapshot-model [3].)

As something of a side-issue, we should remind ourselves that there are, of course, many different sensory modalities at work in the animal and robot kingdoms. Tactile sensors (such as whiskers or bumpers) are essentially proximal sensory mechanisms, extracting reliable information from only the immediate environmental surroundings. However, to achieve many complex behaviours (such as behaviours involving navigation strategies more sophisticated than something like wall-following), distal sensing is required [10]. Vision is probably the most studied of distal sensing capacities, and it is all too tempting to treat visual perception as the paradigmatic model for *all* distal perception. But it is certainly not obvious that the functional principles of different perceptual capacities (even within either the proximal or the distal classes) differ only in the ways in which the environmental data are encoded and extracted. There may be generalizations to be made across different modalities, but they have to be demonstrated, not assumed.

A recognition of the intimacy between perception and action has resulted in the concept of *active perception* becoming part of the conceptual armoury of researchers investigating adaptive behaviour in animals and artefacts. The aim of this section is to make plausible one version of the active perception thesis, i.e., the view that perception should be conceptualized an an activity performed by an *autonomous agent* in the context of some *adaptive behaviour.* To proceed, then, we need at least working definitions of the two key terms included in this initial characterization. (Here I merely make explicit views which, I take it, are adopted by the simulation of adaptive behaviour community in general.) An *autonomous agent* is as a fully integrated, self-controlling, adaptive system which, while in continuous long-term interaction with its environment, actively behaves so as to achieve certain goals. So for a system to be an autonomous agent, it must exhibit *adaptive behaviour*, behaviour which increases the chances that that system can survive in a noisy, dynamic, uncertain environment. We should identify a system as an adaptive system only in those cases where it is useful to attribute survival-based purpose and purposes to that system [27]. So rivers don't count as adaptive systems, but moths do. Naturally-occurring adaptive behaviour is the result of evolutionarily determined pressures on the survival and reproduction prospects of embodied creatures. Hence the class of naturally-occurring autonomous agents includes humans, non-human mammals, fish and insects. Artificial autonomous agents — henceforth 'animats' [35] — can be real autonomous robots with actual sensory-motor mechanisms, or simulated agents embedded in simulated environments.[3]

To sketch the nature and advantages of an active account of perception, I shall, over the next few sub-sections, help myself to the results of existing empirical studies, and focus on the following:

- real-time interaction with an environment,

---

[2] For explicit rejections of snapshot-vision, see, among others, [2, 3, 17, 31].

[3] As with most (all?) definitions of concepts, there are potential problem cases. By the definitions offered here, some plants might count not only as adaptive systems, but also as autonomous agents. I shall just stipulate that, in the context of this paper, the class of autonomous agents excludes plants. To me such a move is intuitively correct; but I accept that some may find it more than a a little arbitrary.

- the intimate interlinking of perception and action through close sensory-motor couplings between agent and environment,
- the role of the environment in explaining perceptual capacities,
- the use of specialized perception-action modules in robots, and the hypothesized existence of such modules in animals,
- the attainment of overall competence in certain perceptual abilities, even given far-from-optimal performance in relevant sub-competences,
- the part played in perception by sensors with dynamic properties,
- the nature of representation in active perception.

Action and Interaction

In a process of ongoing interaction with an environment, temporal constraints are hardly ever arbitrary. Lurking behind the sense-model-plan-act methodology is a crucial premiss to the effect that, even given accuracy problems resulting from noisy or drifting sensory-motor mechanisms, it is still *possible* to build an adequate, stored world-model, and to manipulate that model in real-time. This is required so that, for the purposes of planning action, operating in an actual world can be ignored in favour of the internal representations. But, as adaptive behaviour researchers (and others) have often observed, in scenarios where an autonomous agent's domain of activity is a dynamically changing and/or uncertain environment, a commitment to the necessity of maintaining an accurate internal world-model could well be a devastating error. Due to an explosion in the demands placed on representational and computational resources, the problem becomes intractable on the time-scales relevant to the realization of adaptive behaviour (a fact which would signal the untimely end of many a predator-threatened animal).

