Misrepresentations

Inman Harvey

Evolutionary and Adaptive Systems Group
Centre for Computational Neuroscience and Robotics
University of Sussex
Brighton, UK
inmanh@sussex.ac.uk

Abstract

The concept of “representations”, and particularly “internal representations”, can be controversial in Cognitive Science and AI. It is suggested here that much time-wasting confusion could be avoided if participants in such controversies came to recognize the variety of different senses, often incompatible, in which such terms are used. A hypothesis is presented as to why there is so much reluctance to recognize this. Once such fruitless controversies are swept aside through linguistic hygiene, there remain interesting real problems, which are eminently appropriate for being tackled by an Artificial Life methodology.

Introduction

There are many confusions and misunderstandings associated with the term "representation" in Cognitive Science and AI, and by extension in Artificial Life. It will be argued here that most of these problems can be solved (or dissolved) by careful linguistic hygiene. But there remain interesting and genuine problems that can be fruitfully approached using an artificial life, evolutionary robotics methodology.

Representation wars

Artificial life overlaps with AI, in that both tackle the problems involved in synthesising lifelike capabilities; there may be different emphases, perhaps on adaptive behaviour versus rational thought. AI overlaps with cognitive science. All of these are permeated by the positions researchers may take on philosophical issues: what is life, what is cognition, what is mind? Traditional GOFAI (Good Old Fashioned AI) approaches to these questions have often framed answers in terms of "representations", or "internal representations". These sort of notions made no sense to me, working in a GOFAI department, and some of our early artificial life experiments at the beginning of the 90s had as one motivation the intent to make such issues explicit (Cliff et al 1993; Harvey et al 1993). Using evolutionary robotics techniques, we evolved simple minimally cognitive agents to perform simple tasks, and then challenged the GOFAI theorists to try and identify just where these so-called "internal representations" were.

We used genetic algorithms to evolve the connectivity, the connection weights, and the temporal parameters for real-time recurrent artificial neural networks, that formed the "nervous systems" for agents interacting dynamically with their environment. This is within the Dynamical Systems approach to understanding cognition, and our motivation for this continuing line of research at Sussex (Harvey et al 1997, Harvey et al 2005) is very much in sympathy with similar research by Beer (Beer 1995, Beer 2000) and others. What were the responses of the GOFAI theorists?

Some of them claimed that there simply must be internal representations in these agents, though there was little agreement on just where to find them. Others claimed that maybe for the simple cognitive tasks internal representations were not necessary, but for more complex forms of cognition — "representation-hungry tasks" — they would be essential. Some claimed it was logically necessary for brains (or minds) to have representations; others claimed it was a pragmatic necessity — they simply could not imagine how one could design brains to work in any other way.

It became clear that there were many conflicting notions of representation being bandied about. I was prepared to offer my definition of the term (Harvey 1996), but to my surprise it was extremely difficult, if not impossible, to pin down the other protagonists in these representation wars to offer their own definition of the term. In the way that I use the term, I have never ever had any internal representation of any kind in my head — except in the most casual, metaphorical sense. When I say I have a map of Brighton in my head, this is shorthand for saying I can navigate as if I had a map in my hand, I can visualise the configuration of a map as if it were in front of me — but I certainly do not mean that there literally is a map in my head. Others disagreed; but what was the nature of this disagreement?

The lack of communication became so dispiriting that eventually it seemed to me a waste of time continuing these debates. It is perfectly possible to do artificial life style research into cognitive systems, from an enactive or Dynamical Systems perspective, without ever having to engage in such discussions on representations. However it

seems that many people are still being sucked into the same old futile confusions (Grush, 1997; Clark and Grush, 1999; Grush, 2004; Wheeler, 2005; Rowlands 2006; Gallagher, forthcoming), so that here we revisit the representation wars yet again.

The position presented here is that there ***are*** important interesting issues concerning representations that can be usefully tackled with an artificial life, evolutionary robotics approach. But that almost all the debate I see on this topic has nothing to do with such issues, but is rather symptomatic of confused and unclear talk, specifically: multiple incompatible usages for the term "representation".

