Available online at www.sciencedirect.com
ScienceDirect
Artificial Intelligence 171 (2007) 1137–1160
Artificial Intelligence
www.elsevier.com/locate/artint

Why Heideggerian AI failed and how fixing it would require making it more Heideggerian ☆

Hubert L. Dreyfus

University of California, Department of Philosophy, 314 Moses Hall 2390, Berkeley, CA 94720-2390, USA

Available online 10 October 2007

1. The convergence of computers and philosophy

When I was teaching at MIT in the early sixties, students from the Artificial Intelligence Laboratory would come to my Heidegger course and say in effect: "You philosophers have been reflecting in your armchairs for over 2000 years and you still don't understand how the mind works. We in the AI Lab have taken over and are succeeding where you philosophers have failed. We are now programming computers to exhibit human intelligence: to solve problems, to understand natural language, to perceive, and to learn."[1] In 1968 Marvin Minsky, head of the AI lab, proclaimed: "Within a generation we will have intelligent computers like HAL in the film, *2001*."[2]

As luck would have it, in 1963, I was invited by the RAND Corporation to evaluate the pioneering work of Alan Newell and Herbert Simon in a new field called Cognitive Simulation (CS). Newell and Simon claimed that both digital computers and the human mind could be understood as *physical symbol systems*, using strings of bits or streams of neuron pulses as symbols representing the external world. Intelligence, they claimed, merely required making the appropriate inferences from these internal representations. As they put it: "A physical symbol system has the necessary and sufficient means for general intelligent action."[3]

As I studied the RAND papers and memos, I found to my surprise that, far from replacing philosophy, the pioneers in CS had learned a lot, directly and indirectly from the philosophers. They had taken over Hobbes' claim that reasoning was calculating, Descartes' mental representations, Leibniz's idea of a "universal characteristic"—a set of primitives in which all knowledge could be expressed,—Kant's claim that concepts were rules, Frege's formalization of such rules, and Russell's postulation of logical atoms as the building blocks of reality. In short, without realizing it, AI researchers were hard at work turning rationalist philosophy into a research program.

At the same time, I began to suspect that the critical insights formulated in existentialist armchairs, especially Heidegger's and Merleau-Ponty's, were bad news for those working in AI laboratories—that, by combining rational-

☆ The essence of this article will appear in the book "The Mechanization of Mind in History", MIT Press.
E-mail address: dreyfus@cogsci.berkeley.edu.

[1] This isn't just my impression. Philip Agre, a PhD student at the AI Lab at that time, later wrote:

> I have heard expressed many versions of the propositions ... that philosophy is a matter of mere thinking whereas technology is a matter of real doing, and that philosophy consequently can be understood only as deficient.

Philip E. Agre, *Computation and Human Experience* (Cambridge: Cambridge University Press, 1997), 239.

[2] Marvin Minsky as quoted in a 1968 MGM press release for Stanley Kubrick's *2001: A Space Odyssey*.

[3] Newell, A. and Simon, H.A., Computer Science as Empirical Inquiry: Symbols and Search, *Mind Design*, John Haugeland, Edt. (Cambridge, MA, MIT Press), 1988.

0004-3702/$ – see front matter
doi:10.1016/j.artint.2007.10.012

ism, representationalism, conceptualism, formalism, and logical atomism into a research program, AI researchers had condemned their enterprise to reenact a failure.

2. Symbolic AI as a degenerating research program

Using Heidegger as a guide, I began to look for signs that the whole AI research program was degenerating. I was particularly struck by the fact that, among other troubles, researchers were running up against the problem of representing significance and relevance—a problem that Heidegger saw was implicit in Descartes' understanding of the world as a set of meaningless facts to which the mind assigned what Descartes called values, and John Searle now calls functions.[4]

But, Heidegger warned, values are just more meaningless facts. To say a hammer has the function of being for hammering leaves out the defining relation of hammers to nails and other equipment, to the point of building things, and to the skills required when actually using the hammer—all of which reveal the way of being of the hammer which Heidegger called *readiness-to-hand*. Merely assigning formal function predicates to brute facts such as hammers couldn't capture the hammer's way of being nor the meaningful organization of the everyday world in which hammering has its place. "[B]y taking refuge in 'value'-characteristics," Heidegger said, "we are ... far from even catching a glimpse of being as readiness-to-hand."[5]

Minsky, unaware of Heidegger's critique, was convinced that representing a few million facts about objects including their functions, would solve what had come to be called the commonsense knowledge problem. It seemed to me, however, that the deep problem wasn't storing millions of facts; it was knowing which facts were relevant in any given situation. One version of this relevance problem was called "the frame problem." If the computer is running a representation of the current state of the world and something in the world changes, how does the program determine which of its represented facts can be assumed to have stayed the same, and which would have to be updated?

As Michael Wheeler in his recent book, *Reconstructing the Cognitive World*, puts it:

> [G]iven a dynamically changing world, how is a nonmagical system ... to take account of those state changes in that world ... that matter, and those unchanged states in that world that matter, while ignoring those that do not? And how is that system to retrieve and (if necessary) to revise, out of all the beliefs that it possesses, just those beliefs that are relevant in some particular context of action?[6]

Minsky suggested that, to avoid the frame problem, AI programmers could use what he called frames—descriptions of typical situations like going to a birthday party—to list and organize those, and only those, facts that were normally relevant. Perhaps influenced by a computer science student who had taken my phenomenology course, Minsky suggested a structure of essential features and default assignments—a structure Husserl had already proposed and already called a frame.[7]

But a system of frames isn't *in* a situation, so in order to select the possibly relevant facts in the current situation one would need frames for recognizing situations like birthday parties, and for telling them from other situations such as ordering in a restaurant. But how, I wondered, could the computer select from the supposed millions of frames in its memory the relevant frame for selecting the birthday party frame as the relevant frame, so as to see the current relevance of, say, an exchange of gifts rather than money? It seemed to me obvious that any AI program using frames to organize millions of meaningless facts so as to retrieve the currently relevant ones was going to be caught in a

[4] John R. Searle, *The Construction of Social Reality* (New York: The Free Press, 1995).

[5] Martin Heidegger, *Being and Time*, J. Macquarrie & E. Robinson, Trans. (New York: Harper & Row, 1962), 132–133.

[6] Michael Wheeler, *Reconstructing the Cognitive World: The Next Step* (Cambridge, MA: A Bradford Book, The MIT Press, 2007), 179.

[7] Edmund Husserl, *Experience and Judgment* (Evanston: Northwestern University Press, 1973), 38.

To do the same job, Roger Schank proposed what he called *scripts* such as a restaurant script. "A script," he wrote, "is a structure that describes appropriate sequences of events in a particular context. A script is made up of slots and requirements about what can fill those slots. The structure is an interconnected whole, and what is in one slot affects what can be in another. A script is a predetermined, stereotyped sequence of actions that defines a well-known situation." R.C. Schank and R.P. Abelson, *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures* (Hillsdale, NJ: Lawrence Erlbaum, 1977) 41. Quoted in: *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, John Preston and Mark Bishop, Eds (Oxford: Clarendon Press, 2002).

regress of frames for recognizing relevant frames for recognizing relevant facts, and that, therefore, the frame problem wasn't just a problem but was a sign that something was seriously wrong with the whole approach.

Unfortunately, what has always distinguished AI research from a science is its refusal to face up to and learn from its failures. In the case of the relevance problem, the AI programmers at MIT in the sixties and early seventies limited their programs to what they called micro-worlds—artificial situations in which the small number of features that were possibly relevant was determined beforehand. Since this approach obviously avoided the real-world frame problem, MIT PhD students were compelled to claim in their theses that their micro-worlds could be made more realistic, and that the techniques they introduced could be generalized to cover commonsense knowledge. There were, however, no successful follow-ups.[8]

The work of Terry Winograd is the best of the work done during the micro-world period. His "blocks-world" program, SHRDLU, responded to commands in ordinary English instructing a virtual robot arm to move blocks displayed on a computer screen. It was the parade case of a micro-world program that really worked—but of course only in its micro-world. So to produce the expected generalization of his techniques, Winograd started working on a new Knowledge Representation Language, (KRL). His group, he said, was "concerned with developing a formalism, or 'representation,' with which to describe ... knowledge." And he added: "We seek the 'atoms' and 'particles' of which it is built, and the 'forces' that act on it."[9]

But this approach wasn't working. Indeed, Minsky has recently acknowledged in *Wired Magazine* that AI has been brain dead since the early 70ies when it encountered the problem of commonsense knowledge.[10] Winograd, however, unlike his colleagues, was scientific enough to try to figure out what had gone wrong. So in the mid 70ies we began having weekly lunches to discuss his problems in a broader philosophical context. Looking back, Winograd says: "My own work in computer science is greatly influenced by conversations with Dreyfus."[11]

After a year of such conversations, and after reading the relevant texts of the existential phenomenologists, Winograd abandoned work on KRL and began including Heidegger in his Computer Science courses at Stanford. In so doing, he became the first high-profile deserter from what was, indeed, becoming a degenerating research program. John Haugeland now refers to the symbolic AI of that period as Good Old Fashioned AI—GOFAI for short—and that name has been widely accepted as capturing its current status. Indeed, Michael Wheeler argues that a new paradigm is already taking shape. He maintains:

> [A] Heideggerian cognitive science is ... emerging right now, in the laboratories and offices around the world where embodied-embedded thinking is under active investigation and development.[12]

Wheeler's well informed book could not have been more timely since there are now at least three versions of supposedly Heideggerian AI that might be thought of as articulating a new paradigm for the field: Rodney Brooks' behaviorist approach at MIT, Phil Agre's pragmatist model, and Walter Freeman's neurodynamic model. All three approaches implicitly accept Heidegger's critique of Cartesian internalist representations, and, embrace John Haugeland's slogan that cognition is embedded and embodied.[13]

---

[8] After I published, *What Computers Can't Do* in 1972 and pointed out this difficulty among many others, my MIT computer colleagues, rather than facing my criticism, tried to keep me from getting tenure on the grounds that my affiliation with MIT would give undeserved credibility to my "fallacies," and so would prevent the AI Lab from continuing to receive research grants from the Defense Department.

The AI researchers were right to worry. I was considering hiring an actor to impersonate an officer from DARPA having lunch with me at the MIT Faculty Club. (A plan cut short when Jerry Wiesner, the President of MIT, after consulting with Harvard and Russian computer scientists, and reading my book himself, personally granted me tenure.) I did, however, later get called to Washington by DARPA to give my views, and the AI Lab did loose DARPA support during what has come to be called the AI Winter.