But are these difficulties products of the way in which perception and action are conceptualized in architectures committed to decomposition by function? Evidence that the answer to this question may well be "yes" can be gleaned from a consideration of animat control systems in which perception and action are intimately intertwined through the realization of tight feedback loops between the agent's sensory-motor mechanisms and the environment. For example, Franceschini *et al.* [12, 13] describe how real-time visual guidance of a path-finding autonomous mobile robot, demonstrating obstacle avoidance as one self-contained competence, was achieved through the speeds of its drive and steering motors being adjusted via purely local visual feedback loops. This close sensory-motor coupling enables the fully-developed robot to function in certain classes of environments in which it had not been debugged, and to succeed not only at path-finding tasks involving stationary spatially located goals (for which the hardwired control system was specifically designed) but also at similar tasks involving non-stationary goals (for which the hardwired control system was not specifically designed). This 'natural adaptibility' of the control system rests on the intimate link between perception and action. Indeed it seems that — within certain restrictions on how fast a moving goal is travelling in relation to the robot's sensory-motor capacities — *any* goal-finding architecture featuring close sensory-motor couplings will have a natural adaptibility to transfer from environments featuring a static goal to those featuring dynamic goals. (This adaptibility will fail if, in the static-goal case, the sensing mechanism has been picking up not properties of the goal, but properties of some other fixed object in sensory range — such as a window — which had a fixed relation to the stationary goal).[4]

Franceschini *et al.*'s robot (more on which below) is an example of a *behaviour-based* control architecture, as pioneered by Brooks and his colleagues.[5] The 'behaviour-based' approach advocates architectures with no central reasoning systems and no manipulable symbolic representations. Given the intuition that a situated agent should operate by continuously referring to its sensors as opposed to some internal representation, the process of attempting to build a centrally stored, objective world model is rejected as constituting a positive hindrance to real-time activity in a messy environment. Individual behaviour-producing systems, called 'layers', are designed to be individually capable of — and to be generally responsible for — connecting the robot's sensing and motor-activity in the context of, and in order to achieve, some ecologically relevant behaviour. So each layer is closely coupled to the robot's environment along what might be called a 'channel of ecological significance.' Starting with layers which achieve simpler behaviours (such as 'avoid objects' and 'explore'), layers are added, one at a time, to a debugged, working robot, so that overall behavioural competence increases incrementally. The layers run in parallel, affecting each other only by means of suppression or inhibition mechanisms.

Any account of perception based on tight sensory-motor couplings effectively makes a prediction that, in order to achieve ongoing perceptually guided activity, autonomous agents have to depend on the world to do its share of the work by providing not just surprises, but informational continuity.[6]

Specificity and Behavioural Success

Once we treat perception as an ecologically-embedded phenomenon, it seems that at least some mechanisms which we see as playing a particular role in perception are orientated towards certain specific adaptive behaviours. (It is definitional of behaviour-based layers that specific adaptive transitions from sensing to motor-behaviour — e.g. obstacle avoidance — are, in a sense, 'wired-in.') The presence of such systems would seem to mesh poorly with a view according to which the job of perception is to recover task-independent objective information from the world, and to deliver that information to a central system for the purposes of planning action.

To appreciate the potential applicability of this specificity claim to naturally-occurring control systems, consider Webb's robot-implementation of a hypothesized mechanism for cricket phonotaxis (the ability to track a conspecific auditory signal) [32]. Female crickets locate potential mates by tracking an auditory advertisement produced by the male. A cricket's control system has to be seen as a layered architecture of parallel behaviours. Each of these layers involves specialized links between sensing and motor-behaviour. This is clearly analogous to behaviour-based robot control architectures (Webb, 1993, p.1092). Phonotaxis is realized by one of these behaviour-producing layers. It is a specialized sensory-motor mechanism which is activated by the conspecific call of a potential mate. As Webb puts it (p.1092), "[There] is no need to process sounds in general, provided this sound has the right motor effects. Indeed, it may be advantageous to have such specificity

---

[4] Many thanks to Dave Cliff for discussion of this point.

[5] See, for example, [6, 20, 22].

[6] As Boden [5] reminds us, the observation that the best source of information about the world is the world itself occasionally surfaced in mainstream A.I., as did the thought that we should be wary of hallucinating world-models where close couplings are at work. Unfortunately (as Boden observes) these observations were no sooner made than forgotten by most A.I. researchers.

built in, because it implicitly provides ‘recognition’ of the correct signal through the failure of the system with any other signal.”

Moreover there is no need to hypothesize two separate processes — one to recognize the song, and one to determine the direction in which to move. This is fortunate, since the most obvious way for the cricket to ‘compute’ the direction to move — by comparing the intensity of the sound on each side — is seemingly unavailable. Crickets are too small, relative to the wavelength of the song, for there to be much in the way of intensity difference between the cricket’s ears. It appears that the following mechanism (hypothesized to exist in the cricket, and shown to function successfully in Webb’s robot) may solve the adaptive problem: the cricket’s ears are connected so that the movement of each tympanum is a cancellation of the sound directly outside, and the sound transferred from the opposite tympanum, the relative phase of which will depend on the way the cricket is orientated to the sound source.