The Pavement Problem

Imagine an international conference of urban traffic specialists. They meet for extended discussions on the question of what restrictions are desirable, for safety reasons, to place on bicycles riding on the pavement. The discussions are intense, confused and chaotic; there is a complete lack of agreement. The participants returned home demoralised, and astonished at the obtuseness of many of the suggestions that they had heard. It turns out that many of those present were not aware that "pavement" means the part of the street that cars drive on in North American English, and the side of the street (or sidewalk) where pedestrians walk in British English. Even worse, it turns out that some present actually *were* aware of this crucial ambiguity, but did not feel it necessary to comment that this lay at the root of all the confusion.

Fortunately — we hope — urban traffic experts are not so stupid. Their work is important and can save lives; they need to think and explain themselves clearly. What a pity that so many philosophers do not live up to these basic standards.

When people use the term "representation" in the context of cognitive science, it is not nearly so simple as having merely one or two well-defined meanings. There is a constellation of different, overlapping academic interests, from neuroscience to philosophy of mind, each with their own different perspectives on the term. Furthermore, the richness of the variety of usages of the term "representation" means that it often has multiple, incompatible referents within any one such academic discipline. In the rest of this paper I shall categorise some of these usages; and then I shall speculatively offer a possible explanation why people are so often so very reluctant to clarify which of these many possible senses they are using.

The Everyday Sense of the Word Representation

Before going into any technical sense in which the word might be used, let us look at its everyday meaning. Consulting a dictionary (www.dictionary.net, based on Websters) for the term "represent" (since "representations" include 'the act of representing, in any sense of the verb') we find a list of 8 variants of increasing sophistication. I shall return, below, to the last two listed, but can summarise briefly here the first six, as everyday meanings: to represent is to present again or anew, to present by means of something standing in the place of, to exhibit the counterpart or image of, to typify. This extends to serving as a sign or symbol of — words represent ideas or things.

In this sense, it seems to me that human beings are supremely representation users, ever since the dawn of cave art and of language. Our use of representations is, above all, what differentiates us from other animals; we can argue about grey areas such as chimpanzees' use of sign language, or the sexual displays of peacocks, but our human usage of representations is orders of magnitude more complex and comprehensive. We humans live in language and culture, as a fish lives in the water.

What can we do, now that we have this useful trick of representing? If we want to draw somebody's attention to a cat, instead of bringing a cat out of one's pocket and waving it around, one can draw a picture of a cat, or even more conveniently use the word "cat". This is a sophisticated and incredibly useful trick, so useful that we spend much of our childhood learning how to use and extend our capacities for representation. At the most sophisticated end of the spectrum, we can reason with mathematical symbols, and write computer programs so that machines can reason for us.

Representation is a relational term

In this basic sense, representation is a relational term like North or Twin. 'Brighton is to the north' is ambiguous without a context — north in relation to where? Brighton is north of Paris and south of London, so a disagreement as to whether it is or is not "to the north" can sometimes be easily settled by establishing this context. Likewise, twin is a relational concept. Any number of exhaustive tests on a child cannot settle the question of whether that child is or is not a twin. Twinness can refer only to its relationship with a second child. Very often the context — the relational partners to these relational terms — are implicit and obvious, leaving no scope for disagreement. In the representation wars, across different disciplines and across different sets of starting assumptions, it is crucial to recognize that there simply is no such universal agreement on the context.

A symbol **P** is used by a person **Q** to represent, or refer to, an object **R** to a person **S**. Nothing can be referred to without somebody to do the referring. Normally **Q** and **S** are members of a community that have come to agree on their symbolic usages, and training as a mathematician involves learning the practices of such a community. The vocabulary of symbols can be extended by defining them in terms of already-recognised symbols.

The English language and the French language are systems of symbols used by people of different language communities for communicating about their worlds, with their similarities and their different nuances and clichés. The languages themselves have developed over thousands of years, and the induction of each child into the use of its native language occupies a major slice of its early years. The fact that, nearly all the time we are talking English, we are doing so to an English-speaker (including when we talk to ourselves), makes it usually an

unnecessary platitude to explicitly draw attention to the community that speaker and hearer belong to.

Since symbols and representation stand firmly in the linguistic domain, another attribute they possess is that of some element of arbitrariness (from the perspective of an observer external to the communicators). When I raise my forefinger with its back to you, and repeatedly bend the tip towards me, the chances are that you will interpret this as `come here'. This particular European and American sign is just as arbitrary as the Turkish equivalent of placing the hand horizontally facing down, and flapping it downwards. Different actions or entities can represent the same meaning to different communities; and the same action or entity can represent different things to different communities.

