Cognition in Context: Phenomenology, Situated Robotics and the Frame Problem

Michael Wheeler
Department of Philosophy
University of Stirling

Abstract

The frame problem is the difficulty of explaining how non-magical systems think and act in ways that are adaptively sensitive to context-dependent relevance. Influenced centrally by Heideggerian phenomenology, Hubert Dreyfus has argued that the frame problem is, in part, a consequence of the assumption (made by mainstream cognitive science and artificial intelligence) that intelligent behaviour is representation-guided behaviour. Dreyfus' Heideggerian analysis suggests that the frame problem dissolves if we reject representationalism about intelligence and recognize that human agents realize the property of thrownness (the property of being always already embedded in a context). I argue that this positive proposal is incomplete until we understand exactly how the properties in question may be instantiated in machines like us. So, working within a broadly Heideggerian conceptual framework, I pursue the character of a representation-shunning thrown machine. As part of this analysis, I suggest that the frame problem is, in truth, a two-headed beast. The intra-context frame problem challenges us to say how a purely mechanistic system may achieve appropriate, flexible and fluid action within a context. The inter-context frame problem challenges us to say how a purely mechanistic system may achieve appropriate, flexible and fluid action in worlds in which adaptation to new contexts is open-ended and in which the number of potential contexts is indeterminate. Drawing on the field of situated robotics, I suggest that the intra-context frame problem may be neutralized by systems of special purpose adaptive couplings, while the inter-context frame problem may be neutralized by systems that exhibit the phenomenon of continuous reciprocal causation. I also defend the view that while continuous reciprocal causation is in conflict with representational explanation, special-purpose adaptive coupling, as well as its associated agential phenomenology, may feature representations. My proposal has been criticized recently by Dreyfus, who accuses me of propagating a cognitivist misreading of Heidegger, one that, because it maintains a role for representation, leads me seriously astray in my handling of the frame problem. I close by responding to Dreyfus' concerns.

1. Reviving the Frame Problem

I'm prepared to bet that most cognitive scientists would agree, in broad terms at least, with the following thought: any system worthy of the epithet 'intelligent' must be able to retrieve from its memory just those items or bodies of stored information that are most relevant to its present context, and then decide how to use, update, or weight that information, in contextually appropriate ways, in processes such as belief fixation and action selection. But many of those same researchers would also agree that we don't yet know how to explain this cognitive achievement scientifically – not really. The qualification 'not really' is important, because the prevailing view would be that the broad shape of the explanation has been known for a very long time, although the details remain to be settled. Intelligent behaviour is representation-guided behaviour. Intelligence is essentially a matter of sensitivity to information, and the capacity of certain systems to realize such informational sensitivity is explained at root by the fact that those systems are able to build, store and manipulate internal representations. Call this the *orthodox representationalist story,* henceforth *ORS.*

Of course, if all ORS says is that intelligent agents are those agents that are configured (in our case by evolution and learning) so that they retrieve just those representations that are appropriate to a given context and then update, in contextually appropriate ways, just those representations that need to be updated, then the putative explanation on the table would be no more than a restatement, in representational language, of the phenomenon to be explained. An immediate response here may be that I am failing to see the scope and power of ORS, that part of what it means for an agent to build, store and manipulate internal representations is precisely for that agent to construct representations of any context in which it finds itself, and then to use those representations to guide its processes of search, selection and update. This idea will be assessed later. For now, as we edge our way into the issues, it is more important that we draw a general lesson from the way in which the response explicitly targets the question of context. That lesson is that phrases such as 'appropriate to a given context' and 'in contextually appropriate ways' are not innocent flourishes of the language in which we describe intelligence. Rather, they highlight a necessary feature of the phenomenon. In other, more dramatic, words, a theory of intelligence that fails to give us a satisfactory explanation of how cognitive processes achieve context-sensitivity is not an incomplete theory of intelligence, it is no theory of intelligence at all.

As I shall use the term, *the frame problem* is the difficulty of explaining how non-magical systems (machines like us) think and act in ways that are adaptively sensitive to context-dependent relevance. There was a time when the frame problem so conceived was a hot research topic in cognitive science, especially in artificial intelligence (AI). These days it may seem a little passé, which is a shame since (with sincere apologies to various researchers) it's not as if anybody ever actually solved the problem. Maybe we just got bored reading and writing about it. In any case, it is high time we had a frame problem revival. As a contribution to this revival, the present paper has three main aims. The first (sections 2 and 3) is to explore a seminal analysis of the frame problem due to Hubert Dreyfus, an analysis in which Heideggerian phenomenology is deployed as an analytical

tool. The second (section 4) is to present what I take to be the beginnings of a potential solution to the frame problem, a solution that combines Heideggerian phenomenology and cognitive-scientific naturalism as mutually constraining influences. In pursuing these first two goals, the paper revolves around an updated (and so one hopes improved) version of an argument that I've presented before, although in its previous incarnation it was just one thread in a bigger project and so was realized in a somewhat diffuse and partially buried way (Wheeler, 2005).[1] This previous attempt on my part to navigate the issues has been criticized recently by Dreyfus (2008). To the extent that these criticisms are on target, they apply just as much to the reworked treatment presented here as they did to the original one, so the third main aim of this paper (section 5) is to respond to Dreyfus' concerns. Let's begin, then, with a closer look at the frame problem itself.

2. Stalking the Frame Problem

The frame problem originated in logicist AI, where it first emerged as the problem of characterizing, using formal logic, those aspects of a state that are not changed by an action (see e.g. Genesereth and Nilsson, 1987). It soon turned out, however, that this original frame problem was just one dimension of a more general and multi-faceted difficulty confronted by any broadly mechanistic account intelligence, given the context-embeddedness and relevance-sensitivity of cognition. (See the range of discussions in (Pylyshyn, 1987).) As a result the term 'frame problem' has taken on a wider meaning, although its historical roots are reflected in the fact that the ultimate litmus test for any proposed solution is still widely considered to be whether one could engineer a robot – a wholly mechanistic system acting in the real (dynamic, physical) world – that solved or avoided the problem precisely by implementing the proposal in question. Expressing a generalized version of the worry, Fodor famously describes the frame problem as 'the problem of putting a "frame" around the set of beliefs that may need to be revised in the light of specified newly available information' (Fodor, 1983: pp.112-13). This formulation still isn't general enough, however, because knowing what to do is just as context-sensitive as knowing what to believe. Thus the frame problem is not merely a problem associated with belief-fixation, it is a problem associated with belief-fixation and action-selection. But now if we interpret 'thinking' broadly, so as to cover both of these cognitive phenomena, another of Fodor's characterizations does get to the heart of the matter. Fodor (1987) glosses the frame problem as the problem of how a robot, when working out the consequences of its actions (and thus what it should do), could ever be completely confident that it hadn't failed to consider something important. The frame problem is thus 'Hamlet's problem: when to stop thinking… viewed from an engineer's perspective' (Fodor, 1987: p.140).