The message is that having dedicated mechanisms that link sensing to motor-behaviour may be a highly effective and efficient tactic. Any such mechanism may be highly specific to some particular adaptive behaviour of high ecological significance, as in the case of cricket phonotaxis. Or it may negotiate a domain reflecting behavioural patterns which are grouped together — naturally — by the capacities of the system and the way it is embedded in its ecological niche. For instance, the same fundamental tracking mechanisms may well be used for tracking all sorts of animate entities of ecological significance, whether those entities constitute predators or potential mates. No doubt the way in which evolution works has ensured that the organization and flow of control will be opportunistic and far from optimal, but that’s life.

It would be a mistake for the unconvinced to claim that such ecological specificity could only possibly be found in the ‘simple’ control systems of creatures such as insects. Neurophysiological studies of monkeys have demonstrated the existence of specialized cortical areas for behaviours that use foveal motion and for behaviours that use peripheral motion [18]. And Ballard [3] argues that once we include in our theory of vision both the structure of the environment and the self-motion of the animate agent, ongoing visual perception in general (humans included) may well be achieved by a very large number of distinct special-purpose processes. These processes exploit different visual cues, and are not theoretically independent of the environmentally embedded behaviours in which those processes occur.

The embedding of perceptual capacities in an active situated agent also means that it would be a mistake to judge the success of the special-purpose mechanisms themselves according to any criterion of optimal performance *by that mechanism in its own sub-domain, abstracted from the overall behavioural success of the complete agent.* For instance, Horswill and Brooks [16] present evidence that a two-layer behaviour-based robot performing visual segmentation and motion tracking tasks can, as a complete agent, behave adequately (and adaptively) in an unconstrained, dynamic, real-world environment, whilst being rather poor at each of those tasks considered individually. The interactions between the two layers and between the agent and its environment suffice to secure overall behavioural adequacy.

Dynamic Sensing

Now I wish to consider two of the most fundamental yet sophisticated mechanisms for achieving visually guided behaviour — saccades (rapid relocations of gaze-direction occurring without visual feedback during their execution) and target-tracking. In humans, whose eyes, necks and bodies are all movable, saccades (which transfer the image of an object from the visual periphery to the fovea) and foveal tracking are largely oculomotor. However, insects (whose bodies are not so pliable) are generally compelled to move their whole bodies to achieve the same results. For instance, male hoverflies of the species *Syritta* have evolved foveal vision for tracking females. (Females do not track males, and so have not evolved foveal vision.) When the male locates a female, he turns so that she falls within his fovea. Analysis demonstrates that these turns are made without visual feedback, and are functionally equivalent to the saccades of the human visual system.[7]

Once a target-image has been transferred to the fovea, the job of the visuo-motor system is to keep it there. Human foveal tracking movements (at low velocities at least) are continuous and smooth, unlike the discrete, jerky movements which characterize saccades. *Syritta* also executes smooth tracking movements (again of the whole body) in order to perform foveal tracking. If an image moves away from the fovea, the human-visual system tracks it by a series of saccades. And, once again, there is evidence to suggest that *Syritta* executes equivalent saccadic movements.

Finally, consider the fact that humans are able to track a target that is moving sinsuoidally more accurately than one which is moving chaotically. One school of thought would no doubt treat this as an indication that certain visuo-motor behaviours are achieved by inference through computational calculation. But another possibility is that the relevant visual systems are 'anticipatory' rather than (strictly) predictive. To generalize to creatures in general, the idea is that an internal dynamical tracking system is tuned (through evolution and/or learning) to certain ecologically relevant dynamics of the creature's environment (e.g., the likely trajectories of certain animate objects). Thus the visuo-motor system, once activated, would 'expect,' as a product of its very nature, certain types of behaviour from those portions of the creature's environment with which that creature (or its ancestors) had had meaningful regular interaction. And when activated by objects which behave contrary to those dynamics, that system would be less likely to track those objects successfully. The only further assumption required is that most portions of a creature's environment generally behave in ways which can, in principle, be anticipated.[8]

So all this suggests that naturally evolved sensory mechanisms are not passive receivers and conveyers of environmental information, but dynamic processes embedded in the ecologically significant behaviours performed by active autonomous agents. Also notice that it is by examining such processes as they function *in their ecological context,* that it is possible to explain *why,* for example, there are senses in which the visuo-motor mechanisms of insects and humans are not dissimilar — despite the structural differences between the eyes of insects and humans. It is not merely a fact of evolutionary continuity, in the sense that the human visuo-motor system is an adaptation of previously existing mechanisms in simpler creatures. It is also that all animals which rely on vision have to solve similar

---

[7] All the observational evidence I present regarding saccades and tracking in hoverflies is drawn from Collett and Land [11]. The similarities between the visuo-motor behaviour of insects and humans are discussed in detail by Land [19]. Ballard [3] and Brooks and Stein [9] have stressed the role of processes such as saccadic eye-movements in achieving complex visually guided behaviour in robots. And this is one case where the same principles of operation can be applied to a radically different sensory modality. Predatory aquatic bugs make saccadic movements (in order to locate the source of surface ripples) using receptors in their legs [26].