[9] Winograd, T. (1976). Artificial Intelligence and Language Comprehension, in *Artificial Intelligence and Language Comprehension*, Washington, D.C.: National Institute of Education, 9.

[10] *Wired Magazine*, Issue 11:08, August 2003.

[11] *Heidegger, Coping, and Cognitive Science*, *Essays in Honor of Hubert L. Dreyfus*, Vol. 2, Mark Wrathall Ed. (Cambridge, MA: The MIT Press, 2000), iii.

[12] Michael Wheeler, *Reconstructing the Cognitive World*, 285.

[13] John Haugeland, "Mind embodied and embedded," *Having Thought: Essays in the Metaphysics of Mind* (Cambridge, MA: Harvard University Press, 1998), 218.

3. Heideggerian AI, stage one: Eliminating representations by building behavior-based robots

Winograd sums up what happened at MIT after he left for Stanford:

> For those who have followed the history of artificial intelligence, it is ironic that [the MIT] laboratory should become a cradle of "Heideggerian AI." It was at MIT that Dreyfus first formulated his critique, and, for twenty years, the intellectual atmosphere in the AI Lab was overtly hostile to recognizing the implications of what he said. Nevertheless, some of the work now being done at that laboratory seems to have been affected by Heidegger and Dreyfus.[14]

Here's how it happened. In March 1986, the MIT AI Lab under its new director, Patrick Winston, reversed Minsky's attitude toward me and allowed, if not encouraged, several graduate students, led by Phil Agre and John Batali, to invite me to give a talk.[15] I called the talk, "Why AI Researchers should study *Being and Time*." In my talk I repeated what I had written in 1972 in *What Computers Can't Do*: "[T] he meaningful objects ... among which we live are not a *model* of the world stored in our mind or brain; *they are the world itself*."[16] And I quoted approvingly a Stanford Research Institute report that, "It turned out to be very difficult to reproduce in an internal representation for a computer the necessary richness of environment that would give rise to interesting behavior by a highly adaptive robot,"[17] and concluded that "this problem is avoided by human beings because their model of the world is the world itself."[18]

The year of my talk, Rodney Brooks, who had moved from Stanford to MIT, published a paper criticizing the GOFAI robots that used representations of the world and problem solving techniques to plan their movements. He reported that, based on the idea that "the best model of the world is the world itself," he had "developed a different approach in which a mobile robot uses the world itself as its own representation—continually referring to its sensors rather than to an internal world model."[19] Looking back at the frame problem, he writes:

> And why could my simulated robot handle it? Because it was using the world as its own model. It never referred to an internal description of the world that would quickly get out of date if anything in the real world moved.[20]

Brooks's approach is an important advance, but Brooks's robots respond only to *fixed isolable features* of the environment, not to context or changing significance. Moreover, they do not learn. They are like ants, and Brooks

---

[14] Terry Winograd, "Heidegger and he Design of Computer Systems," talk delivered at Applied Heidegger Conference, Berkeley, CA, Sept. 1989. Cited in H. Dreyfus, *What Computers Still Can't Do*, Introduction to the MIT Press edition, xxxi.

[15] Not everyone was pleased. One of the graduate students responsible for the invitation reported to me: "After it was announced that you were giving the talk, Marvin Minsky came into my office and shouted at me for 10 minutes or so for inviting you."

[16] Hubert Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason.* MIT Press, 1992, 265–266.

[17] Ibid, 300.

[18] Ibid.

[19] Rodney A. Brooks. "Intelligence without Representation," *Mind Design*, John Haugeland, Ed. (The MIT Press, 1988), 416. (Brooks's paper was published in 1986.) Haugeland explains Brooks's breakthrough using as an example Brooks's robot, Herbert:

> Brooks uses what he calls "subsumption architecture", according to which systems are decomposed not in the familiar way by local functions or faculties, but rather by global *activities* or *tasks*. ... Thus, Herbert has one subsystem for detecting and avoiding obstacles in its path, another for wandering around, a third for finding distant soda cans and homing in on them, a fourth for noticing nearby soda cans and putting its hand around them, a fifth for detecting something between its fingers and closing them, and so on... fourteen in all. What's striking is that these are all complete input/output systems, more or less independent of each other. (John Haugeland, *Having Thought: Essays in the Metaphysics of Mind* (Cambridge, MA: Harvard University Press, 1998), 218.)

[20] Brooks gives me credit for "being right about many issues such as the way in which people operate in the world is intimately coupled to the existence of their body," (Ibid. 42.) but he denies the direct influence of Heidegger:

> In some circles, much credence is given to Heidegger as one who understood the dynamics of existence. Our approach has certain similarities to work inspired by this German philosopher (for instance, Agre and Chapman 1987) but our work was not so inspired. It is based purely on engineering considerations. ("Intelligence without Representation," 415). [Rodney A. Brooks, *Flesh and Machines: How Robots Will Change Us*, Vintage Books (2002), 168.]

aptly calls them “animats.” Brooks thinks he does not need to worry about learning, putting it off as a concern for possible future research.[21] But by operating in a fixed world and responding only to the small set of possibly relevant features that their receptors can pick up, Brooks’ animats beg the question of changing relevance and so finesse rather than solve the frame problem.

Still, Brooks comes close to an existential insight spelled out by Merleau-Ponty, viz. that intelligence is founded on and presupposes the more basic way of coping we share with animals, when he says:[22]

> The “simple” things concerning perception and mobility in a dynamic environment ... are a necessary basis for “higher-level” intellect. ... Therefore, I proposed looking at simpler animals as a bottom-up model for building intelligence. It is soon apparent, when “reasoning” is stripped away as the prime component of a robot’s intellect, that the dynamics of the interaction of the robot and its environment are primary determinants of the structure of its intelligence.[23]

Brooks is realistic in describing his ambitions and his successes:

> The work can best be described as attempts to emulate insect-level locomotion and navigation. ... There have been some behavior-based attempts at exploring social interactions, but these too have been modeled after the sorts of social interactions we see in insects.[24]

Surprisingly, the modesty Brooks exhibited in choosing to first construct simple insect-like devices did not deter Brooks and Daniel Dennett from repeating the extravagant optimism characteristic of AI researchers in the sixties. As in the days of GOFAI, on the basis of Brooks’ success with insect-like devices, instead of trying to make, say, an artificial spider, Brooks and Dennett decided to leap ahead and build a humanoid robot. As Dennett explained in a 1994 report to The Royal Society of London:

> A team at MIT of which I am a part is now embarking on a long-term project to design and build a humanoid robot, Cog, whose cognitive talents will include speech, eye-coordinated manipulation of objects, and a host of self-protective, self-regulatory and self-exploring activities.[25]

Dennett seems to reduce this project to a joke when he adds in all seriousness: “While we are at it, we might as well try to make Cog crave human praise and company and even exhibit a sense of humor.”[26]

Of course, the “long term project” was short lived. Cog failed to achieve any of its goals and the original robot is already in a museum.[27] But, as far as I know, neither Dennett nor anyone connected with the project has published an account of the failure and asked what mistaken assumptions underlay their absurd optimism. In a personal communication Dennett blamed the failure on a lack of graduate students and claimed that:

> Progress was being made on all the goals, but slower than had been anticipated.[28]

If progress was actually being made, however, the graduate students wouldn’t have left, or others would have continued to work on the project. Clearly some specific assumptions must have been mistaken, but all we find in Dennett’s assessment is the implicit assumption that human intelligence is on a continuum with insect intelligence,

---

[21] “Can higher-level functions such as learning occur in these fixed topology networks of simple finite state machines?” he asks. But he offers no response. (“Intelligence without Representation,” *Mind Design*, 420.)

[22] See, Maurice Merleau-Ponty, *The Structure of Behavior*, A.L. Fisher, Trans. (Boston: Beacon Press, 2nd edition 1966).

[23] Brooks, “Intelligence without representation”, 418.

[24] Brooks, From earwigs to humans, *Robotics and Autonomous Systems*, vol. 20, 1997, 291.

[25] Daniel Dennett, The practical requirements for making a conscious robot, *Philosophical Transactions of the Royal Society of London*, A, v. 349, 1994, 133–146.

[26] Ibid. 133.

[27] Although, as of going to press in 2007, you couldn’t tell it from the Cog web page. (www.ai.mit.edu/projects/humanoid-robotics-group/cog/.)

[28] Private communication. Oct. 26, 2005.

and that therefore adding a bit of complexity to what has already been done with animats counts as progress toward humanoid intelligence. At the beginning of AI research, Yehoshua Bar-Hillel called this way of thinking the first-step fallacy, and my brother at RAND quipped, "It's like claiming that the first monkey that climbed a tree was making progress towards flight to the moon."

In contrast to Dennett's assessment, Brooks is prepared to entertain the possibility that he is barking up the wrong tree. He soberly comments that:

> Perhaps there is a way of looking at biological systems that will illuminate an inherent necessity in some aspect of the interactions of their parts that is completely missing from our artificial systems .... I am not suggesting that we need go outside the current realms of mathematics, physics, chemistry, or biochemistry. Rather I am suggesting that perhaps at this point we simply do not *get it*, and that there is some fundamental change necessary in our thinking in order that we might build artificial systems that have the levels of intelligence, emotional interactions, long term stability and autonomy, and general robustness that we might expect of biological systems.[29]

We can already see that Heidegger and Merleau-Ponty would say that, in spite of the breakthrough of giving up internal symbolic representations, Brooks, indeed, doesn't get it—that what AI researchers have to face and understand is not only why our everyday coping couldn't be understood in terms of inferences from symbolic representations, as Minsky's intellectualist approach assumed, but also why it can't be understood in terms of responses caused by fixed features of the environment, as in Brooks' empiricist model. AI researchers need to consider the possibility that embodied beings like us take as input energy from the physical universe, and respond in such a way as to open themselves to a world organized in terms of their needs, interests, and bodily capacities without their *minds* needing to impose meaning on a meaningless given, as Minsky's frames require, nor their *brains* converting stimulus input into reflex responses, as in Brooks's animats.

Later I'll suggest that Walter Freeman's neurodynamics offers a radically new basis for a Heideggerian approach to human intelligence—an approach compatible with physics and grounded in the neuroscience of perception and action. But first we need to examine another approach to AI contemporaneous with Brooks' that actually calls itself Heideggerian.

4. Heideggerian AI, stage 2: Programming the ready-to-hand

In my talk at the MIT AI Lab, I introduced Heidegger's non-representational account of the absorption of Dasein (human being) in the world. I also explained that Heidegger distinguished two modes of being: the *readiness-to-hand* of equipment when we are involved in using it, and the *presence-at-hand* of objects when we contemplate them. Out of that explanation and the lively discussion that followed, grew the second type of Heideggerian AI—the first to acknowledge its lineage.