It might seem that what one needs to do here is find a mechanistic way of pragmatically limiting the search space to be explored. In this spirit Holland et al. (1986) write: 'a processing mechanism of the sort we favor circumvents the problem of the potential relevance of everything in the knowledge store by pragmatically selecting limited areas

---

[1] Some passages in sections 2-4 of this paper adapt textual material from (Wheeler, 2005; 2008).

of information to explore... By tending to fire the strongest and most goal-appropriate rules, a constrained search through the space of relevant information can be carried out'. But how does the system itself select the most goal-appropriate rules and information? One natural thought (in line with ORS, see above) is that it should deploy stored heuristics (internally represented rules of thumb) or representations of context that determine which of its stored bodies of information are relevant in the present situation. All this does, unfortunately, is push the problem one stage back. For how does the system decide which of its stored heuristics or potentially context-specifying representations are relevant? Another, higher-order set of heuristics or representations would seem to be required. But of course the same problem will re-emerge at that higher level. So, depending on how one looks at it, a combinatorial explosion or infinite regress beckons. Dreyfus gives voice to just this sort of point, when he considers the plight of an AI-programmed computer faced with incoming environmental data:

> The significance to be given to each logical element [each internally represented piece of data] depends on other logical elements, so that in order to be recognized as forming patterns and ultimately forming objects and meaningful utterances each input must be related to other inputs by rules. But the elements are subject to several interpretations according to different rules and which rule to apply depends on the context. For a computer, however, the context itself can only be recognized according to a rule…
>
> …[T]o pick out two dots in a picture as eyes one must have already recognized the context as a face. To recognize this context as a face one must have distinguished its relevant features such as shape and hair from the shadows and highlights, and these, in turn, can be picked out as relevant only in a broader context, for example, a domestic situation in which the program can expect to find faces. This context too will have to be recognized by its relevant features, as social rather than, say, meteorological, so that the program selects as significant the people rather than the clouds. But if each context can be recognized only in terms of features selected as relevant and interpreted in terms of a broader context, the AI worker is faced with a regress of contexts. (Dreyfus, 1992: pp.288-9)

An infinite regress would be bad enough, but may not be the worst of it. As Horgan and Tienson (1994) point out, the context-sensitivity of cognition cannot be achieved by a system first retrieving an inner structure (an item of information or a heuristic), and then deciding whether or not it is relevant, as that would take us back to square one. But then how can the system assign relevance until the structure has been retrieved? The result is a kind of cognitive paralysis.

Despite what has been said so far, one might be tempted to think that the frame problem is a difficulty only for ideally rational systems in search of optimal behaviour, and that it is pretty much irrelevant in the case of real human beings, who famously enjoy only bounded rationality (Shanahan, 2006). According to the bounded rationality model

(Simon, 1955), human beings get by (we satisfice rather than optimize) on the basis of certain cognitive tricks, limitations and constraints that characterize our less-than-ideal rationality. These might include ignoring features that are shared by incompatible options, imperfect recall (here an adaptively beneficial property), and, given that there are cognitive-resourcing costs associated with gathering and holding information, making decisions about what one should know.[2] In the present context, however, it is questionable just what the bounded rationality picture buys us. In some cases, for example in deciding what to know, it looks as if determining when to use the trick might itself be a context-sensitive affair, in which case the frame problem continues to apply directly. Indeed, the response from bounded rationality would be reduced to nothing more than a fancy version of the appeal to heuristics to reduce the informational search space. But perhaps the idea is that the tricks, limitations and constraints apply in a context-independent manner, in which case it is desperately unclear that such across-the-board processing restrictions will reliably have the effect of blunting the search problem in the right kind of way. For sure, some options and some information won't be considered, but unless the pairing down of the search space is done in a relevance-guided manner, there seems to be no reason to think that the boundary between considered and ignored will in any way reflect the boundary between relevant and irrelevant, so the frame problem will remain in force.

About now one might be moved to wonder why AI hasn't simply ground to a halt in the jaws of the frame problem. According to many theorists (e.g. Dreyfus and Dreyfus, 1988; Brooks, 1991; Cliff, 1994; Dreyfus, 2008), most AI researchers (classical and connectionist) have typically managed to side-step the frame problem by assuming that real-world problem-solving can be treated as a kind of messy and complicated approximation to reasoning (or learning) in artificially restricted worlds that are relatively static, essentially closed, and feature some small number of contexts of action. In such worlds, all the contexts that could possibly arise may be identified and defined, alongside all the factors that could possibly count as relevant within each of them. So the system designer can take comprehensive and explicit account of the effects of every action or change. And if this strategy incurs a prohibitive processing load, the designer can either (i) work on the assumption that nothing changes unless it is explicitly said to change by some rule, or (ii) use carefully targeted relevancy heuristics. The upshot is that in such well-defined and well-behaved problem-domains, the frame problem is no more than a nuisance. The fact is, however, that the actual world often consists of an indeterminate number of dynamic, open-ended, complex scenarios in which context-driven and context-determining change is common and ongoing, and in which vast ranges of cognitive space might, at any time, contain the relevant psychological elements. It is in this sort of world that the frame problem really bites, and in which the problem-solving strategies that have been successful in toy worlds and toy domains will fail to deliver.

---

[2] This list comes from Rubinsetin (1998), although the author himself remains largely neutral about the psychological validity of the constraints and strategies in question.

3. From a Phenomenological Point of View

ORS (the orthodox representationalist story) is the generic picture of cognition on which mainstream cognitive science and AI (classical and connectionist) is built. Now I want to forge an explicit link between ORS and the frame problem, by briefly exploring an analysis due to Hubert Dreyfus. As I shall interpret him, Dreyfus offers us a tripartite diagnosis of why ORS invites the frame problem. This diagnosis turns on phenomenological insights drawn from the work of Heidegger (e.g. 1927) and Merleau-Ponty (e.g. 1962). However, given the Heideggerian character of the steps towards a solution to the frame problem that I shall take later in this paper, I shall place the present interpretative emphasis on the Heideggerian dimension of Dreyfus' analysis, as he himself sometimes does (e.g. Dreyfus 1990).

The first part of Dreyfus' analysis turns on the massively holistic character of contextual significance, as revealed by Heideggerian phenomenology. Famously, Heidegger argues that we ordinarily encounter entities as *equipment*, that is, as being *for* certain sorts of tasks (typing, cooking, hair-care, and so on). Entities so encountered have their own distinctive kind of intelligibility that Heidegger calls *readiness-to-hand*. He introduces the term *involvements* to capture the significance of equipmental entities (the ways in which they are involved) in our everyday activities and tasks. Crucially, for Heidegger, an involvement is not a stand-alone structure, but rather a link in a network of intelligibility that he calls a *totality of involvements*. Thus I am currently working with a computer (an involvement that Heidegger calls a *with-which*), in the practical context of my office (an *in-which*), in order to write this paper (an *in-order-to*), which is aimed towards developing an analysis of the frame problem (a *towards-this*), for the sake of my academic work, that is, for the sake of my being an academic (a *for-the-sake-of-which*). Totalities of involvements constitute the contexts of everyday human activity, the contexts that, as we have seen, determine what is relevant at a given time and thereby provide the backdrop for the frame problem. To glimpse the frame problem here, one needs to appreciate the extent of Heidegger's holism. Once one begins to trace a path through a network of involvements, one will inevitably traverse vast regions of involvement-space. Thus connections will be traced not only from computers to offices to paper-writing to academia, but also from computers to computer games to negotiating with my son over how much time he should be allowed to play such games to good parenting. This behaviour will refer back to many other behaviours (taking my son to football training on a Saturday morning, taking him to see Hibernian FC play) and thus to many other items of equipment (footballs, football boots, replica kits, match-day programmes), and so on. Contextual significance is thus massively holistic.

Having adopted Heidegger's picture, Dreyfus argues that the massively holistic character of contextual significance presents ORS with a serious difficulty. As we have seen, one way in which the mechanistic agent designed according to the ORS blueprint is supposed to achieve context-sensitivity is by internally representing the contexts within which belief-fixation or action-selection needs to occur. From a Heideggerian perspective, however, any attempt to internally reconstruct highly distributed and interconnected networks of involvements, by building inner representations of those networks (e.g. as

atomic nodes and the links between them), looks to be a prohibitively difficult, and perhaps even an infinite, task. With Heideggerian holism on the table, one might think it unsurprising that ORS will encounter the combinatorial explosion that is one tell-tale sign of the frame problem.