[8] Miller and Freyd [25] present psychological evidence for the existence, in humans, of internal dynamical mechanisms tuned to the different properties of motion possessed by different classes of object. They also point out that protean (adaptively unpredictable) behaviour on the part of predator-threatened prey can be explained as adaptations to such anticipatory tracking mechanisms. If a predator's visuo-motor dynamics are entrained to the expected movements of some prey, any prey acting in a way which diverges from those anticipatory dynamics will secure an adaptive advantage.

adaptive tasks such as navigation and the tracking of other animate creatures [19]. These claims share a common theme which has run through my account of active perception, namely that those interested in endowing artefacts with perceptual capacities had better take notice of data from the biological sciences. For example, detailed information about the compound eye of the fly provided the basis for the visual system implemented on the Franceschini *et al.* robot; and, as we have already seen, the control structure of Webb's robot was closely based on that of the cricket. The biological sciences include not only neuroscience, but also ethology, behavioural ecology and evolutionary theory. Understanding how an autonomous agent is embedded in its ecological niche is important to explaining that agent's perceptual capacities. Moreover, by taking account of ecological embeddedness (as indeed both Franceschini *et al.* and Webb do), we run less risk of hallucinating mechanisms which are either unnecessary or impractical, given the resources provided by the agent and its environment.

Representations

In many cases, using a theory of active perception removes any need to postulate inner processes for which representational interpretations would be appropriate (e.g., Webb's cricket-robot). The internal states postulated by the active perception thesis are decentralized, non-manipulable, essentially *active* structures, used in the context of a specific behaviour. All of this is in obvious contrast to the all-purpose, task-independent, object-centred world-models favoured in mainstream A.I.. *At the very least* then it seems that if any notion of 'representation' makes sense, it must be one which has undergone a fundamental transformation. One possibility is the development of structures that are agent-centred (as opposed to object-centred) in that they are defined through the relations that ecologically significant environmental entities bear to the agent itself. For example, Franceschini *et al.* [12] implement a 'goal pursuit' layer (in parallel with the obstacle avoidance layer discussed earlier) which functions by constantly defining a robot-egocentric map of obstacles in polar coordinates, in relation to the instantaneous direction in which the robot is heading. Notice that this dynamical structure does not sacrifice the active perception insights regarding close environmental couplings and ecological embeddedness. The map is not an objective representation which is stored, recalled and updated; rather it is agent-centred and 'built on the fly.'

4 Meaningful Worlds

How far have we come? It seems that once we start to think about perception as an ecologically-embedded activity performed in the service of some adaptive behaviour, standard assumption after standard assumption about the nature of perception and the mechanisms realizing it has to be called into question. Far from a model according to which perception and action can be investigated separately, we have moved to a model in which it makes much more sense to think in terms of a unitary phenomenon, the *perception-action cycle.*[9] As Merleau-Ponty [24] observes,

> ...[when] the eye and the ear follow an animal in flight, it is impossible to say "which started first" in the exchange of stimuli and responses. Since all the movements of the organism are always conditioned by external influences,

---

[9] Variations on the theme of the perception-action cycle are provided by Merleau-Ponty [24], Arbib [2] and Varela *et al.* [31].

> one can, if one wishes, readily treat behaviour as an effect of the mileu. But in the same way, since all the stimulations which the organism receives have in turn been possible only by its preceding movements which have culminated in exposing the receptor organ to external influences, one could say that behaviour is the first cause of all the stimulations (p.13).

So with reference to some specific event, perception and action are inextricably intertwined. Moreover, how an animal (or animat) exploits some sensory capacity is determined by the adaptive behaviour in which that autonomous agent is engaged. But if the way in which we think about perception has to be overhauled, then I suggest that we need to find a theoretical language which is in harmony with (and does justice to) the insights and implications of the new account.