This new approach took the form of Phil Agre's and David Chapman's program, *Pengi*, which guided a virtual agent playing a computer game called Pengo, in which the player and penguins kick large and deadly blocks of ice at each other.[30] Their approach, which they called "interactionism," was more self-consciously Heideggerian than Brooks's, in that they attempted to capture what Agre called "Heidegger's account of everyday routine activities."[31] In his book, *Computation and Human Experience*, Agre takes up where my talk left off:

> I believe that people are intimately involved in the world around them and that the epistemological isolation that Descartes took for granted is untenable. This position has been argued at great length by philosophers such as Heidegger and Merleau-Ponty; I wish to argue it technologically.[32]

---

29 Brooks, From earwigs to humans, 301. (The missing idea may well be Walter Freeman's. See below.)

30 Philip E. Agre, *The Dynamic Structure of Everyday Life*, MIT AI Technical Report 1085, October 1988, chapter 1, Section A1a, 9.

31 Agre, *Computation and Human Experience*, 243. His ambitious goal was to "develop an alternative to the representational theory of intentionality, beginning with the phenomenological intuition that everyday routine activities are founded in habitual, embodied ways of interacting with people, places, and things in he world."

32 Ibid, xi.

Agre's interesting new idea is that the world of Pengo in which the Pengi agent acts is made up, not of present-at-hand objects with properties, but of possibilities for action that trigger appropriate responses from the agent. To program this situated approach, Agre used what he called "deictic representations." He tells us:

> This proposal is based on a rough analogy with Heidegger's analysis of everyday intentionality in Division I of *Being and Time*, with objective intentionality corresponding to the present-at-hand and deictic intentionality corresponding to the ready-to-hand.[33]

And he explains:

> [Deictic representations] designate, not a particular object in the world, but rather a role that an object might play in a certain time-extended pattern of interaction between an agent and its environment.[34]

Looking back on my talk at MIT and rereading Agre's book I now see that, in a way, Agre understood Heidegger's account of readiness-to-hand better than I did at the time. I thought of the ready-to-hand as a special class of *entities*, viz. equipment, whereas the Pengi program treats what the agent responds to purely as *functions*. For Heidegger and Agre the ready-to-hand is not a *what* but a *for-what*.[35] But not just that the hammer is for hammering. As Agre saw, Heidegger wants to get at something more basic than simply a class of objects defined by their use. At his best Heidegger would, I think, deny that a hammer in a drawer has readiness-to-hand as its way of being. Rather, he sees that, *for the user*, equipment is encountered as *a solicitation to act*, not *an entity* with a function feature. He notes that: "When one is wholly devoted to something and 'really' busies oneself with it, one does not do so just alongside the work itself, or alongside the tool, or alongside both of them 'together'."[36] And he adds: "the peculiarity of what is proximally ready-to-hand is that, in its readiness-to-hand, it must, as it were, withdraw in order to be ready-to-hand quite authentically."[37]

As usual with Heidegger, we must ask: What is the phenomenon he is pointing out? In this case he wants us to see that, to observe our hammer or to observe ourselves hammering undermines our skillful coping. We can and do observe our surroundings while we cope, and sometimes, if we are learning, monitoring our performance as we learn improves our performance in the long run, but in the short run such attention interferes with our performance. For example, while biking we can observe passers by, or think about philosophy, but if we start observing how we skillfully stay balanced, we risk falling over.

Heidegger struggles to describe the basic way we are drawn in by the ready-to-hand. The Gestaltists would later talk of "*solicitations*." In *Phenomenology of Perception* Merleau-Ponty speaks of "*motivations*" and later, of "*the flesh*." All these terms point at what is not objectifyable—a situation's way of directly drawing from one a response that is neither caused like a reflex, nor done for a reason.

---

33 Ibid. 332.

34 Ibid. 251. As Beth Preston sums it up in her paper, "Heidegger and Artificial Intelligence:" *Philosophy and Phenomenological Research* 53 (1), March 1993: 43–69:

> What results is a system that represents the world not as a set of objects with properties, but as current functions (what Heidegger called in-order-tos). Thus, to take a Heideggerian example, I experience a hammer I am using not as an object with properties but as an in-order-to-drive-in-this-nail.

35 Heidegger himself is not always clear about the status of the ready-to-hand. When he is stressing the holism of equipmental relations, he thinks of the ready-to-hand as equipment, and of equipment as things like lamps, tables, doors, and rooms that have a place in a whole nexus of other equipment. Furthermore, he holds that breakdown reveals that these interdefined pieces of equipment are made of present-at-hand stuff that was there all along. (*Being and Time*, 97.) At one point Heidegger even goes so far as to include the ready-to-hand under the categories that characterize the present-at-hand:

> We call '*categories*'—characteristics of being for entities whose character is not that of Dasein. ... Any entity is either a "*who*" (existence) or a what (present-at-hand in the broadest sense.) *Being and Time* 70.

36 Heidegger, *Being and Time*, 405.

37 Ibid. 99.

In his 1925 course, *Logic: The Question of Truth* Heidegger describes our most basic experience of what he later calls "pressing into possibilities" not as dealing with the desk, the door, the lamp, the chair and so forth, but as directly responding to a "what for":

> What is first of all 'given' ... is the 'for writing,' the 'for going in and out,' the 'for illuminating,' the 'for sitting.' That is, writing, going-in-and-out, sitting, and the like are what we are a priori involved with. What we know when we 'know our way around' and what we learn are these 'for-what's.'[38]

It's clear here that, in spite of what some interpreters take Heidegger to be suggesting in *Being and Time*, this basic experience has no *as-structure*.[39] That is, when absorbed in coping, I can be described *objectively* as using a certain door *as* a door, but I'm not *experiencing* the door *as* a door. Normally there is no "I" and no experiencing of the door at all but simply pressing into the possibility of going out. The important thing to realize is that, when we are pressing into possibilities, there is no *experience* of an *entity* doing the soliciting; just the immediate response to a solicitation. (When solicitations don't pan out, what then is disclosed is the world of interconnected equipment, and I can then step back and perceive things *as* things, and act for reasons.[40])

But Agre's Heideggerian AI did not try to program this experiential aspect of being drawn in by a solicitation. Rather, with his deictic representations, Agre *objectified* both the functions and their situational relevance for the agent. In Pengi, when a virtual ice cube defined by its function is close to the virtual player, a rule dictates a response, e.g. kick it. No skill is involved and no learning takes place.

So Agre had something right that I was missing—the transparency of the ready-to-hand—but he nonetheless fell short of programming a Heideggerian account of everyday routine activities. For Heidegger, the ready-to-hand is not a fixed function, encountered in a predefined type of situation that triggers a predetermined response that either succeeds or fails. Rather, as we have begun to see and will soon see further, readiness-to-hand is experienced as a *solicitation* that calls forth a *flexible response* to the *significance* of the current situation—a response which is experienced as either improving one's situation or making it worse.

Moreover, although he proposed to program Heidegger's account of everyday routine activities, Agre doesn't even try to account for how our experience feeds back and changes our sense of the significance of the next situation and what is relevant in *it*. In putting his virtual agent in a virtual micro-world where all possibly relevance is determined beforehand, Agre didn't try to account for how we learn to respond to new relevancies, and so, like Brooks, he finesses rather than solves the frame problem.

Merleau-Ponty's work, on the contrary, offers a nonrepresentational account of the way the body and the world are coupled that suggests a way of avoiding the frame problem. 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

---

38 Heidegger, *Logic: The Question of Truth*, Trans. Thomas Sheehan manuscript. *Gesamtausgabe*, Band 21, 144. 2008.

39 Heidegger goes on immediately to contrast the total absorption of coping he has just described with the as-structure of thematic observation:

> Every act of having *things* in front of oneself and *perceiving them* is held within [the] disclosure of those *things*, a disclosure that things get from a *primary meaningfulness*s in terms of the what-for. Every act of *having something in front of oneself and perceiving it* is, in and for itself, a 'having' s*omething as somethin*g.

To put it in terms of *Being and Time*, the as-structure of equipment goes all the way down in *the world,* but not in the way the world shows up in our absorbed coping. It is poor phenomenology to read the self and the as-structure into our experience when we are coping at our best.

40 There is a third possible attitude. Heidegger calls it responding to signs. Then I am sensitive to possibly relevant aspects of my environment and take them into account as I cope. We normally do this when driving in traffic, and the master potter, for example, is alert to the way the pot she is making may be deviating from the normal.

the perceptual world the *intentional arc*. He says: “Cognitive life, the life of desire or perceptual life—is subtended by an ‘intentional arc’ which projects round about us our past, our future, [and] our human setting.”[41]

5. Pseudo Heideggerian AI: Embedded, embodied, extended mind

As if taking up from where Agre left off with his objectified version of the ready-to-hand, in *Reconstructing the Cognitive World* Wheeler tells us:

> [O]ur global project requires a defense of action-oriented representation. ... [A]ction-oriented representation may be interpreted as the subagential reflection of online practical problem solving, as conceived by the Heideggerian phenomenologist. Embodied-embedded cognitive science is implicitly a Heideggerian venture.[42]

He further notes:

> As part of its promise, this nascent, Heideggerian paradigm would need to indicate that it might plausibly be able either to solve or to dissolve the frame problem.[43]

And he suggests:

> The good news for the reoriented Heideggerian is that the kind of evidence called for here may already exist, in the work of recent *embodied-embedded cognitive science*.[44]

He concludes:

> Dreyfus is right that the philosophical impasse between a Cartesian and a Heideggerian metaphysics can be resolved empirically via cognitive science. However, he looks for resolution in the wrong place. For it is not any alleged empirical failure on the part of orthodox cognitive science, but rather the concrete empirical success of a cognitive science with Heideggerian credentials, that, if sustained and deepened, would ultimately vindicate a Heideggerian position in cognitive theory.[45]

I agree that it is time for a positive account of Heideggerian AI and of an underlying Heideggerian neuroscience, but I think Wheeler is the one looking in the wrong place. Merely by supposing that Heidegger is concerned with *problem solving* and action oriented *representations,* Wheeler’s project reflects not a step beyond Agre but a regression to aspects of pre-Brooks GOFAI. Heidegger, indeed, claims that skillful coping is basic, but he is also clear that, all coping takes place on the background coping he calls being-in-the-world that doesn’t involve any form of representation at all.[46]

Wheeler’s cognitivist misreading of Heidegger leads him to overestimate the importance of Andy Clark’s and David Chalmers’ attempt to free us from the Cartesian idea that the mind is essentially inner by pointing out that in thinking we sometimes make use of external artifacts like pencil, paper, and computers.[47] Unfortunately, this argument for the extended mind preserves the Cartesian assumption that our basic way of relating to the world is by using propositional representations such as beliefs and memories whether they are in the mind or in notebooks in the world. In effect,

[41] Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. C. Smith (Routledge & Kegan Paul, 1962), 136.
[42] Wheeler, *Reconstructing the Cognitive World*, 222–223.
[43] Ibid. 187.
[44] Ibid. 188.
[45] Ibid. 188–189.
[46] Merleau-Ponty says the same:

> [T]o move one’s body is to aim at things through it; it is to allow oneself to respond to their call, which is made upon it independently of any representation. (*Phenomenology of Perception*, 139.)