The second part of Dreyfus' analysis turns on the issue of skills. According to Dreyfus, to have a skill is to 'come into a situation with a readiness to deal with what normally shows up in that sort of situation' (Dreyfus, 1990: p.117). In other words, it is to be equipped with a prior capacity to be flexibly sensitive to what is (normally) relevant in that kind of context. This takes us back to the notion of equipment. For Heidegger, we achieve our primary relationship with equipment not by looking at the entity in question, or by some detached intellectual or theoretical study of it, but rather by skillfully manipulating it in a hitch-free manner, where at least part of what it means to manipulate an item of equipment skillfully is to be sensitive to contextual factors. Following others, I shall call this primary mode of engagement with entities *smooth coping*. Smooth coping is the principal instantiation of Dreyfusian skills. Moreover, smooth coping realizes a mode of knowledge. As Heidegger puts it: the 'kind of dealing which is closest to us is... not a bare perceptual cognition, but rather that kind of concern which manipulates things and puts them to use; and this has its own kind of 'knowledge'' (Heidegger, 1927: p.95). In contemporary philosophical terminology, this mode of knowledge would standardly be identified as a form of knowledge-how.

Although smooth coping is the principal way in which Dreyfusian skills are manifested, it is not, I think, the only way. According to Heidegger, *un-readiness-to-hand* is a species of intelligibility that emerges when smooth coping is disturbed by broken or malfunctioning equipment, discovered-to-be-missing equipment, or in-the-way equipment. When confronted by the un-ready-to-hand, the human agent adopts a strategy of practical problem solving (e.g. executing a repair). But notice that the broken, malfunctioning, missing or obstructive status of un-ready-to-hand entities is defined relative to a particular equipmental context. Thus a driver does not encounter a punctured tyre as a lump of rubber of a measurable mass; she encounters it as a damaged item of equipment, that is, as the cause of a disruption to her driving activity. Heidegger is clear that when 'something cannot be used – when, for instance, a tool definitely refuses to work – it can be conspicuous only in and for dealings in which something is manipulated' (Heidegger, 1927: p.406). So practical problem solving in the domain of the un-ready-to-hand remains context-sensitive. *In at least some cases*, this will be another manifestation of a Dreyfusian skill, and thus of knowledge-how. The qualification 'in at least some cases' allows for examples of practical problem solving that approximate, without quite becoming, the kind of theoretical reasoning distinctive of science, a detached reflective engagement that reveals entities in the context-independent mode of intelligibility that Heidegger calls *presence-at-hand*.[3]

---

[3] We should register the following subtlety in Heidegger's analysis. According to Heidegger it is an essential fact about human cognition that it always operates within some sort of context of activity. However, when an agent adopts the detached theoretical attitude, she enters a special kind of context of activity, one in which the entities thereby

For the advocate of ORS, a Dreyfusian skill (flexible sensitivity to a context) must result from the inner deployment of context-specifying representations. From a Heideggerian perspective, however, this attempt to explain the phenomenon appears seriously misguided. Here is my reconstruction of Dreyfus' argument:

1. The representations favoured by ORS are a form of knowledge-that.
2. A Dreyfusian skill is a kind of knowledge-how.
3. Knowledge-how cannot be reduced to knowledge-that.
So
4. ORS must fail to explain Dreyfusian skills.

According to Dreyfus, then, ORS radically misconceives the kind of knowledge that underlies the way in which we enter situations with a readiness to deal with what normally shows up in those situations. It is perhaps unsurprising that what results is a difficulty like the frame problem.

To bring the third, final and most important dimension of Dreyfus' analysis into view, consider the following quotation from Heidegger.

> What we 'first' hear is never noises or complexes of sounds, but the creaking waggon, the motor-cycle. We hear the column on the march, the north wind, the woodpecker tapping, the fire crackling... It requires a very artificial and complicated frame of mind to 'hear' a 'pure noise'. The fact that motor-cycles and waggons are what we proximally hear is the phenomenal evidence that in every case Dasein [the human agent], as Being-in-the-world, already dwells alongside what is ready-to-hand within-the-world; it certainly does not dwell proximally alongside 'sensations'; nor would it first have to give shape to the swirl of sensations to provide a springboard from which the subject leaps off and finally arrives at a 'world' [a context]. Dasein, as essentially understanding, is proximally alongside what is understood. (Heidegger, 1927: p.207)

As Heidegger lays things out, his opponent thinks that the human agent is in primary epistemic contact with a set of context-independent primitives (e.g. raw sense data, such as an experience of a patch of red) to which context-dependent significance must somehow be added by cognitively downstream processing. By contrast, Heidegger's own view is that the agent is in primary epistemic contact not with bare context-independent elements, but rather with equipment, the kind of entity that comes already laden with context-dependent significance. This is an aspect of the phenomenon that

---

disclosed are fully context-independent. So although there is, in Heidegger's picture, no such thing as a fully decontextualized subject, there are fully decontextualized objects.

Heidegger calls *thrownness*, the fact that the human agent always finds herself located in a meaningful context (what Heidegger means here by a 'world') in which things matter to her.

There seems little doubt that ORS, as realized in mainstream cognitive science, sides with Heidegger's opponent here. For example, Marr (1982) famously assumed that the main task confronting vision is to derive representations of the 3D shapes of objects from 2D arrays of light intensity values at the retina. Here both the representational input to vision (the 2D array of light intensity values at the retina) and its output (models of the 3D shapes of objects) are context-independent representations. Contextual significance is a cognitively downstream addition. In Dreyfus' analysis, this explains why the inner representations appealed to by ORS require additional elements (such as relevancy heuristics) in order to determine which of those representations are appropriate to any particular context. We might put it like this: All the level 0 elements in an ORS-style cognitive architecture are context-independent in nature, so context-dependent relevance needs to be assigned, at level 1, by further elements. But these level 1 elements require yet further elements, at level 2, to determine their correct contextual application. Why? Because those level 1 elements are, in the first instance, context-independent in nature. But the level 2 elements in question require yet further elements, at level 3, to determine their correct contextual application. Why? Because those level 2 elements are, in the first instance, context-independent in nature… and so on. On this analysis, it is unsurprising that the repeated application of the orthodox strategy (the use of progressively higher order relevance-assigning elements) succeeds only in pushing the problem of relevance further and further back. The intrinsic context-independence of the elements in play, at whatever level of depth they are first applied, explains this regress.

So, as I interpret him, Dreyfus holds, on Heideggerian grounds, that the frame problem can be traced to three things: (i) that ORS's representation-building strategy runs aground in the face of contextual holism, (ii) that our ability to behave in context-sensitive ways is conceived by ORS as a form of knowledge-that when in fact it is a form of knowledge-how; and most importantly (iii) that an infinite regress results from the need for an ORS-designed cognitive system to assign relevance to its intrinsically context-free representations. So do we now call time on ORS and its cognitive-scientific adherents? Not quite. To see why ORS is not yet out of the game, we need to begin by noting that although I have stressed the critical aspect of Dreyfus' analysis, it can also be read as offering a Heideggerian solution to the frame problem, or, more accurately, as offering a Heideggerian *dissolution of* the frame problem. This has two prongs:

1. Reject all forms of representationalism about intelligence. In a recent paper, Dreyfus declares that 'for Heidegger, *all* representational accounts are part of the problem' (Dreyfus, 2008: p.358). This move frees the cognitive scientist from the need to represent context and removes the temptation to treat skills as a form of knowledge-that rather than a mode of knowledge-how.

2. Argue that since human agents are characterized by thrownness, they are always already embedded in some meaningful context. Because of this, human agents are

never in the position of having to add contextual significance to context-independent primitives (see Dreyfus, 1992: p.262-3).

The idea, then, is that for a representation-shunning thrown agent, the frame problem simply doesn't arise.