One possibility is to employ a modified version of Gibson's concept of an 'affordance' [14]. According to Gibson, invariants of structured stimulus information in ambient light specify such things as the way surfaces are laid out, whilst invariant combinations of invariants specify what Gibson calls the *affordances* of the environment — the meaningful possibilities for interactive behaviour that the terrain, elements, objects and animals in a creature's environment present to that creature. If a rock is of the appropriate size and density, it can be grasped and thrown by a primate, i.e., it *constitutes* a missile for that animal. This is part of its fundamental 'way-of-being' for primates. Affordances are the value-rich objects of ecological worlds which 'point both ways,' i.e., they are properties of the environment taken with reference to (the activity of) the observer. A set of affordances defines a specific environmental niche, i.e., an inherently meaningful ecological world. In short, on the ecological view, the relation between perceivers and their meaningful worlds is one of mutuality or cospecification, and visual perception is a matter not of recovery by way of intermediate psychological structures such as representations, but of 'direct information pickup' on the part of perceiving agents who are 'tuned' to invariants in the structured ambient light. Gibson argues that the process of information pickup requires the concept of a perceptual system. For example, the eye is a perceptual organ that is one of a pair of mobile eyes, set in a head that can turn; and that head is part of a body that can move around in the environment. The whole eye-head-brain-body system is the perceptual system. There are five perceptual systems corresponding to five modes of active attention, namely looking, listening, touching, tasting, and smelling. Unlike sensory receptors, perceptual systems are not stimulated; they are activated in the presence of stimulus information.

The ecological approach, with its stress on perception as an active process of seeking environmental information, should be attractive to the active perception researcher. However, there are problems. The Gibsonian commitment is to a form of direct realism about the way perceivers register environmentally embedded information. The bodies of information available to the various perceptual systems are 'out there' in the world, waiting to be picked up. So understanding, say, visual perception becomes a matter of specifying the invariant structures in ambient light to which the visual system is tuned. The mechanisms underpinning the cospecification of organism and environment are presumably supposed to be explained by way of this 'attunement.' But, without further argument, a strategy in which environmental information can be specified independently of the organism (and to which the organism is attuned) does not appear to secure any 'deep' mutuality of organism and environment, perceiver and world. And if, as suggested earlier, we wish to reject the idea that perception is a process of inferential recovery of a pregiven world, then it is precisely that deep notion of mutuality that would constitute a powerful way of securing that rejection.[10]

Before confronting that issue of cospecification head-on, we need to place the theory of affordances in the adaptive behaviour framework assumed throughout this paper. So the intrinsic 'good' or 'ill' of an affordance would generally need to be 'cashed out' in terms of positive or negative survival and reproduction prospects for the agent. Consider an adaptive problem such as catching some prey. What will be detectable (via sensory information) will be environmental objects, events, or situations which Miller and Freyd [25] call 'fitness affordances.' Just like their Gibsonian ancestors, Miller and Freyd's evolutionarily defined fitness affordances 'point both ways.' A prey is only a prey with respect to some predator. It may, itself, be a predator to some other creature on which *it* preys. Consequently, fitness affordances are irreducibly cospecificationary, whilst being perceiver-independent in the sense that "fitness effects are imposed by natural selection whether the organism likes it or not; they cannot be eliminated through subjective denial or wishful thinking" (Miller and Freyd, 1993, p.16). This modification, whilst significant, is clearly in the spirit of Gibson's approach.

But now how can we explicate 'organism-environment cospecification'? Here we can turn to a different conceptual language — that of dynamical systems theory[11] — to provide a way of conceptualizing organism-environment relations in which this somewhat vague concept becomes much more concrete.

A nervous system is a complex dynamical neural network that constitutes the basis of the control system for a situated agent. (In general, the agent's sensory-motor mechanisms should be thought of as part of its control system.) This control system is embedded in a continually changing physical medium, with which it interacts in such a way that the network's intrinsic dynamics are regularly perturbed as a result of the agent's sensory-motor activity. The agent's control system and the medium in which that control system operates can be conceptualized as two *coupled* dynamical systems. Two theoretically separable dynamical systems are said to be coupled when they are bound together in a mathematically describable way, such that, at any particular moment, the state of either system fixes the dynamics of the other system; that is, each system fixes the principles governing change in the other system. In formal terms, this means that some of the parameters of each system either become, or become functions of, some of the state variables of the other.[12]

The situation in which one system fixes the dynamics of another system through coupling is not to be equated with a relation according to which one system specifies the state of a second. Each system biases the intrinsic possibilities for change already present in the other. The relation is one of influence of dynamics rather than specification of static state. If we begin by thinking of an animal nervous system as a non-coupled dynamical system, then we can conceptualize its intrinsic dynamics as generating a space of possible perturbations which the system can undergo as a result of coupling to a physical medium. Then, through sensory-motor activity, the dynamics of an animal's nervous system are continually perturbed in accordance with the adaptive couplings 'discovered' by evolution.[13]

In effect, we have now explicated the mechanism of attunement. It is the discovery, by Darwinian natural selection, of adaptive sensory-motor couplings. This evolutionary pro-

---

[10] Varela *et al.* make similar criticisms of the Gibsonian framework (Varela *et al.,* 1991, pp.202-5).