In the more general case, and particularly in the field of connectionism and cognitive science, when talking of representation (in the sense outlined above) it is imperative to make clear who the users of the representation are. In particular it should be noted that where one and the same entity can represent different things to different observers, conceptual confusion can easily arise. When in doubt, always make explicit the **Q** and **S** when **P** is used by **Q** to represent **R** to **S**.

Of course, it is open for people to choose to use the word in some different, technical, sense that may not fit into this format; but then it is obligatory to make clear just how this different sense is defined.

The Homuncular Representation

We shall return to the more sophisticated meanings listed for "represent" in the dictionary further on, but let us first explore common extensions of the basic meanings. When we try and explain or describe complex systems, to other people or to ourselves, then it is standard practice, good common sense, to draw on metaphors from everyday life. We often represent the component parts of the mechanism in terms of homunculi, or little imaginary people, who are performing different functions in coordination with each other. The thermostat measures the temperature, and tells the central heating boiler when it should switch on — for many purposes this shorthand, that personifies (or "homuncularises") the different components, is so much clearer and more useful than any detailed mechanical description. Notice how there are two levels of representation going on here: at one level we are representing the thermostat (*to the reader*) as, in effect, a little homunculus; and at another level the signal travelling down the wire from the thermostat represents the temperature or the command that the thermostat-homunculus "intends" (metaphorically) to convey *to the boiler-homunculus*. Here, for clarification, I have *italicized* the different "recipients" **S** for the two different instances of representation.

Now what are the requirements for this signal to represent the temperature? Firstly, there must be some correspondence or covariance between variations in temperature and variations in signal. Secondly, the signal must play some functional role, in communicating with whatever plays the role of a receiver of the signal. There is, in my view, a third requirement that may be more controversial. There needs also to be a further metaphorical homunculus that acts as the sender of the signal. The metaphor requires one homunculus communicating with another one.

When we follow somebody on a walk in the countryside, they could leave an indication of which route to take at a fork in the path by drawing an arrow; this is a representation of the desired direction. Alternatively, we could just trace their footprints in the mud; but we would not call these a representation, or at any rate not in the same sense. Likewise, I would suggest, the homunculi metaphor implicitly requires the active connivance of both sender and receiver — the arrow counts, but the footprints do not.

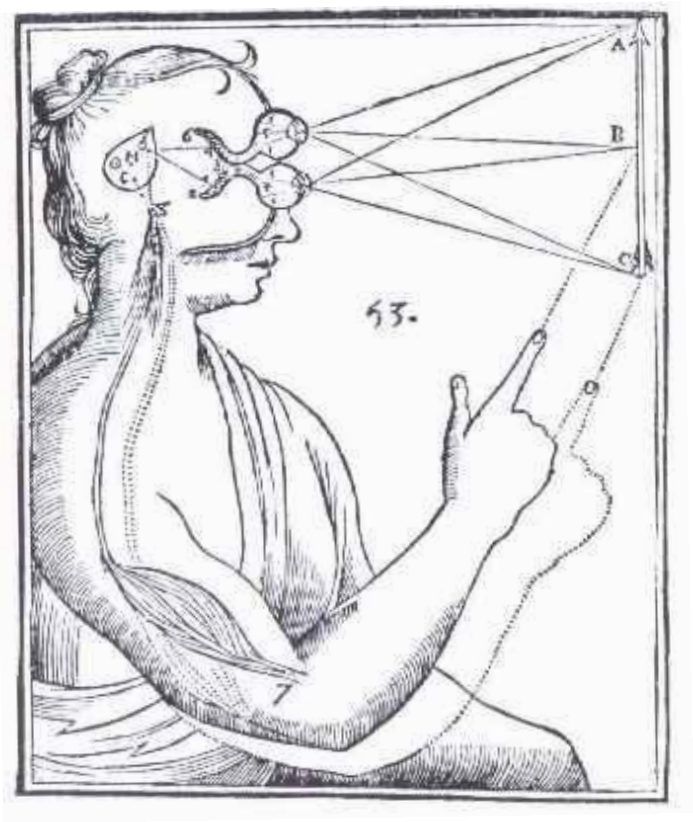
Not everyone will agree with this third requirement. Fair enough, but then you are using the word in a different sense from that which I propose. We should acknowledge and recognize that there are these different senses of the word representation, we should make explicit which sense we intend if we are to avoid confusion.