Unfortunately, from the perspective of cognitive science, this Heideggerian dissolution of the frame problem is radically incomplete. There is something like a Kuhnian argument here. Kuhn (1970) argued that a crisis in an established scientific paradigm is no more than a necessary precondition for the rejection of that paradigm. A crisis turns into a scientific revolution (a transition between paradigms) only when an established paradigm is successfully challenged by a genuine rival. Ideally, this rival should either demonstrate the potential to solve the problems faced by the previous paradigm, or avoid those problems altogether, in the sense that they simply do not arise for it. (Measuring the weight of phlogiston was a significant problem for some pre-Lavoisierian chemists, but a non-issue for Lavoisier.) In general, then, Kuhnian reasoning supports the following thought. Whatever problems mainstream cognitive science may currently face, there is no justifiable basis for fundamental change unless a plausible suggestion for an alternative explanatory framework (preferably one for which there is some good evidence of early empirical success) has emerged. So now let's apply this Kuhnian reasoning more precisely. What it tells us is that even if ORS, in the form of mainstream cognitive science, has no good answer to the frame problem, that is no reason for cognitive science to reject its tenets in favour of Dreyfus' proposed Heideggerian alternative, *unless* a plausible solution to, or dissolution of, the frame problem is available within that new approach. But didn't we just get offered such a dissolution? Well, no, not *from the perspective of cognitive science*. This is because we don't as yet have any account of how Dreyfus' proposal could be realized by the kind of non-magical, mechanistic system that provides the material for cognitive-scientific explanation. What we need, then, is a plausible example of a *representation-shunning thrown machine*.

This is a point that, in mildly different terms, I have made before (Wheeler, 2005). Recently, Dreyfus has taken up the challenge. Later in this paper I shall have something to say about his own specific suggestion for how a Heideggerian cognitive science might go. First, however, I want to present my own suggestions for how to begin to develop a Heideggerian response to the frame problem that takes seriously the Kuhnian argument. As we shall see, this response embraces the thought that part of the story will be a mechanistic explanation of thrownness, but it refuses the Dreyfusian call to shun representations altogether. That tension between Dreyfus' analysis and my own will be explained in the next section and its resolution will form part of the business of section 5.

4. Thrown Machines

Dreyfus is surely right that thrownness must be part of any Heideggerian solution to the frame problem. In addition, as we have just seen, if we are to give a response to the frame problem that will satisfy the cognitive scientist, we need to supplement phenomenological analysis and provide an account of the mechanisms that causally

explain how an agent may be thrown. In my view this account has already been given (although not in explicitly Heideggerian terms) within the area of contemporary AI sometimes known as *situated robotics*. Researchers in situated robotics favour the design and construction of complete robots that are capable of integrating perception and action in real time so as to generate fast and fluid embodied adaptive behaviour. In addition, such roboticists shun the classical cognitive-scientific reliance on detailed internal world models, on the grounds that such structures are computationally expensive to build and keep up to date. Instead they adopt a design strategy according to which the robot regularly senses its environment to guide its actions (examples in a moment). It is this specific behaviour-generating strategy that marks out a robot as *situated* (Brooks, 1991).[4] Against this background, one of the key ideas from situated robotics is that much of the richness and flexibility of intelligence is down not to general-purpose processes of reasoning and inference, but to integrated suites of special-purpose adaptive couplings that combine neural mechanisms (or their robotic equivalent), non-neural bodily factors, and environmental elements, as 'equal partners' in the behaviour-generating strategy (again, examples in a moment). In my view, situated special-purpose adaptive couplings may make it intelligible to us how it is that unmysterious causal processes, such as those realized subagentially in brains, can give rise to the agential level phenomenon of thrownness. To unpack this claim I shall return to an example I have used a number of times before, because it makes the key point so clearly.

Consider the ability of the female cricket to find a mate by tracking a species-specific auditory advertisement produced by the male. According to Barbara Webb's robotic model of the female cricket's behaviour, here, roughly, is how the phonotaxis system works (for more details, see Webb, 1993; 1994; or the discussion in Wheeler, 2005). The basic anatomical structure of the female cricket's peripheral auditory system is such that the amplitude of her ear-drum vibration will be higher on the side closer to a sound-source. Thus, if some received auditory signal is indeed from a conspecific male, all the female needs to do to reach him (all things being equal) is to continue to move in the direction indicated by the ear-drum with the higher amplitude response. So how is it that the female tracks only the correct stimulus? The answer lies in the activation profiles of two interneurons (one connected to each of the female cricket's ears) that mediate between ear-drum response and motor behaviour. The decay rates of these interneurons are tightly coupled with the specific temporal pattern of the male's song, such that signals with the wrong temporal pattern will simply fail to produce the right motor-effects.

Now, here is Webb's own explanation of why the mechanism just described is adaptively powerful: 'Like many other insects, the cricket has a simple and distinctive cue to find a mate, and consequently can have a sensory-motor mechanism that works for this cue and nothing else: there is no need to process sounds in general, provided this specific sound has the right motor effects. Indeed, it may be advantageous to have such specificity built

---

[4] In contemporary philosophy of mind and cognitive science, the term 'situated' is often used to mean 'environmentally embedded'. In this paper I shall use it exclusively to name the specific behaviour-generating strategy described in the main text, which one might think of as a form of environmental embedding.

in, because it implicitly provides 'recognition' of the correct signal through the failure of the system with any other signal' (Webb, 1993: p.1092). So the situated special-purpose adaptive coupling that constitutes the cricket phonotaxis mechanism works correctly *only in the presence of the right, contextually relevant input*. A reasonable gloss on this picture is that, rather than starting outside of context and having to find its way in using relevancy heuristics and so on, the cricket's special purpose mechanism, in the very process of being activated by a specific environmental trigger, brings a context of activity along with it, implicitly realized in the very operating principles which define that mechanism's successful functioning. Here, context is not something that certain causal mechanisms must reconstruct, once they have been triggered. Rather, context is something that is *always there at the point of triggering*, in the adaptive fabric of the activated mechanism. This, I suggest, is the subagential mechanistic mark of thrownness.[5]

There is a worry here. Shaun Gallagher is concerned that, in my account, 'the term 'context' is bouncing a little too freely between agent level [phenomenological] and subagential level [mechanistic] discourse' (Gallagher, 2007). If I understand this correctly, the criticism rests on the thought that context is an exclusively agential level

---

[5] One might complain that the cricket simply doesn't have a context, so whatever the adaptive contribution of the phonotaxis mechanism may be, it cannot causally explain context-sensitive behaviour. This argument is not compelling because the opening premise is arguably false. The meaningful and thus normatively loaded character of non-human animal behaviour is manifestly obvious, and although Heidegger got himself into all sorts of trouble over the issue of non-human animals (Derrida, 1989), I see no insurmountable barrier to conceiving such meaning and normativity in terms of contexts (networks of significance). That said, the criticism might be resurrected in the following form: although the cricket has a context, it is not a context in anything like the human sense, so whatever the adaptive contribution of the phonotaxis mechanism may be, it cannot causally explain context-sensitive behaviour in a way that throws light on human-level context-sensitivity. This is a more serious worry, but I remain unmoved. Of course insect-level contexts are not the same as human-level contexts. At a minimum the former are presumably comprised entirely of evolutionarily determined norms, whereas the latter involve a complex combination of evolutionary and cultural determination. But this genuine difference should not blind us to the fact that the context-sensitivity in question, whatever the source of the contexts themselves, may be causally achieved by a similar underlying process, that is, by the activity of situated special-purpose adaptive couplings. At this point my critic might reply that the cricket's phonotaxis mechanism is hard-wired by evolution, whereas many of the routines that enable human beings to navigate cultural contexts will be installed by learning. However, with regard to the present issue, this is not a difference that makes a difference. As long as the notion of a 'mechanism' is not understood in some overly restrictive manner, installed routines of the sort at issue will count as cognitive mechanisms; and there is no reason to think that the learned status of those routines must prevent them from being situated special-purpose adaptive couplings. For further discussion of how to think about animal worlds from a broadly Heideggerian perspective, see (Wheeler, 1995; 2005).

phenomenon that cannot legitimately appear (explicitly or implicitly) at the mechanistic, subagential level. If that is what one thinks, then my claim that there is a sense in which one can find context 'down there' in the causal mechanisms seems to endorse a reductionist picture that, whatever its plausibility outside of Heideggerian circles (not very high, I suspect), is anyway clearly not available within them. So what should I say? Gallagher's worry alerts us to the fact we need to be careful about what 'context in the causal mechanisms' might mean. I have argued that we need to identify unmysterious causal mechanisms at the subagential level that *make it intelligible to us* how certain phenomenological descriptions, such as thrownness, could be true of whole agents. However, this demand *need* not be heard reductively, because intelligibility may be secured without reduction. Crucially this remains true, I think, even if we decide to *use some of our familiar agential terms as explanatory terms of art at the subagential level.*