[47] See, Clark, A. and Chalmers, D., “The extended mind,” *Analysis* 58 (1): 7–19, 199.

while Brooks happily dispenses with representations where coping is concerned, all Chalmers, Clark, and Wheeler give us as a supposedly radical new Heideggerian approach to the human way of being in the world is to note that memories and beliefs are not necessarily *inner* entities and that, therefore, *thinking* bridges the distinction between *inner and outer representations*.

Heidegger's important insight is not that, when we solve problems, we sometimes make use of representational equipment outside our bodies, but that *being-in-the-world* is more basic than *thinking* and solving problems;that it is not representational at all. That is, when we are coping at our best, we are drawn in by solicitations and respond directly to them, so that the distinction between us and our equipment—between inner and outer—vanishes.[48] As Heidegger sums it up:

> I *live* in the understanding of writing, illuminating, going-in-and-out, and the like. More precisely: as Dasein I am—in speaking, going, and understanding—an act of understanding dealing-with. My being in the world *is* nothing other than this already-operating-with-understanding in this mode of being.[49]

Heidegger and Merleau-Ponty's understanding of embedded embodied coping, then, is not that the *mind* is sometimes *extended into the world* but rather that all such problem solving is derivative, that in our most basic way of being, that is, as absorbed skillful copers, we are not minds at all but *one with the world*. Heidegger sticks to the phenomenon, when he makes the strange-sounding claim that, in its most basic way of being, "Dasein is its world existingly."[50]

When you stop thinking that mind is what characterizes us most basically but, rather, that most basically we are absorbed copers, the inner/outer distinction becomes problematic. There's no easily askable question as to whether the absorbed coping is in me or in the world. According to Heidegger, intentional content isn't in the mind, nor in some third realm (as it is for Husserl), nor in the world; it isn't anywhere. It's an embodied way of being-towards. Thus for a Heideggerian, all forms of *cognitivist* externalism presuppose a more basic *existential* externalism where even to speak of "externalism" is misleading since such talk presupposes a contrast with the internal. Compared to this genuinely Heideggerian view, extended-mind externalism is contrived, trivial, and irrelevant.

6. What motivates embedded/embodied coping?

But why is Dasein called to cope at all? According to Heidegger, we are constantly solicited to improve our familiarity with the world. Five years before the publication of *Being and Time* he wrote:

> Caring takes the form of a looking around and seeing, and as this circumspective caring it is at the same time ... concerned about developing its circumspection, that is, about *securing and expanding its familiarity* with the objects of its dealings.[51]

---

[48] As Heidegger puts it: "The self must forget itself if, lost in the world of equipment, it is to be able 'actually' to go to work and manipulate something." *Being and Time*, 405.

[49] Heidegger, *Logic*, 146. It's important to realize that when he uses the term "understanding," Heidegger explains (with a little help from the translator) that he means a kind of know-how:

> In German we say that someone can *vorstehen* something—literally, stand in front of or ahead of it, that is, stand at its head, administer, manage, preside over it. This is equivalent to saying that he *versteht sich darauf*, understands in the sense of being skilled or expert at it, has the know-how of it. (Martin Heidegger, *The Basic Problems of Phenomenology*, A. Hofstadter, Trans. Bloomington: Indian University Press, 1982, 276.)

[50] Heidegger, *Being and Time*, 416. To make sense of this slogan, it's important to be clear that Heidegger distinguishes the human *world* from the physical *universe*.

[51] Heidegger, *Phenomenological Interpretations in Connection with Aristotle,* in *Supplements: From the Earliest Essays to Being and Time and Beyond*, John Van Buren, Ed. (State University of New York Press, 2002), 115. My italics.

This away of putting the source of *significance* covers both animals and people. By the time he published *Being and Time*, however, Heidegger was interested exclusively in the special kind of significance found in the world opened up by human beings who are defined by the stand they take on their own being. We might call this *meaning*. In this paper I'm putting the question of uniquely human meaning aside to concentrate on the sort of significance we share with animals.

This pragmatic perspective is developed by Merleau-Ponty, and by Samuel Todes.[52] These heirs to Heidegger's account of familiarity and coping describe how an organism, animal or human, interacts with what is objectively speaking the meaningless physical universe in such a way as to cope with an environment organized in terms of that organism's *need to find its way around*. All such coping beings are motivated to get a more and more refined and secure sense of the specific objects of their dealings. According to Merleau-Ponty:

> My body is geared into the world when my perception presents me with a spectacle as varied and as clearly articulated as possible...[53]

In short, in our skilled activity we are drawn to move so as to achieve a better and better grip on our situation. For this movement towards maximal grip to take place one doesn't need a mental representation of one's goal nor any problem solving, as would a GOFAI robot. Rather, acting is experienced as a steady flow of skillful activity in response to the situation. When one's situation deviates from some optimal body-environment gestalt, one's activity takes one closer to that optimum and thereby relieves the "tension" of the deviation. One does not need to know what the optimum is in order to move towards it. One's body is simply drawn to lower the tension.

That is, if things are going well and I am gaining an optimal grip on the world, I simple respond to the solicitation to move towards an even better grip and, if things are going badly, I experience a pull back towards the norm. If it seems that much of the time we don't experience any such pull, Merleau-Ponty would no doubt respond that the sensitivity to deviation is nonetheless guiding one's coping, just as an airport radio beacon doesn't give a warning signal unless the plane strays off course, and then, let us suppose, the plane gets a signal whose intensity corresponds to how far off course it is and the intensity of the signal diminishes as it approaches getting back on course. The silence that accompanies being on course doesn't mean the beacon isn't continually guiding the plane. Likewise, the absence of felt tension in perception doesn't mean we aren't being directed by a solicitation.

As Merleau-Ponty puts it: "Our body is not an object for an 'I think', it is a grouping of lived-through meanings that moves towards its equilibrium."[54] Equilibrium being Merleau-Ponty's name for the zero gradient of steady successful coping. Moreover, normally, we do not arrive at equilibrium and stop there but are immediately taken over by a new solicitation.

7. Modeling situated coping as a dynamical system

Describing the phenomenon of everyday coping as being "*geared into*" the world and *moving towards "equilibrium"* suggests a dynamic relation between the coper and the environment. Timothy van Gelder calls this dynamic relation between coper and environment *coupling*, explaining its importance as follows:

> The fundamental mode of interaction with the environment is not to represent it, or even to exchange inputs and outputs with it; rather, the relation is better understood via the technical notion of coupling. ...
>
> The post-Cartesian agent manages to cope with the world without necessarily representing it. A dynamical approach suggests how this might be possible by showing how the internal operation of a system interacting with an external world can be so subtle and complex as to *defy* description in representational terms—how, in other words, cognition can *transcend* representation.[55]

---

[52] See, Samuel Todes, *Body and World* (Cambridge, MA: The MIT Press, 2001). Todes goes beyond Merleau-Ponty in showing how our world-disclosing perceptual experience is structured by the structure of our bodies. Merleau-Ponty never tells us what our bodies are actually like and how their structure affects our experience. Todes points out that our body has a front/back and up/down orientation. It moves forward more easily than backward, and can successfully cope only with what is in front of it. He then describes how, in order to explore our surrounding world and orient ourselves in it, we have to balance ourselves within a vertical field that we do not produce, be effectively directed in a circumstantial field (facing one aspect of that field rather than another), and appropriately set to respond to the specific thing we are encountering within that field. For Todes, then, perceptual receptivity is an embodied, normative, skilled accomplishment, in response to our need to orient ourselves in the world. Clearly, this kind of holistic background coping is not done for a reason.

[53] Merleau-Ponty, *Phenomenology of Perception*, 250. (Trans. modified.)

[54] Ibid, 153.

[55] Timothy Van Gelder, Dynamics and cognition, *Mind Design II*, John Haugeland, Ed., A Bradford Book (Cambridge, MA: The MIT Press, 1997), 439, 448.

Van Gelder shares with Brooks the existentialist claim that thinking such as problem solving, is grounded in a more basic relation of body and world. As van Gelder puts it:

> Cognition can, in sophisticated cases, [such as breakdowns, problem solving, and abstract thought] involve representation and sequential processing; but such phenomena are best understood as emerging from a dynamical substrate, rather than as constituting the basic level of cognitive performance.[56]

This dynamical substrate is precisely the causal basis of the skillful coping first described by Heidegger and worked out in detail by Merleau-Ponty and Todes.

Van Gelder importantly contrasts the rich interactive temporality of real-time on-line coupling of coper and world with the austere step by step temporality of thought. Wheeler helpfully explains:

> [W]hilst the computational architectures proposed within computational cognitive science require that inner events happen in the right order, and (in theory) fast enough to get a job done, there are, in general, no constraints on how long each operation within the overall cognitive process takes, or on how long the gaps between the individual operations are. Moreover, the transition events that characterize those inner operations are not related in any systematic way to the real-time dynamics of either neural biochemical processes, non-neural bodily events, or environmental phenomena (dynamics which surely involve rates and rhythms).[57]

> Computation is thus paradigmatically austere:

> Turing machine computing is digital, deterministic, discrete, effective (in the technical sense that behavior is always the result of an algorithmically specified finite number of operations), and temporally austere (in that time is reduced to mere sequence).[58]

Ironically, Wheeler's highlighting the contrast between rich dynamic temporal coupling and austere computational temporality enables us to see clearly that his appeal to extended minds as a Heideggerian response to Cartesianism leaves out the essential temporal character of embodied embedding. Clark and Chalmers's examples of extended minds manipulating representations such as notes and pictures are clearly cases of temporal austerity—no rates and rhythms are involved.

Wheeler is aware of this possible objection to his backing both the *dynamical systems* model and the *extended mind* approach. He asks: "What about the apparent clash between continuous reciprocal causation and action orientated representations? On the face of it this clash is a worry for our emerging cognitive science."[59] But instead of engaging with the incompatibility of these two opposed models of *ground level intelligence,* Wheeler suggests that we must somehow combine them and that "this question is perhaps one of the biggest of the many challenges that lie ahead."[60]

Wheeler, however, hopes he can combine these approaches by appealing to the account of involved problem solving which Heidegger calls dealing with the unready-to-hand. Wheeler's point is that, unlike detached problem solving with its general representations, the unready-to-hand requires situation-specific representations. But, as we have seen, for Heidegger all un-ready-to-hand coping takes place on the background of an even more basic nonrepresentational holistic coping that allows copers to orient themselves in the world.