Is this signal travelling down the wire from the thermostat ***really*** a representation? Well yes, in this sense: the coin that I use on my chessboard to replace a lost piece ***really*** does represent the black king. Within the rules of the game it plays this role. When we used the homuncular metaphor there are these rules for the game.

I should stress that when I call this form of explanation "metaphorical" and "homuncular" this is certainly not intended to be disparaging, nor to mean that such explanations are any sense illegitimate or second rate. They are legitimate and invaluable forms of explanation, absolutely essential when we are talking in terms of functions, and we simply could not attempt to start to understand complex systems without employing them. We explain the strange and unfamiliar in terms of the familiar, and what could be more familiar than examples of inter-personal communication from our everyday life. But we should be aware of the baggage that this form of explanation carries with it.

More sophisticated senses of the term representation

Returning to definitions of "represent" in the online dictionary, items seven and eight were "to bring a sensation of into the mind or sensorium", and (Metaphorically) "to form or image again in consciousness, as an object of cognition or apprehension (something which was originally apprehended by direct presentation". Now personally I have trouble with these definitions. Apart from any other considerations, sensorium is not in my everyday vocabulary, and I have to go and look that up. My suspicion is that these definitions carry with them a certain baggage, certain philosophical assumptions that I may not share. For the purposes of this paper, however, I need only comment that these senses either do, or do not, share the basic properties outlined above. If they do share these properties, then to avoid confusion we



From Descartes (1662), *De Homine.* Visual information travels to the brain through hollow optic nerves. It continues on to the Pineal body, that regulates the flow of animal spirits to the nerves, and thence the muscles.

must be willing to spell out the **P**, **Q**, **R**, **S**. If they do not share these properties then we must acknowledge that these are different senses of the term.

As an aside on that final definition of represent above, we can notice the differences and similarities between the propositions "I imagined seeing a car" and "I saw an imaginary car". Superficially they appear to say the same thing, but the second version carries overtones of asserting the existence of "an imaginary car". With this second version, we might be more tempted to ask just where this imaginary car is. If it is not on the street, then where is it — in the consciousness, in the mind, in the brain? I would like to suggest that asking these questions is as foolish as asking where the twinness of a twin resides — is it in the DNA, where should we look for it? For the purposes of this paper, however I do not need to argue further in favour of my philosophical views or against other people's. I merely need to observe that different people use the term representation (and imagination) in different ways, and we must acknowledge that there are different senses of the terms; think "pavement" and "pavement".

Neuroscience

One role of a neuroscientist is to attempt to explain how the mechanisms of the brain allow us to see (and indeed even to imagine seeing) objects in the world around us. This is the same task as Descartes set himself, though perhaps nowadays we have less emphasis on the pineal gland and the immaterial spirit. Consider the image of an arrow as projected onto the back of the retina. In the picture shown, we have multiple representational layers of meaning. This is a reproduction — a representation — of the original woodcut first printed in Descartes' De Homini ('On Mankind'), 1662. We can see sketched on the right-hand side a representation of a vertical arrow. For the subject pictured here, of course, this is not a representation but rather it is a real arrow in front of it. If we perform our linguistic hygiene carefully, we should be able to avoid any confusion as to whether this is, or is not, a representation by clarifying the contextual relationship within which we are using the term. But then what about the image projected on the retina?

We, the observers of Descartes's woodcut, can see this projection, this representation. But the subject pictured there cannot see this, no more than they can see their own blindspot. It is the arrow in the world in front of it that the subject sees. We can, however, construct a metaphorical homunculus story where there is a chain of information flow, a pipeline, such that this retinal image is an intermediate carrier of information from the external world to some further internal mechanisms. Descartes certainly had one version of the story, and nowadays neuroscientists have different versions of such a story.

So the world (conceived as a homunculus) is conveying the information about the arrow in the world via this retinal image, this representation, to — to whom? To a posited receiving homunculus that, for the neuroscientist, may be some further subsystem of the brain. This could be a further staging post on the pipeline, as the pineal gland in Descartes's version. This is closely related to the example of the thermostat representing the room temperature, or the thresholded temperature, by an electrical signal sent to the central heating boiler. Is it a representation or not — that depends on which context you are spelling out; think "pavement" and "pavement".