One story about how this might work may be given if we apply a moral drawn from McDowell (1994). McDowell argues that when we attribute mental content to whole agents, we are saying something that is literally true or false. By contrast, cognitive-scientific explanations that attempt to account for agential phenomena by describing informational transactions within a subagential control system, and which thus attribute representational contents to neurally realized vehicles, are engaged in a practice of attributing only *metaphorical* content to those vehicles. This is no threat to the explanatory credentials of cognitive science however, because, in McDowell's own words, 'it is surely clear, at least in a general way, how content-attribution that is only 'as if' can even so pull its weight in addressing a genuine explanatory need: the question is what enables us animals to be the semantic engines we are' (McDowell, 1994: p.199). McDowell's claim that subagential content must be metaphorical in character is only one way of playing out the key idea here, which is a good thing because that claim is unnecessarily strong. As long as there is a real and important difference between agential and subagential content (e.g. in the conditions for what counts as a representation, or in how to fix content), such that the former cannot be reduced to the latter, then we can have our cake (we can use a term such as 'content' at the subagential level) and eat it too (we don't need to 'go metaphorical' about the subagential use of that term). It seems to me that the lesson here generalizes to other terms whose conceptual home is agential level discourse but which, to address a genuine explanatory need, may become terms of art at the subagential level; and that includes the term 'context'. When we talk of context-sensitivity at the agential level, we describe a necessary aspect of what it is for an intelligent agent to competently inhabit its environment. When we talk of context being 'brought along' by certain subagential mechanisms, of context being 'always there at the point of subagential triggering', or of context being 'woven into the adaptive fabric of the activated mechanism', we are giving further expression to the point that such mechanisms do not causally underpin agential level context-sensitivity by virtue of building structures that subagentially represent context, but rather by virtue of the fact that they are situated special-purpose adaptive couplings (more on this below). This should go some way towards allaying Gallagher's worry.

Next I want to argue that we should refuse Dreyfus' invitation to shun representations. Rather, we should tame those troublesome entities and put them in their place. This might

seem like an odd strategy for me to pursue. After all, the cricket's phonotactic mechanism is surely non-representational in character, which appears to provide some support for Dreyfus' anti-representational stance. But although, as a matter of fact, many situated-robotics-style explanations of intelligent action appeal to non-representational adaptive mechanisms, representations have not been excised entirely from the overall picture. Rather, the situated roboticist's characteristic upgrading of the behaviour-generating contributions made by the non-neural body and the environment sometimes leads to the traditional reliance on representational elements being reconfigured rather than rejected altogether. Crucially for the argument of this paper, this reconfiguration allows the mechanisms concerned to continue to display the subagential mechanistic mark of thrownness, as exhibited in the case of Webb's cricket robot.

To illustrate these points I shall turn to another example that I have used before. Franceschini et al. (1992) set themselves the task of building a robot that navigates its way to a light source while avoiding obstacles. The resulting system achieves this goal by executing a sequence of movements, each of which is generated in the following way. A primary visual system, inspired, in part, by the compound eye of the fly, features a layer of elementary motion detectors (EMDs). Since these components are sensitive only to movement, the primary visual system is blind at rest. What happens, however, is that the EMD layer uses relative motion information, generated by the robot's own bodily motion during the immediately preceding movement in the sequence, to build a temporary snap map of detected obstacles, constructed using an egocentric coordinate system. Then, in an equally temporary motor map, information concerning the angular bearings of those detected obstacles is fused with information concerning the angular bearing of the light source (supplied by a supplementary visual system) and a direction-heading for the next movement is generated. This heading is as close as possible to the one that would take the robot straight towards the light source, adjusted so that the robot avoids all detected obstacles.

Notice that the ways in which objects are represented by this robot's action-generating maps are deeply dependent upon the specific obstacle-avoiding-homing context and the manner in which the robot uses its own history of physical movement and its close coupling to its local environment to structure its behaviour. The shape, absolute position, and/or orientation of objects are neither calculated nor stored. Consider, for example, objects other than lights. These are located as distally located edges fixed by contrast points in the optic flow. The obstacle-avoidance mechanism treats these contrast points as revealing regions of the environment from which to steer away, defined in terms of angular bearings relative to the robot itself. Thus objects (other than lights) are represented as avoidance-regions or motion-barriers in an egocentrically defined space. Pulling out the key lessons, we can say that the representations in question are *action-specific*, in that they are tailored to the job of producing the particular navigational behaviour required and are designed to represent the world in terms of specifications for possible actions; they are *egocentric*, in that the snap map of detected obstacles features an agent-centred coordinate system, and the motor map exploits agent-based angular bearings; and they are *intrinsically context-dependent*, in that, as in the case of the cricket phonotaxis mechanism, the explicit representation of context is eschewed in favour of

situated special-purpose adaptive couplings that implicitly define the context of activity in their basic operating principles. This is, of course, the aforementioned subagential mechanistic mark of thrownness.

Borrowing a term (although not its precise usage) from Andy Clark (1997), I shall call states that have the foregoing profile, *action-oriented representations*. But now isn't it the case that my resurrection of a representational form of explanation has a decidedly un-Heideggerian ring to it? I think not, because, as I interpret Heidegger, the mechanism-based action-oriented representations that I have just described have phenomenological counterparts in the arena of practical problem solving in the domain of the un-ready-to-hand. To see how this works, we need to do some groundwork on the relationship between agential level representations and the subject-object dichotomy. It is, I think, hard to see how to make sense of an agent enjoying psychological *re-presentations* of its world unless that agent is already in some way understood to be a subject over and against an independent world of objects, with the cognitive distance between agent and world that such an arrangement implies. This suggests that the presence of a subject-object dichotomy is necessary for agential, and thus for phenomenological, representation. But what about the reverse dependency? Is the presence of agential representation necessary for there to be a subject-object dichotomy? Perhaps in this case 'necessary' is too strong. However, where there exists a subject-object dichotomy there exists the issue of how that subject gains epistemic access to the independent world of objects that it inhabits. The issue here might be put by way of the following question: how is it that an agent is able to gain competent and appropriate epistemic access to its world, given that it is not merely distinguishing itself from that world, but distinguishing itself from that world in a particular way – that is, precisely as a subject distinguished from a collection of independent objects? Although an answer to this question may not strictly necessitate an explanation in terms of structures that stand in for or encode worldly states of affairs, that is, in terms of representations, it certainly invites one.

The first of these dependencies helps us to come to terms with the Heideggerian thought that smooth coping in the domain of the ready-to-hand has a non-representational phenomenology. Smooth coping involves a form of awareness in which there are no subjects and no objects, only the experience of the ongoing task (e.g. typing). If the presence of a subject-object dichotomy is necessary for agential representation, then the non-representational character of smooth coping is no mystery. To see the importance of the second dependency, let's think about what happens when smooth coping is disrupted and we enter the realm of un-readiness-to-hand. At this point a cognitive distance is introduced between agent and entity, a distance that may be understood as the gradual emergence of the subject-object dichotomy. Heidegger puts this in terms of the way in which, in un-readiness-to-hand, readiness-to-hand is gradually usurped by presence-at-hand. Thus the human agent is progressively revealed as a detached theoretical reasoner over and against entities revealed as context-independent objects to which that subject has access.[6] For example, Heidegger writes: '[t]he more urgently… we need what is missing, and the more authentically it is encountered in its un-readiness-to-hand, all the

[6] For how to interpret this point, see note 3 above.