[11] Abraham and Shaw offer a friendly, but thorough, introduction to dynamical systems theory [1].

[12] For various dynamical systems approaches, see Beer [4], Yamauchi and Beer [36], van Gelder [30], Husbands *et al.* [17], Smithers [28, 29], and Wheeler [33, 34].

[13] These points about the relations between nervous system and world are influenced, in particular, by the work of Maturana [23] and Varela *et al.* [31].

cess operates on the nervous system of the organism. So, given the sense of 'information' relevant to the Gibsonian framework, the mechanisms resulting in attunement are not informationally sensitive, and so cannot be tuned to informational properties. This may look as if we are further away from organism-environment cospecification. But notice that our consideration of adaptive attunement has confined itself to a level of explanation characterized by physical processes of stimulus and response, occurring in the networks making up the sensory-motor and nervous systems. That is not the place to locate meaningful ecological 'objects' such as fitness affordances. So where should one look?[14]

The dynamical systems perspective provides a framework in which the dividing line between the 'internal' states of the agent and the 'external' states of the environment is something to be imposed in context, relative to the interests of the observer. This is because whilst it is useful, under certain circumstances, to think of *agent and environment,* (and not merely sensory-motor control system and physical medium) as separate, but coupled, dynamical systems, it is equally valid to redescribe the coupled agent-environment system as one larger dynamical system, in which the observed patterns of interaction between the agent-system and the environment-system are properties of that larger system [4]. In fact, it is perfectly legitimate to think of the one agent-environment system as *primary,* and to consider any analysis in terms of coupled dynamical systems as an *abstraction* that may not always be the best method of understanding what is going on, and, in some cases, may not even be practically possible.[15]

Now we have a way of cashing out the mutuality of organism and environment — the cospecificationary nature of fitness affordances. If we think of agent and environment as separate but coupled dynamical systems, with interactive causal commerce, then we can say that, through natural selection, evolutionary processes have tuned the agent's sensory-motor mechanisms to certain patterns of stimuli. *At the ecological level,* we have the activity of perception. Meaningful ecological properties (fitness affordances) are *brought forth* on the basis of those adaptive sensory-motor couplings (cf. the enactive approach to cognitive science [31]). From this perspective, fitness affordances can be seen as embedded properties of the single agent-environment dynamical system, thus ensuring agent-environment mutuality. So the inherently meaningful nature of ecological worlds is explained by this enactive process, a process which takes place on the basis of the dynamical coupling of a sensory-motor control system and a physical medium.

Mechanisms and Meanings

The broad features of the proposed philosophical framework for conceptualizing perception are now in place. The development of the framework was guided, initially, by work on the simulation of adaptive behaviour, and by studies in the biological sciences. But some might feel that the philosophical framework has become somewhat autonomous from the ongoing practical investigations of active perception. In an attempt to allay these fears, I shall close this paper by highlighting just a few ways in which the empirical work (described in section 3) and the philosophical suggestions fit together to suggest an overall way of thinking.

Consider the use of specific sensory-motor couplings along channels of ecological significance. The physical stimuli to which, say, individual behaviour-based layers respond are

---

[14] To prevent any misunderstandings, I should point out that this stress on the role of evolution does not ignore the existence of lifetime learning abilities in some animals, or downplay the importance of that capacity. What is implied is that learning amounts to a process of modifying pre-existing evolutionarily determined couplings.

[15] Many thanks to Phil Husbands for discussion of this point.

not, in themselves, meaningful properties of the agent's world. The way in which the complete agent, embedded in its environment, proceeds actively to exploit those stimuli brings forth meaning. It is this enactive process that *makes* a particular input-output transition a channel of ecological significance. But how an agent exploits patterns of stimuli is, of course, just the cumulative result of the ongoing operations of sets of adaptive sensory-motor couplings, which have either survived the process of natural selection, or have been developed by the human designer. Hence it is on the basis of the way in which the female cricket's nervous system is tuned to the auditory stimulus produced by the male, that the male's auditory output becomes a call-to-mate, a meaningful property (fitness affordance) of the female cricket's ecological world. This slots in nicely alongside the conceptual claim that the perception-action cycle is grounded in — as opposed to the same process as — the sensory-motor coupling of control systems to patterns of physical stimuli. And notice that Webb's cricket-robot neatly sidesteps the potentially serious problem of how the cricket might compare sound-intensities, precisely by exploiting the physics of its (the robot's/the cricket's) body and its physical environment.