A taxonomy of representations

I have spelt out above my own understanding of what I mean when I use the term representation in an everyday sense. I am aware that other people may disagree on some of these terms and conditions. When we moved to discussion of "internal representations" in cognitive science I personally am unwilling to change any of these terms and conditions (unless a clearly different technical sense of the term is defined and agreed upon). It follows necessarily that it makes no sense at all to talk of internal representations in the brain, except in the limited homuncular sense that I have outlined above. These are limited in that they are not representations "for me", rather they are representations "for some homunculus". Even though I can imagine seeing a unicorn, or visualise a map of

Brighton, this does not mean that there is a representation (for me) of either in my brain. There are structures or changes in my neural circuitry associated with my ability to imagine these things; but these structures or changes are simply not representational in the sense that I have spelt out.

Other people use the term in different ways, so let us make a start on classifying the different dimensions of meaning involved. To clarify, this is not so much an attempt to find the ***real*** meaning of the word (though I have my personal preferences), but rather an attempt to make visible the variety of possible senses that are used, just as road traffic experts should make explicit the two interpretations of "pavement".

1. Is representation a relational term or not?
2. If relational, is it essential that we should be able to contextualise the sender only, or the receiver only, or both or neither?
3. For internal representations, are they internal to the mind, the brain, or both or neither?
4. Is an internal representation a concept at a personal level, a sub-personal level, a neuro-anatomic level, or what?
5. Is it a concept within a functional explanation, or some other form of explanation?
6. Are such representations intentional or non-intentional concepts?
7. For internal representations, if you believe that there is a receiver, then is this receiver the person, or some substructure of the brain, or something else?
8. Is an internal representation of a cat in front of my eyes the same thing, or different from, the internal representation when the cat has gone but I am thinking of it?
9. Suppose somebody can navigate around Brighton without a map. Is the claim that they must have an internal representation of Brighton a logical claim that adds nothing to the previous statement; or an empirical claim that the nervous system underlying this cognitive capacity must be organised in some particular way?

In my experience, a room full of cognitive scientists and philosophers will, when challenged, produce a whole range of different responses to these questions. Not only are there these differences between different people, but also the same person may use the same word in different contexts with different senses. Yet typically, even when this is pointed out to them, this does not often seen to make them want to clarify their meaning in future discussions. I shall shortly suggest a hypothesis as to why this is so.

Presentations and Representations

Grush (1997: expanded in Grush 2004) does make a start at defining the terms he uses, and in particular makes distinctions between “presentations” and “representations”; also between “simulations” and “emulations”. To briefly summarise, he distinguishes between the two senses in item 8 listed above. He calls the former sense, as with direct sensory inputs of a cat, ‘presentational’, and reserves the term ‘representational’ for ‘counterfactual presentations’ such as considering the cat when it is not there.

Such a careful distinction is commendable, is consonant with the sentiments of this paper, and is regrettably all too rare amongst contributors to these debates. However Grush does not go far enough. In leading up to this, he states:

> If this second definition, and my gloss on it, are correct, then a representation is a part of a three-way relationship which also includes a user and a target. So far so good. Some may quibble over the need for a user, but that is not where the real problem lies. The real problem has been, and continues to be, the choice of states for which theorists attempt to give a representational analysis. Specifically, sensory states have been used as a model for representational states, the idea presumably being that sensory states represent the world to the subject. (Grush 1997)

So he is committed to a relational sense of the term, but less concerned about just how many partners there are in such a relationship, seeing this as a quibble and not where the real problem lies. Now my personal practice is normally to use the term with 4 such partners (**P**, **Q**, **R** and **S**), but I am prepared to engage in a discussion where the participants have agreed to use a different (e.g. 2 or 3-partner version) provided this is done openly and consistently. But it is simply unacceptable to leave it open to different participants in a discussion to have different senses in mind — and a sense that requires 2 partners must be a different sense to one that requires 3.