15

more obtrusive… does that which is ready-to-hand become – so much so, indeed, that it seems to lose its character of readiness-to-hand. It reveals itself as just present-at-hand and no more, which cannot be budged without the thing that is missing' (Heidegger, 1927: p.103). But now where one has a subject-object dichotomy, even a minimal one, the invitation to representation kicks in, making it eminently plausible that the agent's epistemic access to the world will ultimately be secured by a representational route. This has implications for our understanding of Dreyfusian skills. As we have seen, for Dreyfus, to have a skill is to be equipped with a prior capacity to be flexibly sensitive to what is (normally) relevant in a certain context. The primary manifestation of such skills is during smooth coping in the domain of the ready-to-hand, where the associated phenomenology is non-representational in character. However, Dreyfusian skills are also manifested in certain cases of practical problem solving in the domain of the un-ready-to-hand. What the present analysis tells us is that, in this second arena, those skills will have a representational phenomenology.

So what will the representations characteristic of the un-ready-to-hand be like? We can answer this question, I believe, if we compare the domain of un-readiness-to-hand with that of presence-at-hand. When revealed as present-at-hand (e.g. by detached theoretical reflection) an entity will be experienced in terms of properties that are action-neutral, specifiable without essential reference to the representing agent, and context-independent. Moreover, according to Heidegger, this group of properties will also characterize the contents of the agent's related representational states.[7] In the domain of the un-ready-to-hand, by contrast, an entity will be experienced in terms of properties that are action-specific, egocentric and dependent on a particular context of activity. Moreover, this second group of properties will also characterize the contents of the agent's related representational states. So, in the domain of un-readiness-to-hand, we should expect to find representations that are action-specific, egocentric, and intrinsically context-dependent. And that is the profile of subagential action-oriented representation, as I have characterized it.

Now, because context is woven into the fundamental operating principles by which situated special-purpose adaptive couplings function, such mechanisms do not face the difficulties of assigning relevance or of representing massively holistic networks of contextual significance that Dreyfus highlights in his analysis of the frame problem. Even where action-oriented representations are involved, the fact that those structures function as part of a situated special-purpose adaptive coupling means that relevance is guaranteed and what is represented is not context. To the extent that what is being represented by action-oriented representations remains knowledge *that* the environment is thus and so, it is a thus and so that is encoded in action-specific and agent-relative terms (e.g. 'region-to-be-avoided over there'). Such content makes sense only against the backdrop of intrinsic context-embeddedness provided by the rest of the mechanism, so the knowledge-that in question is not doing the specific job that opens the door to Dreyfus'

---

[7] It does not follow from the fact that some of the agent's representational states have contents with these properties that the agent herself is fully decontextualized (again see note 3 above).

worries about knowledge-how and knowledge-that, that is, it is not doing the job of explaining how we are equipped with a prior capacity to be flexibly sensitive to what is (normally) relevant in a particular context.

It seems we are making headway. Unfortunately, it is time to face up to a serious difficulty. To bring this difficulty into view, we need to make a distinction between two different dimensions of the frame problem. The first, call it the *intra-context frame problem*, challenges us to say how a purely mechanistic system might achieve appropriate, flexible and fluid action *within* a context. The second, call it the *inter-context frame problem* challenges us to say how a purely mechanistic system might achieve appropriate, flexible and fluid action in worlds in which adaptation to new contexts is open-ended and in which the number of potential contexts is indeterminate. Earlier I made the point that the frame problem really bites in cases of the latter. We can now develop this point in an alternative way. In effect, I have been arguing in this section that the intra-context frame problem may be solved by a purely special-purpose mechanism, or by some suite of such mechanisms, perhaps featuring action-oriented representations, perhaps not. But from what I've said so far it remains mysterious how any collection of purely special-purpose mechanisms, whether or not they feature action-oriented representations, could ever solve the inter-context frame problem. So our new goal is to identify one or more mechanisms that might causally explain adaptive flexibility on a scale sufficient to account for open-ended adaptation to new contexts. Here is a suggestion.

Most work in connectionist cognitive science has tended to concentrate on network architectures that, in effect, limit the range and complexity of the dynamics available to such a system. Restricting features include: neat symmetrical connectivity; noise-free processing; update properties which are based either on a global, digital pseudo-clock or on methods of stochastic change; units which are uniform in structure and function; activation passes that proceed in an orderly feed-forward fashion; and a model of neurotransmission in which the effect of one neuron's activity on that of a connected neuron will simply be either excitatory or inhibitory, and will be mediated by a simple point-to-point signalling process. Quite recently, however, some researchers have come to favour a class of connectionist machines with richer system dynamics, so-called dynamical neural networks (henceforth DNNs).

What we might, for convenience, call mark-one DNNs feature the following sorts of properties: asynchronous continuous-time processing; real-valued time delays on connections; non-uniform activation functions; deliberately introduced noise; and connectivity which is not only both directionally unrestricted and highly recurrent, but also not subject to symmetry constraints (see e.g. Husbands et al., 1995). Mark-two DNNs add two further twists to the architectural story. In these networks, christened GasNets (Husbands et al., 1998), the standard DNN model is augmented with modulatory neurotransmission (according to which fundamental properties of neurons, such as their activation profiles, are transformed by arriving neurotransmitters), and models of neurotransmitters that diffuse virtually from their source in a cloud-like, rather than a point-to-point, manner, and thus affect entire volumes of processing structures. GasNets

thus provide a platform for potentially rich interactions between two interacting and intertwined dynamical mechanisms – virtual cousins of the electrical and chemical processes in real nervous systems. Diffusing 'clouds of chemicals' may change the intrinsic properties of the artificial neurons, thereby changing the patterns of 'electrical' activity, while 'electrical' activity may itself trigger 'chemical' activity. So, dropping the scare quotes, these biologically inspired machines feature neurotransmitters that may not only transform the transfer functions of the neurons on which they act, but which may do so on a grand scale, as a result of the fact that they act by gaseous diffusion through volumes of brain-space, rather than by electrical transmission along connecting neural wires.

Evolutionary algorithms have been used to design GasNet robot control systems for simple homing and discrimination tasks (Husbands et al., 1998). So what does the analysis of such machines tell us? Viewed as static wiring diagrams, many of the successful controllers appear to be rather simple structures. Typical GasNets feature a very small number of primitive visual receptors, connected to a tiny number of inner and motor neurons by just a few synaptic links. However, this apparent structural simplicity hides the fact that the dynamics of the networks are often highly complex, involving, as expected, subtle couplings between chemical and electrical processes. For example, it is common to find adaptive use being made of oscillatory dynamical sub-networks, some of whose properties (e.g., their periods) depend on spatial features of the modulation and diffusion processes, processes which are themselves determined by the changing levels of electrical activity in the neurons within the network. What seems clear is that GasNets realize a potentially powerful kind of ongoing structural fluidity, one that involves the functional reconfiguration of large networks of components. This is achieved on the basis of multiple simultaneous interactions and complex dynamic feedback loops, such that (a) the causal contribution of each systemic component partially determines, and is partially determined by, the causal contributions of large numbers of other systemic components, and, moreover, (b) those contributions may change radically over time. This is what Clark (1997) dubs *continuous reciprocal causation*, henceforth CRC. At root, then, GasNets are mechanisms of significant adaptive plasticity achieved on the basis of CRC, and although the empirical evidence is far from decisive, it seems plausible that it is this sort of system that, when harnessed and tuned appropriately by selection or learning to operate over different time-scales, may be the mechanistic basis of open-ended adaptation to new contexts.

It is worth pointing out (since it will be relevant to what follows) that, to the extent that subagential representation requires the presence of a kind of internal modularity, one in which communicating subsystems inside the agent's brain are conceptualized as trafficking in information carried by inner representational vehicles, there is a tension between CRC and action-oriented representation. Here I can provide only the barest bones of a justification for this claim. (For the full story, see Wheeler, 2005.) However, two things need to be said by way of support:

1. The representational architecture of our spatial-navigation-robot discussed earlier has exactly the modular profile just mentioned. The

EMD-array is a module that produces object-related information which, realized as the action-oriented snap map of detected obstacles, is consumed (along with additional, goal-related information arriving from the supplementary visual system) by the modular subsystem that produces the motor map. The information carried by this second action-oriented representation is then itself consumed by a further modular subsystem that transforms movement-specifications into actual physical motion.