In conceptualizing the agent control system as a complex dynamical system with its own intrinsic ongoing dynamics, we also avoid a possible misunderstanding about the nature of the ecologically-significant couplings underpinning the perception-action cycle. It would, of course, be ludicrous to suggest that all perceptually guided activity requires only *reactive* architectures, if 'reactive' is taken to mean 'no *significant* internal state transitions are required in the agent-system.' Whilst such reactive couplings may well be used to great effect in achieving certain simple behaviours (such as straightforward obstacle avoidance), the question in most cases of interest will be not "do we need internal state?" but rather "what sorts of internal states are required to achieve this particular behaviour?" (Cf. Beer [4] and Yamauchi and Beer [36] on the role of internal state in dynamical neural networks.) However, the need for internal state is not equivalent to the need for semantically interpretable internal structures such as symbolic (or, come to that, sub-symbolic) representations. By starting from a perspective according to which meanings (fitness affordances) are enacted properties of the agent-environment dynamical system, we are less likely to be tempted into assuming that every meaningful feature or act has, in some way, to be encoded in agent-internal 'psychological' or 'cognitive' structures. Thus it would be courting explanatory disaster to start out by assuming that the fundamental role of the agent-internal mechanisms underpinning adaptive behaviour is to encode for meaning *in anything but the most unrestrained sense of the word 'encode.'* It seems that this implication of the adopted philosophical framework reflects the commitments of much of the empirical work described in section 3 of this paper.

5 Summary

By rejecting the view according to which perception and action can be studied independently, we open the door to an account in which perception is conceptualized as an activity performed by an autonomous agent in the context of some adaptive behaviour. Evidence from adaptive behaviour research and the biological sciences supports this thesis of active perception. But if perception is actively situated, we need a a conceptual framework in which to place the intuitions and implications of the approach. I have endeavoured to describe the form of such a framework. Fitness affordances are brought forth on the basis of the sensory-motor coupling between the agent's control-system and the stimuli to which it reponds. They are properties of the ecological level (single agent-environment system). This level of explanation is the world of meaningful interactions, a world which arises out of sensory-motor activity. Active perception takes place in meaningful worlds.

Acknowledgments

This work was supported by British Academy major award no. 92/1693. Many thanks to Maggie Boden, Matthew Elton and Inman Harvey for their valuable comments on an earlier version of this paper.

References

[1] R. H. Abraham and C. D. Shaw. *Dynamics - The Geometry of Behavior 2nd edition.* Addison-Wesley, Redwood City, California, 1992.

[2] M. A. Arbib. *The Metaphorical Brain 2: Neural Networks and Beyond.* John Wiley and Sons, New York, 1989.

[3] D. H. Ballard. Animate vision. *Artificial Intelligence*, 48:57–86, 1991.

[4] R. Beer. A dynamical systems perspective on autonomous agents. Technical Report 92-11, Case Western Reserve University, Cleveland, Ohio, 1992.

[5] M. A. Boden. Autonomy and artificiality. To be published in *A.I.S.B. Quarterly* and D. Cliff (ed.) *Evolutionary Robotics and Artificial Life* (provisional title), forthcoming.

[6] R. Brooks. A robust layered control system for a mobile robot. A.i. memo 864, M.I.T. A.I. Lab, 1985.

[7] R. Brooks. Intelligence without reason. In *Proceedings of the Twelth International Joint Conference on Artificial Intelligence*, pages 569–95, San Mateo, California, 1991. Morgan Kauffman.

[8] R. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–59, 1991.

[9] R. A. Brooks and L. A. Stein. Building brains for bodies. Artificial Intelligence Lab Report 1439, M.I.T., 1993.

[10] D. Cliff, I. Harvey, and P. Husbands. Explorations in evolutionary robotics. *Adaptive Behavior*, 2:73–110, 1993.

[11] T. S. Collett and M. F. Land. Visual control of flight behaviour in the hoverfly, *syritta pipiens l . Journal of Comparative Physiology*, 99:1–66, 1975.

[12] N. Franceschini, J-M. Pichon, and C. Blanes. Real time visuomotor control: from flies to robots. In *Proceedings of: International Conference on Advanced Robotics, Pisa*, 1991.

[13] N. Franceschini, J. M. Pichon, and C. Blanes. From insect vision to robot vision. *Philosophical Transactions of the Royal Society, series B*, 337:283–94, 1992.

[14] J. J. Gibson. *The Ecological Approach to Visual Perception.* Houghton Mifflin, Boston, 1979.