Representations as the billiard balls of cognitive science

We typically try to explain the complex and unfamiliar in terms that are simpler and more familiar. This is commendable and natural. Physicists will often try to explain atoms or other elementary particles in terms of billiard balls. We do not need to question how billiard balls will travel in a straight line, until they bounce off a wall or collide into each other. If we were perverse, however, we might demand to ask further questions. After all, billiard balls are made out of molecules and atoms, so is this not a circular explanation that does not bottom out anywhere? Well, in one sense *yes*, but in another sense *no,* because that is to misunderstand the role of an explanation. An explanation has to find some level of agreement, where no further questions are asked, and then try to reduce the complex and unfamiliar to this level. So for most purposes we can treat a billiard ball as an *explanans*, the explanatory premises, and this is incompatible with treating the billiard ball as an *explanandum*, that which is to be explained.

After many years of puzzlement, I have formed the tentative hypothesis that this lies at the root cause of cognitive scientists and philosophers being so reluctant to define what they mean by "representations" (Grush here being a partial exception, who has not gone far enough). For them, representations are the billiard balls of cognitive science. They are so familiar to them that they do not need to explain them further. Rather, they use them as part of the premises on which they build their cognitive theories. This makes them annoyed and irritated, just as physicists with billiard balls,

when one seeks an explanation for these premises.

This makes for difficulties if one thinks, as I do, that the capacity of human beings to represent things (common or garden everyday representations, in the external world) is supremely important and interesting; is the *explanandum*, not the *explanans*. In the phylogenetic history of human cognition, surely this counts as one of the Major Transitions (Maynard Smith and Szathmary, 1995). How did organisms start to create patterns in the world as representations of other objects or events in the world? Both the origin of, and the maintenance of, a capacity to represent things are surely amongst the main challenges for cognitive science. Artificial life techniques have a possible role to play here.

Minimal Cognition

What is the relationship between the physical mechanisms incorporated in the physical body of an organism and its behavioural capacities? The minimal cognition route aims to tackle this sort of question by starting at the bottom and building minimal models of minimally cognitive agents in some virtual environment. Evolutionary robotics allows one to evolve nervous systems to generate, we hope, the desired behaviour. When we come to analyse these, the advantage over real organisms is that we have full knowledge of the inner workings of all the mechanisms, and we can manipulate, alter constraints, lesion and otherwise experiment at will.

This kind of minimal cognition experiment (Harvey et al, 1997, 2005; Beer, 1995, 2000) typically has the status of a thought experiment, and an existence proof. If an artificial organism generates behaviour comparable to that of a real biological organism by use of mechanism X, this does not by itself prove that the real organism uses a similar mechanism X. It merely adds mechanism X to the list of possibilities; it is a separate scientific question as to which mechanism the real organism actually uses. The Artificial Life experiment can still be a very useful exercise, particularly if mechanism type X was previously thought unfeasible.

There have been many artificial life studies on communication, and language, but for the most part these have had built into them the possibilities of communication. To look at the major transition of the origin of representations one should perhaps start much earlier than this. Relevant work here is by Di Paolo (Di Paolo 2000) on the origin of social coordination and by Quinn (Quinn et al, 2002, 2003) on team behaviour. These studies look at communication between agents, two at a time (in simulation) in the first case, a group of three (both in simulation and on real robots) in the second case, where they have motives to influence each other's behaviour, and to do so via their actions. In the Quinn example, three agents or robots can sense each other through short range sensors, and move around on the plane. Their task is to travel across the plane in formation, which because of their sensors is only possible if they travel in a column of three. They are initially identical, so to achieve this task the first requirement is that they sort out between them the roles of leader, man in the middle, and tail. In some of the experiments a possible interpretation of what happens is that the symmetry is broken by the first one to make a stereotypical movement, which then determines its role. The others in some sense recognize this movement, and then take on the other roles.

It would be difficult (but not impossible) to claim that there is a fully-fledged representation going on here: "this stereotypical movement of mine, when suitably responded to by you, specifies or determines or represents what roles we should take". But if not fully-fledged, then arguably this is a transitional example. The stereotypical movement can be analysed either as (i) meaningless dynamics that nevertheless results in coordinated behaviour; or (ii) as a symbolic gesture from one agent to others. These two interpretations are not contradictory, they are framed at different levels of description. I suggest that these sorts of minimal cognition experiments are stepping stones on the way to evolving artificial agents where we can be more confident in calling them users of representations.