2. CRC undermines modularity because, as CRC increases, it becomes progressively more difficult to specify distinct and robust causal-functional roles played by reliably reidentifiable parts of the system. The performance of any particular sub-task will increasingly be underpinned by larger and larger numbers of interacting components whose contributions are changing in highly context sensitive ways. Thus modular explanation, and so representational explanation, will be threatened.

So much, I think, is true. Still, what this means is that CRC and action-oriented representation cannot be present *in a single mechanism simultaneously*. Nothing rules out their adaptively beneficial co-existence in the same mechanism over time, or indeed in different mechanisms simultaneously. Intriguingly, preliminary analysis suggests that some GasNets exhibit a kind of transient modularity in which, over time, the effects of the gaseous diffusible modulators drive the network through different phases of modular and non-modular (CRC-based) organization (Husbands, personal communication).

5. Coping Trouble

In a recent volley, Dreyfus (2008) has criticized my treatment of the frame problem. At the heart of Dreyfus' critique is the accusation that I am propagating a 'cognitivist misreading of Heidegger' (p.341).[8] This is not merely a dispute over who gets to call themselves Heideggerian. Dreyfus' point is that my misreading of Heidegger leads me seriously astray in my handling of the frame problem. So what is my alleged mistake? Dreyfus objects to the way in which I locate action-oriented representation and practical problem-solving as part of the Heideggerian account. To be clear, my supposed error is not that I find *a* place for such states and processes. It is agreed on both sides (I think) that the phenomenology of un-readiness-to-hand has the character of representational problem solving. Furthermore, Dreyfus and I agree that it is an empirical question just how much of our everyday experience involves encounters with the ready-to-hand in non-representational smooth coping, as opposed to encounters with the un-ready-to-hand in practical problem solving characterized by action-oriented representations (p.346). However, what Dreyfus is keen to point out is that there exists a phenomenon that is ontologically more basic than smooth coping or action-oriented representational problem solving, a phenomenon to which I allegedly don't do justice and that is critical in any

---

[8] All page numbers in this section refer to (Dreyfus, 2008), unless otherwise noted.

genuinely Heideggerian response to the frame problem. So let's bring this phenomenon into view.

The key here is a distinction that Dreyfus draws between *skillful coping* and *background coping*. In the language that I have been using, skillful coping is another term for Dreyfusian skill, and so covers both smooth coping and the relevant examples of practical problem solving (see above). Background coping, on the other hand, is 'an even more basic nonrepresentational holistic coping that allows copers to orient themselves in the world' (p.345). In more detail:

> background coping is not a traditional kind of intentionality. Whereas the ready-to-hand has conditions of satisfaction, like hammering in the nail, background coping does not have conditions of satisfaction… The important point for Heidegger, but not for Wheeler, is that *all* coping, including unready-to-hand coping, takes place on the background of this basic non-representational, holistic, absorbed, kind of intentionality, which Heidegger calls being-in-the-world. (pp.345-6)

So, for Heidegger and for Dreyfus, background coping is the human agent's fundamental familiarity with her world that underpins both her smooth coping and her action-oriented representational problem solving. In fleshing out this idea, Dreyfus draws on Merleau-Ponty's (1962) notion of the *intentional arc*.

> According to Merleau-Ponty, as an agent acquires skills, those skills are "stored," not as representations in the agent's mind, but as the solicitations of situations in the world. What the learner acquires through experience is not represented at all but is presented to the learner as more and more finely discriminated situations. If the situation does not clearly solicit a single response or if the response does not produce a satisfactory result, the learner is led to further refine his discriminations, which, in turn, solicit ever more refined responses. For example, what we have learned from our experience of finding our way around in a city is "sedimented" in how that city looks to us. Merleau-Ponty calls this feedback loop between the embodied coper and the perceptual world the intentional arc. (p.340)

Crucially, it is precisely the phenomenon of background coping, of knowing one's way around one's world, that, for Dreyfus, promises to dissolve the frame problem, in both its intra-context and inter-context forms. On the basis of our constantly honed background know-how, we respond directly to relevance, with context-bound entities soliciting or summoning us to act in ways shaped by our past experiences. And the capacity for flexible context-switching that lies at the heart of the inter-context frame problem is explained by the fact that I can be summoned not only by the present situation, but also by other situations that, because they have been relevant in the past, lie on the horizon of my experience (p.359). Given the existence, the ontological priority, the non-representational character, and the frame-problem-busting consequences of background

coping, 'a Heideggerian Cognitive Science would require working out an ontology, phenomenology, and brain model that denies a basic role to any sort of representations' (p.347). And that (Dreyfus tells us) just about wraps it up for my 'pseudo-Heideggerian' approach in which action-oriented representational problem solving plays a key part. Taking up the challenge to provide a positive cognitive-scientific story (a 'brain model'), Dreyfus discusses at length the work of neuroscientist Walter Freeman (e.g. Freeman, 2000). What emerges is a vision of the brain as a nonlinear non-representational dynamical system, primed by past experience to actively pick up and enrich significance, one whose constantly shifting attractor landscape physically grounds Merleau-Ponty's intentional arc by causally explaining how newly encountered significances change the whole perceptual world of the agent. (For the details, see pp.347-57.)

Now, I agree with Dreyfus that background coping is the phenomenological structure on the basis of which our context sensitive activity is possible. Moreover, I plead guilty to not having paid sufficient attention to the phenomenon or its mechanistic basis in my account previously. So let me put that right. As Dreyfus himself observes, the causal processes realized in the neurodynamical brain described by Freeman, the causal processes that, Dreyfus has argued, underlie background coping, are plausibly interpreted as an instantiation of CRC. Recall, CRC is causation that involves multiple simultaneous interactions and complex dynamic feedback loops, such that (a) the causal contribution of each systemic component partially determines, and is partially determined by, the causal contributions of large numbers of other systemic components, and, moreover, (b) those contributions may change radically over time. Compare this specification with Freeman's own description of the brain's dynamics.

> I have observed that brain activity patterns are constantly dissolving, reforming and changing, particularly in relation to one another. When an animal learns to respond to a new odor, there is a shift in all other patterns, even if they are not directly involved with the learning.

> I conclude that context dependence is an essential property of the cerebral memory system, in which each new experience must change all of the existing store by some small amount, in order that a new entry be incorporated and fully deployed in the existing body of experience. (Freeman, 2000: p.22; quoted by Dreyfus, pp.352-3).

So Dreyfus and I are in fundamental agreement here. Moreover, I have argued that CRC is a non-representational process, so to the extent that CRC is the mechanistic basis of background coping, I also agree with Dreyfus that the cognitive science of that phenomenon will be one that denies a basic role to any sort of representation.