Heidegger describes this background as "the background of … primary familiarity, which itself is not conscious and intended but is rather present in [an] unprominent way."[61] In *Being and Time* he speaks of "that familiarity in accordance with which Dasein … 'knows its way about' [sich auskennt] in its public environment" (405). This coping is like the ready-to-hand in that it does not involve representations. So Heidegger says explicitly that our background

[56] Ibid.

[57] Wheeler, Change in the rules: Computers, dynamical systems, and Searle, in *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, John Preston and Mark Bishop, Eds. (Oxford: Clarendon Press, 2002), 345.

[58] Ibid. 344, 345.

[59] Wheeler, *Reconstructing the Cognitive World*, 280.

[60] Ibid.

[61] Heidegger, *History of the Concept of Time*, Trans. T. Kisiel (Bloomington, IN: Indiana University Press, 1985), 189.

being-in-the-world, which he also calls transcendence, does not involve representational intentionality, but, rather, makes intentionality possible:

> *Transcendence* is a *fundamental determination of the ontological structure of the Dasein*. ... Intentionality is founded in the Dasein's transcendence and is possible solely for this reason—transcendence cannot conversely be explained in terms of intentionality.[62]

To be more exact, 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. What would it be to succeed or fail in finding ones way around in the familiar world? 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.[63]

This is not a disagreement between Wheeler and me about the relative frequency of dealing with the ready-to-hand and the unready-to-hand in everyday experience. True, Wheeler emphasizes intermittent reflective activities such as learning and practical problem solving, whereas I, like Heidegger, emphasize pervasive activities like going out the door, walking on the floor, turning on and off the lights, etc. The question of the relative frequency of the ready-to-hand and the unready-to-hand modes of being is, Wheeler and I agree, an empirical question.[64]

But the issue concerning the background is not an empirical question. It is an ontological question. And, as we have just seen, Heidegger is clear that the mode of being of the world is not that of a collection of independent modules that define what is relevant in specific situations. It seems to me that Wheeler is on the right track, leaving modular solutions and action oriented representations behind, when he writes:

> [W]here one has CRC [continuous reciprocal causation] one will have a non-modular system. Modularity is necessary for homuncularity and thus, on my account, necessary for representation of any kind. To the extent that the systems underlying intelligence are characterized by CRC, they will be non-representational, and so the notion of action-oriented representation won't help explain them. (Personal communication.)

Wheeler directly confronts my objection when he adds:

> If one could generate the claim that CRC must be the norm at the subagential level from a Heideggerian analysis of the agential level, then the consequence for me would be that, to be Heideggerian, I would have to concede that action-oriented representation will in fact do less explanatory work than I have previously implied. (Personal correspondence continued.)

But Wheeler misses my point when he adds:

> However, this takes us back to the points I made above about the prevalence of unreadiness-to-hand. Action-oriented representations will underlie our engagements with the unready-to-hand. In this domain, I suggest, the effects of CRC will be restricted. And, I think, unreadiness-to-hand is the (factual) norm. (Personal correspondence continued.)

We just agreed, that this is not an *empirical* question concerning the *frequency* of coping with the unready-to-hand but an *ontological* point about the background of *all* modes of coping. If Wheeler wants to count himself a Heideggerian, he does, indeed, "have to concede that action-oriented representation will in fact do less explanatory work than [he] previously implied."

---

[62] Heidegger, *The Basic Problems of Phenomenology*, trans. A. Hofstadter (Bloomington, IN: Indiana University Press, 1982), 162.

[63] Moreover, the background solicitations are constantly enriched, not by adding new bits of information as Wheeler suggests, but by allowing finer and finer discriminations that show up in the world by way of the intentional arc.

[64] We agree too that both these modes of encountering the things in the world are more frequent and more basic than appeal to general-purpose reasoning and goal oriented planning.

Wheeler seems to be looking for a neurodynamic model of brain activity such as we will consider in a moment when he writes:

> [A]lthough there is abundant evidence that (what we are calling) continuous reciprocal causation can mediate the transition between different phases of behavior within the same task, that is not the same thing as switching between contexts, which typically involves a reevaluation of what the current task might be. Nevertheless, I am optimistic that essentially the same processes of fluid functional and structural reconfiguration, driven in a bottom-up way by low-level neurochemical dynamics, may be at the heart of the more complex capacity.[65]

Meanwhile, Wheeler's ambivalence concerning which model is more basic, the representational or the dynamic, undermines his Heideggerian approach. For, as Wheeler himself sees, the Heideggerian claim is that action-oriented coping, as long as it is involved (online, Wheeler would say) is not representational at all and does not involve any problem solving, and that all representational problem solving takes place offline and presupposes involved background coping. Showing in detail how the representational un-ready-to-hand in all its forms depends upon a background of holistic, nonrepresentational coping is exactly the Heideggerian project and would, indeed, be the most important contribution that Heideggerian AI could make to Cognitive Science. Indeed, a Heideggerian Cognitive Science would require working out an ontology, phenomenology, and brain model, that denies a basic role to any sort of representation—even action oriented ones—and defends a dynamical model like Merleau-Ponty's and van Gelder's that gives a primordial place to equilibrium and in general to rich coupling.

Ultimately, we will have to choose which sort of AI and which sort of neuroscience to back, and so we are led to the questions: could the brain in its causal support of our active coping instantiate a richly coupled dynamical system, and is there any evidence it actually does so? If so, could this coupling be modeled on a digital computer to give us Heideggerian AI or at least Merleau-Pontian AI? And would that solve the frame problem?

8. Walter Freeman's Merleau-Pontian neurodynamics

We have seen that our experience of the everyday world (not the universe) is given as already organized in terms of significance and relevance, and that significance can't be constructed by giving meaning to brute facts—both because we don't normally experience brute facts and, even if we did, no value predicate could do the job of giving them situational significance. Yet, all that the organism can receive is mere physical energy. How can such senseless physical stimulation be experienced directly as significant? All generally accepted neuro-models fail to help, even when they talk of dynamic coupling, since they still accept the basic Cartesian model, viz.:

1. The brain *receives input* from the universe by way of its sense organs (the picture on the retina, the vibrations in the cochlea, the odorant particles in the nasal passages, etc.).

2. Out of this stimulus information, the brain abstracts *features*, which it uses to *construct a representation* of the world.

This is supposedly accomplished either (a) by applying rules such as the frames and scripts of GOFAI—an approach that is generally acknowledged to have failed to solve the frame problem. Or (b) by strengthening or weakening weights on connections between simulated neurons in a simulated neural network depending on the success or failure of the net's output as defined by the net designer. Significance is thus added *from outside* since the net is not seeking anything. This approach does not even try to capture the animal's way of actively determining the significance of the stimulus on the basis of its past experience and its current arousal.

Both these approaches treat the computer or brain as a passive receiver of bits of meaningless data, which then have to have significance added to them. The big problem for the traditional neuro-science approach is, then, to understand how the brain binds the relevant features to each other. That is, the problem for normal neuro-science is how to pick out and relate features relevant to each other from among all the independent, isolated features picked up by each of the independent isolated receptors. For example, is the redness that has just been detected relevant to the square or the circle shape also detected in the current input? This problem is the neural version of the frame problem in AI: How can the brain keep track of which facts in its representation of the current world are relevant to which other facts?

[65] Wheeler, *Reconstructing the Cognitive World*, 279.

Like the frame problem, as long as the mind/brain is thought of as passively receiving meaningless inputs that need to have significance and relevance added to them, the binding problem has remained unsolved and is almost certainly unsolvable. Somehow the phenomenologist's description of how the active organism has direct access to significance must be built into the neuroscientific model.

Wheeler has argued persuasively for the importance of a positive alternative in overthrowing established research paradigms. Without such a positive account the phenomenological observation that the world is its own best representation, and that the significance we find in our world is constantly enriched by our experience in it, seems to require that the brain be what Dennett derisively calls "wonder tissue."

Fortunately, there is at least one model of how the brain could provide the causal basis for the intentional arc and so avoid the binding problem. Walter Freeman, a founding figure in neurodynamics and one of the first to take seriously the idea of the brain as a nonlinear dynamical system,[66] has worked out an account of how the brain of an active animal can directly pick up and augment significance in its world. On the basis of years of work on olfaction, vision, touch, and hearing in alert and moving rabbits, Freeman has developed a model of rabbit learning based on the coupling of the rabbit's brain and the environment. He maintains:

> [T]he brain moves beyond the mere extraction of features ... it combines sensory messages with past experience ... to identify both the stimulus and its particular meaning to the individual.[67]

To bring out the structural analogy of Freeman's account to Merleau-Ponty's phenomenological descriptions, I propose to map Freeman's neurodynamic model onto the phenomena Merleau-Ponty has described. Freeman's neurodynamics implies the involvement of the whole brain in perception and action, but for explaining the core of his ideas I'll focus on the dynamics of the olfactory bulb, since his key research was done on that part of the rabbit brain.

8.1. Direct perception of significance and the rejection of the binding problem

While all other researchers assume the passive reception of input from the universe, Freeman, like Merleau-Ponty on the phenomenological level, and Gibson on the (ecological) psychology level, develops a third position between the intellectualist and the empiricist. Merleau-Ponty, Gibson, and Freeman take as basic that the brain is embodied in an animal moving in the environment to satisfy its needs.

Freeman maintains that information about the world is not gained by detecting meaningless features and processing these features step-by-step upwards toward a unified representation. The binding problem only arises as an artifact of trying to interpret the output of isolated cells in the receptors of immobilized organisms. Rather, Freeman turns the problem around and asks: Given that the environment is already significant for the animal, how can the animal select a unified significant figure from the noisy background? This turns the binding problem into a selection problem. As we shall see, however, this selection is not among patterns existing in the world but among patterns in the animal that have been formed by its prior interaction with the world.