In Summary

There are both scientific and philosophical questions on the concepts of representations in cognitive science. Even the scientific questions cannot escape the philosophical issues of just what one might mean by the term. I am suggesting here, following Wittgenstein, that most of the philosophical problems and confusions come from poor linguistic hygiene. They are not issues of substance at all, merely the consequence of carelessness. On one interpretation of philosophy, I have not gone to any great lengths to argue for or against any particular philosophical position on representations. On another interpretation, this insistence on linguistic hygiene, to try to help people escape from the messes and confusions they make for themselves, is actually what philosophy is all about. Recall the difference between "pavement" and "pavement".

The common lack of care in defining terms may possibly, I have suggested, been partly because of the billiard balls role of representations in cognitive science; representations are so often treated as *explanans* rather than *explanandum*. It has always been irritating when those of us who share my opinion that we have no internal representations (as I understand the common sense of the term) in our heads have been branded as "Anti-Representationalists". I insist on calling myself a Representationalist, since I consider our human use of representations to be immensely important and interesting, and deservedly a focus of interest for Artificial Life studies. Representations as *explanandum*, not *explanans*.

This paper has been centred on the philosophical confusions, but there are proper scientific and technical questions to be asked. What are the minimal requirements for artificial agents to be capable of being representation-users? Understanding of this would seem to be a pre-requisite to the discussion of what are the minimal requirements for a sub-part (or module, or homunculus) within a brain or nervous system to be capable of being (metaphorically) representation-users in communicating with other modules. Artificial life methods have already made a tentative start to exploring these

questions, and we can hope for further progress. Artificial Life models give a superb arena in which these tricky, potentially ambiguous, terms can be given a demonstrably explicit sense, open to operational testing.

References

Beer, R.D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence* **72**:173-215.

Beer, R.D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences* **4**(3):91-99.

Clark, A. and Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior* 7 (1): 5-16.

Cliff, D., P. Husbands and I. Harvey (1993). Evolving visually guided robots. In: J.-A. Meyer, H. Roitblat and S. Wilson (eds.), *From Animals to Animats 2: Proc. of the Second Intl. Conf. on Simulation of Adaptive Behavior*, (SAB92), pp. 374--383. MIT Press/Bradford Books, Cambridge MA.

Descartes, R. (1662). *De Homini*

Di Paolo, E. A. (2000). Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adaptive Behavior* 8:1. 25 - 46.

Gallagher, S. (forthcoming). Are minimal representations still representations? *International Journal of Philosophical Studies*.

Grush, R. (1997). The Architecture of Representation. *Philosophical Psychology* 10(1):5-25.

Grush, R. (2004). The emulation theory of representation: motor control, imagery and perception. *Behavioral and Brain Sciences* 27:377-442.

Harvey, I., P. Husbands and D. Cliff (1993). Issues in evolutionary robotics. In: J.-A. Meyer, H. Roitblat and S. Wilson (eds.), *From Animals to Animats 2: Proc. of the Second Intl. Conf. on Simulation of Adaptive Behavior*, (SAB92), pp. 364--373. MIT Press/Bradford Books, Cambridge MA.

Harvey, I. (1996): Untimed and misrepresented: connectionism and the computer metaphor. *AISB Quarterly*, no. 96, pp. 20--27.

Harvey, I., P. Husbands, D. Cliff, A. Thompson, N. Jakobi (1997): Evolutionary Robotics: the Sussex Approach. In *Robotics and Autonomous Systems,* v. 20 pp. 205--224.

Harvey, I., Di Paolo, E., Wood, R., Quinn, M, and E. A., Tuci, (2005). Evolutionary Robotics: A new scientific tool for studying Cognition. *Artificial Life*, 11(1-2), pp. 79-98

Maynard Smith, J., and Szathmary, E. (1995). *The Major Transitions in Evolution.* Oxford University Press.

Quinn, M., Smith, L., Mayley, G. and Husbands, P. (2002). Evolving teamwork and role allocation for real robots. In Standish, R.K., Bedau, M.A. and Abbass, H.A., editors, *Proceedings of 8th International Conference on Artificial Life*, pages 302-311.

Quinn, M., Smith, L., Mayley, G. and Husbands, P. (2003). Evolving controllers for a homogeneous system of physical robots: Structured cooperation with minimal sensors. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, 361:2321-2344.

Rowlands, M. (2006). *Body Language*. Cambridge, MA: MIT Press.

Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge, MA:MIT Press.