So where exactly is the disagreement? I have argued that CRC, when harnessed and tuned appropriately by selection or learning to operate over different time-scales, may well be the mechanistic basis of open-ended adaptation to new contexts. In other words, CRC causally explains fluid and flexible context-switching. And this is where the trouble starts. Dreyfus argues that, in the end, I remain problematically ambivalent about which

21

model of cognitive mechanism, CRC or action-oriented representation, is ontologically more basic (p.347). That's because he thinks that *all* background coping, and not just the aspect that underlies fluid and flexible context-switching, must be explained by CRC. Put another way, Dreyfus thinks that it's CRC that provides a causal dissolution of the intra-context frame problem and not 'just' the inter-context version. And that, in part, is because he thinks that representations are a source of the frame problem. As he puts it: 'Wheeler's own proposal… by introducing flexible action-oriented representations, like any representational approach has to face the frame problem head on' (p.358). But, as far as I can see, not all forms of representation necessarily usher in the frame problem. The intra-context frame problem dissolves in the face of situated special-purpose adaptive couplings that use regular sensing of the environment, rather than detailed world models, to guide their behaviour. And Franceschini et al.'s robot, discussed earlier, shows that action-oriented representations can figure in such mechanisms. So does this mean that I am ambivalent about whether action-oriented representation or CRC is more basic? No it doesn't. To see why, we need to be clear that at root it is not action-oriented representation that dissolves the intra-context frame problem, but rather the presence of situated special-purpose adaptive couplings. Representations may figure in such mechanisms, but they may not. Since CRC plays a critical role in dissolving the inter-context frame problem, but action-oriented representation plays no equivalent role, there is a clear sense in which CRC is the ontologically more basic phenomenon.[9]

If I am ambivalent about anything here, it's about whether CRC or situated special-purpose adaptive coupling is ontologically more basic. Since this is where the real disagreement between Dreyfus and me lies, let me finish with a few words on why I take this view. Background coping should not be thought of as some sort of distinct phenomenological 'module' to which smooth coping or practical problem solving may or may not be added. Rather, background coping is exhibited in, by being a structural precondition for, these context-embedded activities. Relatedly Heidegger (1927) draws a distinction between, on the one hand, the world and, on the other, the worldhood of the world. The former indicates the holistic semantic networks of context-defining involvements that we inhabit, the latter the abstract referential network structure that is shared by all concrete totalities of involvements. Similarly, the mechanistic basis of background coping constitutes a precondition, a *causal-structural* one, for the mechanisms that actually underlie smooth coping and practical problem solving. In fact the mechanistic basis of background coping has the dual character of CRC and situated special-purpose adaptive coupling. CRC is the causal-structural basis of fluid and flexible context-switching. Situated special-purpose adaptive coupling is the causal-structural

---

[9] In a personal communication with Dreyfus which he quotes (p.346), I unfortunately muddied the waters by making the error of treating CRC as the causal basis of the ready-to-hand and action-oriented representation as the causal basis of the un-ready-to-hand. Of course, this was a mistake. The right thing for me to say is that these domains are distinguished by the non-representational or representational status of the situated special-purpose adaptive couplings involved. CRC will explain context-switching in either of the domains.

basis of intra-context sensitivity to relevance. Both are structural phenomena that are realized in particular ways by certain concrete mechanisms (GasNets, action-oriented maps, and so on). Both are, to use a term from Heidegger, equiprimordial. They are also, I believe, the mechanistic wellsprings of our resistance to the frame problem.

Acknowledgments

References

Brooks, R. A. (1991) 'Intelligence Without Reason', in *Proceedings of 12th International Joint Conference on Artificial Intelligence*, San Mateo, California: Morgan Kauffman, 569-95.

Clark, A. (1997) *Being There: Putting Brain, Body, And World Together Again*, Cambridge, Mass.: MIT Press.

Cliff, D. (1994) 'AI and A-Life: Never Mind the Blocksworld', *Artificial Intelligence and Simulation of Behaviour Quarterly*, 87: 16-21.

Derrida, J. (1989) *Of Spirit – Heidegger and the Question*, trans. G. Bennington and R. Bowlby, Chicago and London: University of Chicago Press.

Dreyfus, H. L. (1990) *Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I*, Cambridge, Mass.: MIT Press.

Dreyfus, H. L., (1992) *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge, Mass.: MIT Press.

Dreyfus, H. L. (2008) 'Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian', in P. Husbands, O. Holland, and M. Wheeler (eds), *The Mechanical Mind in History*, Cambridge, Mass.: MIT Press, pp.331-71. (A shortened version of this paper appears in under the same title in *Philosophical Psychology* 20/2 (2007), 247-268. Another version appears under the same title in *Artificial Intelligence*, 171 (2007): 1137-1160. I have worked with the Husbands et al. (eds) version of the text, to which the cited page numbers refer.)

Dreyfus H. L. and Dreyfus S. E. (1988) 'Making a Mind versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint'. *Daedalus*, 117(1): 15-44.

Fodor, J. A. (1983) *The Modularity of Mind*, Cambridge, Mass.: MIT Press.

Fodor, J. A. (1987) 'Modules, Frames, Fridgeons, Sleeping Dogs, and the Music of the Spheres', in (Pylyshyn, 1987), 139-50.

Franceschini, N., Pichon, J. M., and Blanes, C. (1992) 'From Insect Vision to Robot Vision, *Philosophical Transactions of the Royal Society, series B*, 337: 283-94.

Freeman, W. (2000) *How Brains Make Up Their Minds*, Columbia University Press, New York.

Gallagher, S. (2007) 'Review of Michael Wheeler's *Reconstructing the Cognitive World*', *Mind*, 116: 792-6.

Genesereth, M. R. and Nilsson, N. J. (1987) *The Logical Foundations of Artificial Intelligence*, Los Altos: Morgan Kauffman.

Heidegger, M. (1927) *Being and Time*, trans. J. Macquarrie and E. Robinson. Oxford, Basil Blackwell. Translation published in 1962.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., and Thagard, P. T. (1986) *Induction: Processes of Inference, Learning and Discovery*, Cambridge, Mass.: MIT Press.

Horgan, T. and Tienson, J. (1994) 'A Nonclassical Framework for Cognitive Science', *Synthese*, 101: 305-45.

Husbands, P., Harvey, I., and Cliff, D. (1995), 'Circle in the Round: State Space Attractors for Evolved Sighted Robots', *Robotics and Autonomous Systems*, 15: 83-106.

Husbands, P., Smith, T., Jakobi, N., and O'Shea, M. (1998) 'Better Living through Chemistry: Evolving GasNets for Robot Control, *Connection Science*, 10(3/4): 185-210.

Kuhn, T. S. (1970), *The Structure of Scientific Revolutions* (second edition, enlarged), Chicago, University of Chicago Press.

Marr, D. (1982) *Vision: a Computational Investigation into the Human Representation and Processing of Visual Information,* New York, W.H. Freeman.

McDowell, J. (1994) 'The Content of Perceptual Experience', *The Philosophical Quarterly*, 44(175): 190-205.

Merleau-Ponty, M. (1962) *Phenomenology of Perception*, trans. C. Smith, London., Routledge and Kegan Paul.

Pylyshyn, Z. (ed.) (1987) *The Robot's Dilemma*, Norwood, NJ, Ablex.

Rubinsetin, A. (1998) *Modeling Bounded Rationality*, Cambridge, Mass.: MIT Press.

Shanahan, M. (2006) 'The Frame Problem', in *The Stanford Encyclopedia of Philosophy (Spring 2006 Edition)*, E. N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2006/entries/frame-problem/>.

Simon, H. (1955) 'A Behavioral Model of Rational Choice', *The Quarterly Journal of Economics,* 64**:** 99-118.

Webb, B. (1993) 'Modeling Biological Behaviour or 'Dumb Animals and Stupid Robots', in *Pre-Proceedings of the Second European Conference on Artificial Life*, 1090-103.

Webb, B. (1994) 'Robotic Experiments in Cricket Phonotaxis', in Cliff, D., Husbands, P., Meyer, J.-A., and Wilson, S.W. (eds), *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*, Cambridge, Mass.: MIT Press, 45-54.

Wheeler, M. (1995) 'Escaping from the Cartesian Mind-Set: Heidegger and Artificial Life', in F. Moran, A. Moreno, J. Merelo and P, Chacon (eds), *Advances in Artificial Life: Proceedings of the Third European Conference on Artificial Life*, Berlin and Heidelberg: Springer-Verlag.

Wheeler, M. (2005) *Reconstructing The Cognitive World: The Next Step*, Cambridge, Mass.: MIT Press.

Wheeler, M. (2008) 'God's Machines: Descartes on the Mechanization of Mind', in P. Husbands, O. Holland, and M. Wheeler (eds), *The Mechanical Mind in History*, Cambridge, Mass.: MIT Press, pp.307-330.