In Freeman's neurodynamic model, the animal's perceptual system is primed by past experience and arousal to seek and be rewarded by relevant experiences. In the case of the rabbit, these could be carrot smells found in the

---

[66] Wheeler explains:

> [F]or the purposes of a dynamical systems approach to Cognitive Science, a dynamical system may be defined as any system in which there is *state-dependent change*, where systemic change is state dependent just in case the future behavior of the system depends causally on the current state of the system. (*Reconstructing the Cognitive World,* 91.)
>
> [N]onlinear dynamical systems exhibit a property known as *sensitive dependence on initial conditions*, according to which the trajectories that flow from two adjacent initial-condition-points diverge rapidly. This means that a small change in the initial state of the system becomes, after a relatively short time, a large difference in the evolving state of the system. This is one of the distinguishing marks of the phenomenon of chaos. ...
>
> [Consider] the case of two theoretically separable dynamical systems that are bound together, in a mathematically describable way, such that some of the parameters of each system either are, or are functions of, some of the state variables of the other. At any particular time, the state of each of these systems will, in a sense, fix the dynamics of the other system. Such systems will evolve through time in a relation of complex and intimate mutual influence, and are said to be *coupled*. (*Reconstructing the Cognitive World*, 93.)

[67] Walter J. Freeman, The physiology of perception, *Scientific American*, Vol. 242, February 1991, 78.

course of seeking and eating a carrot. When the animal succeeds, the connections between those cells in the rabbit's olfactory bulb that were involved are strengthened according to "the widely accepted Hebbian rule, which holds that synapses between neurons that fire together become stronger, as long as the synchronous firing is accompanied by a reward."[68] The neurons that fire together wire together to form what Hebb called *cell assemblies*. The cell assemblies that are formed by the rabbit's response to what is significant for it are in effect tuned to select the significant sensory input from the background noise. For example, those cells involved in a previous narrow escape from a fox would be wired together in a cell assembly. Then, in an environment previously experienced as dangerous, those cell assemblies sensitive to the smell of foxes would be primed to respond.

Freeman notes that: "For a burst [of neuronal activity] to occur in response to some odorant, the neurons of the assembly and the bulb as whole must first be "primed" to respond strongly to that specific input."[69] And he adds: "Our experiments show that the gain [sensitivity to input] in neuronal collections increases in the bulb and olfactory cortex when the animal is hungry, thirsty, sexually aroused or threatened."[70] So, if a male animal has just eaten and is ready to mate, the gain is turned down on the cell assemblies responsive to food smells, and turned up on female smells. Thus, from the start the cells assemblies are not just passive receivers of meaningless input from the universe but, on the basis of past experience, are tuned to respond to what is significant to the animal given its arousal.

Once we see that the cell assemblies in involved, coping animals respond directly to significant aspects of the environment, we can also see why the binding problem need not arise. The problem is an artifact of trying to interpret the output of isolated cells in the cortex of animals from the perspective of the researcher rather than the perspective of the animal. That is, the researcher, like Merleau-Ponty's intellectualist, interprets the firing of the cells in the sense organ as responding to features of an object-type—features such as orange, round, and tapered that can be specified independently of the object to which they belong. The researcher then has the problem of how the brain binds these isolated features into a representation of, say, a carrot (and adds the function predicate, good to eat). But, according to Freeman, in an active, hungry animal the output from the isolated detector cells triggers a cell assembly already tuned to detect the relevant input on the basis of past significant experience, which, in turn puts the brain into a state that signals to the limbic system eat this now, without the brain ever having to solve the problem of how the isolated features abstracted by the researchers are brought together into the presentation of an object.

Freeman, dramatically describes the brain activity involved:

> If the odorant is familiar and the bulb has been primed by arousal, the information spreads like a flash fire through the nerve cell assembly. First, excitatory input to one part of the assembly during a sniff excites the other parts, via the Hebbian synapses. Then those parts reexcite the first, increasing the gain, and so forth, so that the input rapidly ignites an explosion of collective activity throughout the assembly. The activity of the assembly, in turn, guides the entire bulb into a new state by igniting a full-blown burst.[71]

Specifically, after each sniff, the rabbit's olfactory bulb goes into one of several possible states that neural modelers traditionally call energy states. A state tends toward minimum "energy" the way a ball tends to roll towards the bottom of a container, no matter where it starts from within the container. Each possible minimal energy state is called an *attractor*. The brain states that tend towards a particular attractor no matter where they start in the basin are called that attractor's *basin of attraction*. As the brain activation to pulled into an attractor, the brain in effect selects the meaningful stimulus from the background.

Thus the stimuli need not be processed into a representation of the current situation on the basis of which the brain then has to infer what is present in the environment. Rather on Freeman's account, the rabbit's brain forms a new basin of attraction for each new significant class of inputs. The significance of past experience is preserved in basins of attraction. The set of basins of attraction that an animal has learned form what is called an *attractor landscape*. According to Freeman:

[68] Ibid. 81.
[69] Ibid. 82
[70] Ibid.
[71] Ibid. 83.

> The state space of the cortex can therefore be said to comprise an attractor landscape with several adjoining basins of attraction, one for each class of learned stimuli.[72]

Thus Freeman contends that each new attractor does not *represent,* say, a carrot, or the smell of carrot, or even what to do with a carrot. Rather, the brain's current state is the result of the sum of the animal's past experiences with carrots. What in the physical input is directly picked up and resonated to when the rabbit sniffs, then, is the affords-eating,[73] and the brain state is directly coupled with (or in Gibson's terms resonates to) the affordance offered by the current carrot.

Freeman offers a helpful analogy:

> We conceive each cortical dynamical system as having a state space through which the system travels as a point moving along a path (trajectory) through the state space. A simple analogy is a spaceship flying over a landscape with valley resembling the craters on the moon. An expected stimulus contained in the omnipresent background input selects a crater into which the ship descends. We call the lowest area in each crater an 'attractor' to which the system trajectory goes, and the set of craters basins of attraction in an attractor landscape. There is a different attractor for each class of stimuli that the system [is primed] to expect.[74]

Freeman concludes: "The macroscopic bulbar patterns [do] not relate to the stimulus directly but instead to the *significance* of the stimulus."[75] Indeed, after triggering a specific attractor and modifying it, the stimulus—the impression made on the receptor cells in the sense organ—has no further job to perform. Freeman explains:

> The new pattern is selected by the stimulus from the internal pre-existing repertoire [of attractors], not imposed by the stimulus. It is determined by prior experience with this class of stimulus. The pattern expresses the nature of the class and its significance for the subject rather than the particular event. The identities of the particular neurons in the receptor class that are activated are irrelevant and are not retained[76] ... Having played its role in setting the initial conditions, the sense-dependent activity is washed away.[77]

Thus, as Merleau-Ponty claims and psychological experiments confirm, we normally have no experience of the data picked up by the sense organs.[78]

8.2. Learning and Merleau-Ponty's intentional arc

Thus, according to Freeman's model, when hungry, frightened, etc., the rabbit sniffs around seeking food, runs toward a hiding place, or does whatever else prior experience has taught it is successful. The weights on the animal's neural connections are then changed on the basis of the quality of its resulting experience. That is, they are changed in a way that reflects the extent to which the result satisfied the animal's current need.

Freeman claims his read-out from the rabbit's brain shows that each learning experience with a previously unknown stimulus, or an unimportant stimulus class that is significant in a new way, sets up a new attractor for that class and *rearranges all the other attractor basins in the landscape*:

---

[72] Freeman, *How Brains Make Up Their Minds*, New York: Columbia University Press, 2000, 62. (Quotations from Freeman's books have been reviewed by him and sometimes modified to correspond to his latest vocabulary and way of thinking about the phenomenon.)

[73] Thus Freeman's model might well describe the brain activity presupposed by Gibson's talk of "resonating" to affordances.

[74] Freeman, Nonlinear dynamics of intentionality. *Journal of Mind and Behavior* 18: 291–304, 1997. The attractors are abstractions relative to what level of abstraction is significant given what the animal is seeking.

[75] Freeman, *Societies of Brains: A study in the neuroscience of love and hate*, *The Spinoza Lectures, Amsterdam, Netherlands,* Hillsdale, N.J.: Lawrence Erlbaum Associates, Publisher, 1995, 59. (My italics.)

[76] Freeman, *Societies of Brains*, 66. (My italics.)

[77] Ibid. 67.

[78] Sean Kelly, Content and constancy: Phenomenology, psychology, and the content of perception, forthcoming in *Philosophy and Phenomenological Research*.

> 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. There are no fixed representations, as there are in [GOFAI] computers; there are only significances.[79]

The constantly updated landscape of attractors is presumably correlated with the agent's experience of the changing significance of things in the world, that is, with the intentional arc.

Freeman adds:

> 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. This property contrasts with memory stores in computers ... in which each item is positioned by an address or a branch of a search tree. There, each item has a compartment, and new items don't change the old ones. Our data indicate that in brains the store has no boundaries or compartments.... Each new state transition ... initiates the construction of a local pattern that impinges on and modifies the whole intentional structure.[80]

Merleau-Ponty likewise concludes that, thanks to the intentional arc, no two experiences of the world are ever exactly alike.[81]

It is important to realize how different this model is from any representationalist account. There is no fixed and independent intentional structure in the brain—not even a latent one. There is nothing that can be found in the olfactory bulb in isolation that represents or even corresponds to anything in the world. There is only the fact that, given the way the nerve cell assemblies have been wired on the basis of past experience, when the animal is in a state of arousal and is in the presence of a significant item such as food or a potential predator or a mate, the bulb will go into a certain attractor state. That activity state in the current interaction of animal and environment corresponds to the whole world of the organism with some aspect salient. The activity is not an isolate brain state but only comes into existence and only is maintained as long as, and in so far as, it is dynamically coupled with the significant situation in the world that selected it, and does not exist apart from it. Whereas, as we have seen, in the cognitivist notion of representations, a representation exists apart from what it represents.

Thus Freeman offers a model of learning which is not an associationist model according to which, as one learns, one adds more and more fixed connection, nor a cognitivist model based on off-line representations of objective facts about the world that enable off line inferences as to which facts to expect next, and what they mean. Rather, Freeman's model instantiates the causal basis of a genuine intentional arc in which there are no linear casual connections between world and brain nor a fixed library of representations, but where, each time a new significance is encountered, the whole perceptual world of the animal changes so that the significance that is directly displayed in the world of the animal is continually enriched.

8.3. The perception/action loop

The brain's movement towards the bottom of a particular basin of attraction underlies the perceiver's perception of the significance for action of a particular experience.[82] For example, if a carrot affords eating the rabbit is directly readied to eat the carrot, or perhaps readied to carry off the carrot depending on which attractor is currently activated. Freeman tells us:

---

[79] Freeman, *How Brains Make Up Their Minds*, 22.

[80] Freeman, *Societies of Brains*, 99. (My italics.)

[81] Merleau-Ponty, *Phenomenology of Perception*, 216.

[82] See Sean Kelly, The Logic of Motor Intentionality, Unpublished draft. Also, Corbin Collins describes the phenomenology of this motor intentionality and spells out the logical form of what he calls instrumental predicates. See, "Body Intentionality," Inquiry, Dec. 1988.