[15] J. Haugeland. Mind embodied and embedded. Draft of paper to appear in the proceedings of 'Mind and Cognition: an International Symposium'. Institute of European and American Studies, Academia Sinica, Taipei, Taiwan., 1993.

[16] I. D. Horswill and R. A. Brooks. Situated vision in a dynamic world: Chasing objects. In *Proceedings of the Seventh Annual Meeting of the American Association for Artificial Intelligence*, pages 796–800, St. Paul Minnesota, 1988.

[17] P. Husbands, I. Harvey, and D. Cliff. Circle in the round: State space attractors for evolved sighted robots. *Proceedings of the N.A.T.O. Advanced Study Institute Workshop on the Biology and Technology of Intelligent Autonomous Agents*, forthcoming.

[18] H. Komatsu and R. H. Wurtz. Relation of cortical areas mt and mst to pursuit eye movements iii: interaction with full-field stimulation. *Journal of Neurophysiology*, 60:621–44, 1988.

[19] M. F. Land. Similarities in the visual behavior of arthropods and men. In M. S. Gazzaniga and C. Blakemore, editors, *Handbook of Psychobiology*, chapter 2, pages 49–72. Academic Press, New York, 1975.

[20] P. Maes and R. Brooks. Learning to coordinate behaviors. In *Proceedings of A.A.A.I.-90*, pages 796–802, Boston, Massachusetts, 1990.

[21] D. C. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman and Company, New York, 1982.

[22] M. Mataric. Navigating with a rat brain: a neurobiologically inspired model for robot spatial representation. In J.-A. Meyer and S. Wilson, editors, *From Animals to Animats: Proc. of the First International Conference on Simulation of Adaptive Behavior*, Cambridges, Massachusetts, 1991. M.I.T. Press / Bradford Books.

[23] H. R. Maturana. Biology of cognition. In *Autopoiesis and Cognition: the Realization of the Living.* Reidel, Dordrecht, 1970. Published in 1980, the book contains the piece by Maturana (dating from 1970) and a piece by F. Varela.

[24] M. Merleau-Ponty. *The Structure of Behaviour.* Beacon Press, Boston, 1963. Translation by Alden Fisher.

[25] G. F. Miller and J. J. Freyd. Dynamic mental representations of animate motion: The interplay among evolutionary, cognitive, and behavioral dynamics. Cognitive Science Research Paper 290, University of Sussex, 1993.

[26] R. K. Murphey. Localization of receptors controlling orientation to prey by the back swimmer, *notonecta undulata. Journal of Comparative Physiology*, 84:19–30, 1973.

[27] O. G. Selfridge. Some themes and primitives in ill-defined systems. In O. G. Selfridge, E. L. Rissland, and M. A. Arbib, editors, *Adaptive Control of Ill-Defined Systems*, NATO conference Series, chapter 3, pages 21–6. Plenum, New York, 1984.

[28] T. Smithers. Taking eliminative materialism seriously: a methodology for autonomous systems research. In F. J. Varela and P. Bourgnine, editors, *Toward a Practice of Autonomous Agents: Proc. 1st European Conference on Artificial Life*, pages 31–40, Cambridge, Massachusetts, 1992. MIT Press / A Bradford Book.

[29] T. Smithers. On behaviour as dissipative structures in agent-environment system interaction spaces. Faculty of Information, University of San Sebastian, 1993.

[30] T. J. van Gelder. What might cognition be if not computation? Technical Report 75, Indiana University, Cognitive Sciences, 1992.

[31] F. J. Varela, E. Thompson, and E. Rosch. *The Embodied Mind: Cognitive Science and Human Experience.* MIT Press, Cambridge, Massachusetts, and London, England, 1991.

[32] B. Webb. Modeling biological behaviour or 'dumb animals and stupid robots'. In *Proceedings of the Second European Conference on Artificial Life*, pages 1090–103, 1993.

[33] M. Wheeler. For whom the bell tolls: The roles of representation and computation in the study of situated agents. Cognitive Science Research Paper 320, University of Sussex, 1994.

[34] M. Wheeler. From activation to activity: Representation, computation, and the dynamics of neural network control systems. *A.I.S.B. Quarterly*, 87:36–42, 1994.

[35] S. W. Wilson. Knowledge growth in an artificial animal. In J. J. Grefenstette, editor, *Proceedings of an International Conference on Genetic Algorithms and their Applications*, pages 16–23, Pittsburg, PA and Hillsdale, New Jersey, 1985. Lawrence Erlbaum Associates.

[36] B. M. Yamauchi and R. D. Beer. Sequential behavior and learning in evolved dynamical neural networks. Available as Case Western Reserve University Tech. Report no. CES-93-25. Cleveland, Ohio, Submitted to *Adaptive Behavior*.