> The same global states that embody the significance provide … the patterns that make choices between available options and that guide the motor systems into sequential movements of intentional behavior.[83]

The animal must take account of how things are going and either continue on a promising path, or, if the overall action is not going as well as anticipated, the brain must self-organize so the attractor system jumps to another attractor. This either causes the animal to act in such a way as to increase its sense of impending reward, or the brain will shift attractors again, until it lands in one that makes such an improvement. The attractors can change like switching from frame to frame in a movie film with each further sniff or with each shift of attention. If the rabbit achieves what it is seeking, a report of its success is fed back to reset the sensitivity of the olfactory bulb. And the cycle is repeated.

Freeman's overall picture of skilled perception and action, then, is as follows. The animal, let's say a rabbit sniffing a carrot, receives stimuli that, thanks to prior Hebbian learning, puts its olfactory bulb into a specific attractor basin. For example, the attractor that has been formed by, and amounts to, the brain's classification of the stimulus as affording eating. Along with other brain systems, the bulb *selects* a response. The rabbit is solicited to eat this now. It would be too cognitivist to say the bulb *sends a message*, to the appropriate part of the brain and too mechanistic to say the bulb *causes* the activity of eating the carrot. The meaning of the input is neither in the stimulus nor in a mechanical response directly triggered by the stimulus. Significance is not stored as a memory-representation nor an association. Rather the memory of significance is in the repertoire of attractors as classifications of possible responses—the attractors themselves being the product of past experience.

Once the stimulus has been classified by selecting an attractor that says eat this now, the problem for the brain is just how this eating is to be done. On-line coping needs a stimuli-driven feedback policy dictating how to move rapidly over the terrain and approach and eat the carrot. Here, an actor-critic version of Temporal Difference Reinforcement Learning (TDRL) can serve to augment the Freeman model.

According to TDRL, learning the appropriate movements in the current situation requires learning the expected final award as well as the movements. These two functions are learned slowly through repeated experiences. Then the brain can monitor directly whether the expectation of reward is being met as the rabbit approaches the carrot to eat it. If the expected final reward suddenly decreases due, for example, to the current inaccessibility of the carrot, the relevant part of the brain prompts the olfactory bulb to switch to a new attractor or perspective on the situation that dictates a different learned action, say dragging the carrot with its expected reward.[84] Only after a skill is thus acquired can the current stimuli, plus the past history of responding to related stimuli now wired into cell assemblies, produce the rapid responses required for on-going skillful coping.

8.4. Optimal grip

The animal's movements are presumably experienced by the animal as tending towards getting and maintaining an optimal perceptual take on what is currently significant, and, where appropriate, an ongoing optimal bodily grip on it. As Merleau-Ponty says: "through [my] body I am at grips with the world."[85] Freeman sees his account of the brain dynamics underlying perception and action as structurally isomorphic with Merleau-Ponty's. He explains:

> Merleau-Ponty concludes that we are moved to action by disequilibrium between the self and the world. In dynamic terms, the disequilibrium … puts the brain onto … a pathway through a chain of preferred states, which are learned basins of attraction. The penultimate result is not an equilibrium in the chemical sense, which is a dead state, but a descent for a time into the basin of an attractor … .[86]

Thus, according to Freeman, in governing action the brain normally moves from one basin of attraction to another descending into each basin for a time without coming permanently to rest in any one basin. The body is thereby led to

83 Freeman, *How Brains Make Up Their Minds*, 114.

84 Stuart E. Dreyfus, "Totally Model-Free Learned Skillful Coping." Bulletin of Science, Technology & Society, Vol. 24, No. 3, June 2004, 182–187. This article, however, does not discuss the role of a controlling attractor or the use of expected reward to jump to a new attractor.

85 Merleau-Ponty, *Phenomenology of Perception*, 303.

86 Freeman, *How Brains Make Up Their Minds*, 121.

move *towards* a maximal grip but, instead of remaining at rest when a maximal grip is achieved, the coupled coper is drawn to move on in response to another affordance that solicits the body to take up the same task from another angle, or to turn to the next task that grows out of the current one.

The selected attractor, together with input from the sense organs, then signals the limbic system to implement a new action with its new expected reward. Then again a signal comes back to the olfactory bulb and elsewhere as to whether the activity is progressing as expected. If so, the current attractor and action will be maintained but, if the result is not as expected, with the formation of the next attractor landscape some other attractor will be selected on the basis of past learning. In Merleau-Ponty's terms, Freeman's model, as we have seen, explains the intentional arc—how our previous coping experiences feed back to determine what action the current situation solicits—while the TDRL model keeps the animal moving toward a sense of minimal tension, that is, a least rate of change in expected reward, and hence towards achieving and maintaining what Merleau-Ponty calls a maximal grip.

8.5. Circular causality

Such systems are self-organizing. Freeman explains:

> Macroscopic ensembles exist in many materials, at many scales in space and time, ranging from ... weather systems such as hurricanes and tornadoes, even to galaxies. In each case, the behavior of the microscopic elements or particles is constrained by the embedding ensemble, and microscopic behavior cannot be understood except with reference to the macroscopic patterns of activity ....[87]

Thus, the cortical field controls the neurons that create the field. In Freeman's terms, in this sort of circular causality the overall activity "enslaves" the elements. As he emphasizes:

> Having attained through dendritic and axonal growth a certain density of anatomical connections, the neurons cease to act individually and start participating as part of a group, to which each contributes and from which each accepts direction .... The activity level is now determined by the population, not by the individuals. This is the first building block of neurodynamics.[88]

Given the way the whole brain can be tuned by past experience to influence individual neuron activity, Freeman can claim:

> Measurements of the electrical activity of brains show that dynamical states of Neuroactivity emerge like vortices in a weather system, triggered by physical energies impinging onto sensory receptors. ...[89]

Merleau-Ponty seems to anticipate Freeman's neurodynamics when he says:

> It is necessary only to accept the fact that the physico-chemical actions of which the organism is in a certain manner composed, instead of unfolding in parallel and independent sequences, are constituted ... in relatively stable "vortices."[90]

9. Freeman's model as a basis for Heideggerian AI

According to Freeman, the discreteness of global state transitions from one attractor basin to another makes it possible to model the brain's activity on a computer. The model uses numbers to stand for these discrete state transitions. He notes that:

[87] Ibid. 52.
[88] Ibid. 53.
[89] Freeman, *Societies of Brains*, 111.
[90] Merleau-Ponty, *The Structure of Behavior*, 153.

> At macroscopic levels each perceptual pattern of Neuroactivity is discrete, because it is marked by state transitions when it is formed and ended. ... I conclude that brains don't use numbers as symbols, but they do use discrete events in time and space, so we can represent them ... by numbers in order to model brain states with digital computers.[91]

That is, the states of the model are representations of brain states, not of the features of things in the everyday world. Just as simulated neural nets simulate brain processing but do not contain symbols that represent features of the world, the computer can model the series of discrete state transitions from basin to basin, thereby modeling how, on the basis of past experiences of success or failure, physical inputs are directly perceivable as significant for the organism. But the model is not an intentional being, only a description of such.

Freeman has actually programmed his model of the brain as a dynamic physical system, and so claims to have shown what the brain is doing to provide the material substrate for Heidegger's and Merleau-Ponty's phenomenological account of everyday perception and action. This may well be the new paradigm for the Cognitive Sciences that Wheeler proposes to present in his book but which he fails to find. It would show how the emerging embodied-embedded approach could be step towards a genuinely existential AI. Although, as we shall see, it would still be a very long way from programming human intelligence. Meanwhile, the job of phenomenologists is to get clear concerning the phenomena that must to be explained. That would include an account of how human beings, unlike the so-called Heideggerian computer models we have discussed, don't just ignore the frame problem nor solve it, but show why it doesn't occur.

10. How Heideggerian AI would dissolve rather than avoid or solve the frame problem

As we have seen, Wheeler rightly thinks that the simplest test of the viability of any proposed AI program is whether it can solve the frame problem. We've also seen that the two current supposedly Heideggerian approaches to AI avoid rather than solve the frame problem. Brooks's empiricist/behaviorist approach in which the environment directly causes responses avoids it by leaving out significance and learning altogether, while Agre's action-oriented approach, which includes only a small fixed set of possibly relevant responses, fails to face the problem of *changing* relevance.

Wheeler's own proposal, however, by introducing flexible action-oriented *representations*, like any representational approach, has to face the frame problem head on. To see why, we need only slightly revise his statement of the frame problem (quoted earlier), substituting "representation" for "belief":

> [G]iven a dynamically changing world, how is a nonmagical system ... to retrieve and (if necessary) to revise, out of all the *representations* that it possesses, just those *representations* that are relevant in some particular context of action?[92]

Wheeler's frame problem, then, is to explain how his allegedly Heideggerian system can determine in some systematic way which of the action-oriented representations it contains or can generate are relevant in a current situation, and keep track of how this relevance changes with changes in the situation.

Given his emphasis on problem solving and representations, it is not surprising that the concluding chapter of Wheeler's book, where he returns to the frame problem to test his proposed Heideggerian AI, offers no solution or dissolution of the problem. Instead, he asks us to "give some credence to [his] informed intuitions,"[93] which I take to be on the scent of Freeman's account of rabbit olfaction, that nonrepresentational causal coupling must play a crucial role. But I take issue with his conclusion that:

> *in extreme cases* the neural contribution will be *nonrepresentational* in character. In other cases, *representations* will be active partners alongside certain additional factors, but those representations will be action oriented in

[91] Freeman, *Societies of Brains* 105.
[92] Wheeler, *Reconstructing the Cognitive World,* 179.
[93] Ibid. 279.

character, and so will realize the same content-sparse, action-specific, egocentric, context-dependent profile that Heideggerian phenomenology reveals to be distinctive of online *representational* states at the agential level.[94]

But for Heidegger, *all* representational accounts are part of the problem. Wheeler's account, so far as I understand it, gives no explanation of how online dynamic coupling is supposed to dissolve the online frame problem. Nor does it help to wheel in, as Wheeler does, action-oriented representations and the extended mind. Any attempt to solve the frame problem by giving any role to any sort of representational states, even online ones, has so far proved to be a dead end. It looks like nonrepresentational neural activity can't be understood to be the "extreme case." Rather, such activity must be, as Heidegger, Merleau-Ponty and Freeman contend, our basic way of responding directly to relevance in the everyday world, so that the frame problem does not arise.

Heidegger and Merleau-Ponty argue that, and Freeman demonstrates how, thanks to our embodied coping and the intentional arc it makes possible, we directly respond to relevance and our skill in sensing and responding to relevant changes in the world is constantly improved. In coping in a particular context, say a classroom, we learn to ignore most of what is in the room, but, if it gets too warm, the windows solicit us to open them. We ignore the chalk dust in the corners and the chalk marks on the desks but we attend to the chalk marks on the blackboard. We take for granted that what we write on the board doesn't affect the windows, even if we write, "open windows," and what we do with the windows doesn't affect what's on the board. And as we constantly refine this background know-how, the things in the room and its layout become more and more familiar, take on more and more significance, and each thing draws us to act when an action is relevant. Thus we become better able to cope with change. Given our experience in the world, whenever there is a change in the current context we respond to it only if in the past it has turned out to be significant, and even when we sense a significant change we treat everything else as unchanged except what our familiarity with the world suggests might also have changed and so needs to be checked out. Thus, for embedded–embodied beings a local version of the frame problem does not arise.

But the frame problem reasserts itself when we consider changing contexts. How do we sense when a situation on the horizon has become relevant to our current task? When Merleau-Ponty describes the phenomenon, he speaks of one's attention being drawn by an affordance on the margin of one's current experience:

> To see an object is either to have it on the fringe of the visual field and be able to concentrate on it, or else respond to this *summons* by actually concentrating on it.[95]

Thus, for example, as one faces the front of a house, one's body is already being *summoned* (not just *prepared*) to go around the house to get a better look at its back.[96]

Merleau-Ponty's treatment of what Husserl calls the *inner* horizon of the perceptual object, e.g. its insides and back, applies equally to our experience of a situation's *outer* horizon of other potential situations. As I cope with a specific task in a specific situation, other situations that have in the past been relevant are right now present on the horizon of my experience as potentially (not merely possibly) relevant to my current situation.

If Freeman is right, our sense of familiar-but-not-currently-fully-present aspects of what is currently ready-to-hand, as well as our sense of other potentially relevant familiar situations on the horizon of the current situation, might well be correlated with the fact that brain activity is not simply in one attractor basin at a time but is influenced by other attractor basins in the same landscape, as well as by other attractor landscapes which under what have previously been experienced as relevant conditions are ready to draw current brain activity into themselves. According to Freeman, what makes us open to the horizontal influence of other attractors is that the whole system of attractor landscapes collapses and is rebuilt with each new rabbit sniff, or in our case, presumably with each shift in our attention. And after each collapse, a new landscape may be formed on the basis of new significant stimuli—a landscape in which, thanks to past experiences, a different attractor is active.[97] This presumably underlies our experience of being summoned.

---

94 Ibid. 276. (My italics.)

95 Merleau-Ponty, *Phenomenology of Perception*, 67. (My italics.)

96 Kelly, Seeing things in Merleau-Ponty, in *The Cambridge Companion to Merleau-Ponty*.

97 We do not experience these rapid changes of attractor landscapes anymore than we experience the flicker in changes of movie frames. Not everything going on in the brain is reflected in the phenomena.

And, once one correlates Freeman's neurodynamic account with Merleau-Ponty's description of the way the intentional arc feeds back our past experience into the way the world appears to us so that the world solicits from us ever-more-appropriate responses to its significance, we can see that we can be directly summoned to respond appropriately not only to what is relevant in our current situation, but we may be summoned by other familiar situations on the horizon of the present one. Then the fact that we can deal with changing relevance by anticipating what will change and what will stay the same no longer seems unsolvable.

But there is a generalization of the problem of relevance, and thus of the frame problem, that still seems intractable. In *What Computers Can't Do* I gave an example of the possible relevance of everything to everything. In placing a racing bet we can usually restrict ourselves to such relevant facts as the horse's age, jockey, and past performance but there are always other factors such as whether the horse is allergic to goldenrod or whether the jockey has just had a fight with the owner, which in some cases can be decisive. Human handicappers are capable of noticing such anomalies when they come across them.[98] But since anything in experience could be relevant to anything else, for representational/computation AI such an ability seems incomprehensible. Jerry Fodor follows up on my pessimistic example:

> "The problem," he tells us, "is to get the structure of an entire belief system to bear on individual occasions of belief fixation. We have, to put it bluntly, no computational formalisms that show us how to do this, and we have no idea how such formalisms might be developed. ... If someone—a Dreyfus, for example—were to ask us why we should even suppose that the digital computer is a plausible mechanism for the simulation of global cognitive processes, the answering silence would be deafening.[99]

But, if we give up the cognitivist assumption that we have to relate isolated meaningless facts and events to each other, and we see that all facts and events are experienced on the background of a familiar world, we can see the outline of a solution. The handicapper has a sense of which situations are significant. He has learned to ignore many anomalies, such as an eclipse or an invasion of grasshoppers that have so far not turned out to be important, but, given his familiarity with human sports requiring freedom from distraction, he may well be sensitive to the anomalies mentioned above. Of course, given his lack of experience with the new anomaly, it will not show its relevance on its face and summon an immediate appropriate response. Rather, the handicapper will have to step back and *figure out* whether the anomaly is relevant and, if so, how. Unfamiliar breakdowns require us to go off-line and think.

In his deliberations, the handicapper will draw on his background familiarity with how things in the world behave. Allergies and arguments normally interfere with one's doing one's best, etc. Of course, given his lack of experience with this particular situation, any conclusion he reaches will be risky, but he can sense that a possibly relevant situation has entered the horizon of his current task and his familiarity with similar situations will give him some guidance in deciding what to do. While such a conclusion will not be the formal computational solution required by Cognitivism, it is correlated with Freeman's claim that on the basis of past experience, attractors and whole landscapes can directly influence each other.[100] This suggests that the handicapper need not be at a loss; that this extreme version of the frame problem, like all the simpler versions, is an artifact of the atomistic cognitivist/computational approach to the mind/brain's relation to the world.

11. Conclusion

It would be satisfying if we could now conclude that, with the help of Merleau-Ponty and Freeman, we can fix what is wrong with current allegedly Heideggerian AI by making it more Heideggerian. There is, however, a big remaining problem. Merleau-Ponty's and Freeman's account of how we directly pick up significance and improve our sensitivity

98 Hubert L. Dreyfus, *What Computers Can't Do* (New York, NY: Harper and Row, 1997), 258.

99 Jerry A. Fodor, *The Modularity of Mind* (Bradford/MIT Press, 1983), 128–129.

100 Freeman writes: "From my analysis of EEG patterns, I speculate that consciousness reflects operations by which the entire knowledge store in an intentional structure is brought instantly into play each moment of the waking life of an animal, putting into immediate service all that an animal has learned in order to solve its problems, without the need for look-up tables and random access memory systems." Freeman, *Societies of Brains*, 136.

to relevance depends on our responding to what is significant for *us* given our needs, body size, ways of moving, and so forth, not to mention our personal and cultural self-interpretation. If we can't make our brain model responsive to the *significance* in the environment *as it shows up specifically for human beings*, the project of developing an embedded and embodied Heideggerian AI can't get off the ground.

Thus, to program Heideggerian AI, we would not only need a model of the brain functioning underlying coupled coping such as Freeman's; we would also need—and here's the rub—a model of *our particular way of being embedded and embodied s*uch that what we experience is significant for us in the particular way that it is. That is, we would have to include in our program a model of a body very much like ours with our needs, desires, pleasures, pains, ways of moving, cultural background, etc.

So, according to the view I have been presenting, even if the Heideggerian/Merleau-Pontian approach to AI suggested by Freeman is ontologically sound in a way that GOFAI and subsequent supposedly Heideggerian models proposed by Brooks, Agre, and Wheeler are not, a neurodynamic computer model would still have to be given a detailed description of a body and motivations like ours if things were to count as significant for it so that it could learn to act intelligently in *our* world.[101] We have seen that Heidegger, Merleau-Ponty, and Freeman offer us hints of the elaborate and subtle body and brain structures we would have to model and how to model some of them, but this only makes the task of a Heideggerian AI seem all the more difficult and casts doubt on whether we will ever be able to accomplish it.[102]

We can, however, make some progress towards animal AI. Freeman has actually used his brain model to model intelligent devices.[103] Specifically, he and his coworkers have modeled the activity of the brain of the salamander sufficiently to simulate the salamander's foraging and self-preservation capacities. The model seeks out the sensory stimuli that make available the information it needs to reach its goals. Presumably such a simulated salamander could learn to run a maze and so have a primitive intentional arc and avoid a primitive frame problem. Thus, one can envisage a kind of animal Artificial Intelligence inspired by Heidegger and Merleau-Ponty, but that is no reason to believe, and there are many reasons to doubt, that such a device would be a first step on a continuum towards making a machine capable of simulating human coping with what is significant.

[101] Dennett sees the "daunting" problem, but he is undaunted. He optimistically sketches out the task:

> Cog, ... must have *goal-registrations* and *preference-functions* that map in rough isomorphism to human desires. This is so for many reasons, of course. Cog won't work at all unless it has its act together in a daunting number of different regards. It must somehow delight in learning, abhor error, strive for novelty, recognize progress. It must be vigilant in some regards, curious in others, and deeply unwilling to engage in self-destructive activity.

("Consciousness in Human and Robot Minds," IIAS Symposium, *Cognition, Computation and Consciousness*, Kyoto, September 1–3, 1994, in Ito, et al., eds., *Cognition, Computation and Consciousness*, Oxford University Press.)

[102] Freeman runs up against his own version of this problem and faces it frankly: "It can be shown that the more the system is 'open' to the external world (more are the links), the better its neuronal correlation can be realized. However, in the setting up of these correlations also enter quantities which are intrinsic to the system, they are *internal* parameters and may represent (parameterize) subjective attitudes. Our model, however, is not able to provide a dynamics for these variations ...." [W.J. Freeman and G. Vitiello, Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics, 21.]

[103] Freeman writes in a personal communication: "Regarding intentional robots that you discuss in your last paragraph, my colleagues Robert Kozma and Peter Erdí have already implemented my brain model for intentional behavior at the level of the salamander in a Sony AIBO (artificial dog) that learns to run a simple maze. See: Kozma R, Freeman WJ, Erdí P (2003) The KIV model—nonlinear spatio-temporal dynamics of the primordial vertebrate forebrain. *Neurocomputing* 52: 819–826. http://repositories.cdlib.org/postprints/1049, Kozma R, Freeman WJ (2003) Basic principles of the KIV model and its application to the navigation problem. J Integrat. *Neuroscience* 2: 125–145, and also in a prototype Martian Rover at the JPL in Pasadena: Kozma R (2005) Dynamical Approach to Behavior-Based Robot Control and Autonomy Biol Cybern 92(6): 367–379. And also in a prototype Martian Rover at the JPL in Pasadena.