Chapter 7

Metaphorical Mind Fields

Freud often compared the brain to hydraulic and electro-magnetic systems. Leibniz compared it to a mill, and I am told some of the ancient Greeks thought the brain functions like a catapult. At present, obviously, the metaphor is the digital computer.
—*John Searle*

The brain is said to use data, make hypotheses, make choices, and so on, as the mind was once said to have done. In a behavioristic account, it is a person who does these things.
—*B. F. Skinner*

So far, we've considered how our perceptual biases influence our tendency to anthropomorphize the world around us, and how, as big-brained mammals, we often fail to realize that much of the flexible ("intelligent") behavior that we see doesn't require very much in the way of a brain at all. We've also begun to explore some of the scientific biases that exist in psychology, and to see that alternative views are possible. In this chapter, we'll extend this argument and consider in more detail how one particular human bias, the one on which our scientific biases rest, may prevent us from appreciating what natural cognition is all about. Specifically, we'll consider the ways in which our scientific understanding of the world is structured by the use of metaphors, and how this has led to the dominant view of cognition as a brain-based process isolated from the world.[1]

What does it mean to say that we structure and understand our world through metaphor? In our everyday life, it refers to our tendency to understand certain abstract concepts in terms of other, more concrete, experiences. We understand the abstract notion of time, for example, by using spatial metaphors: we "look forward" to spring break, not least

because we've "fallen behind" with our work and deadlines are looming, although we don't worry too much about this, as the "past is behind us." Similarly, we often think of our thoughts and ideas as food: we like something we can "get our teeth into," although we often find that we may have "bitten off more than we can chew."[2] As you may have noticed, we often use metaphors based on the bodily actions we can take in the world (moving through time; chewing on ideas), and, as we'll consider in more detail in the next chapter, this may well occur because most of our understanding of the world is grounded in—and built up from—our ability to act in it, so that even the most abstract of ideas (not excluding mathematical thought, according to some authors)[3] reflect what our bodies can physically achieve.

Using metaphors in this way doesn't mean that we literally believe that our thoughts are food, and that we will starve if we don't get any. Rather, we understand that a similar relationship exists between the equivalent elements in both the concrete and abstract domains and this is why we can draw the comparison: our thoughts can be "intellectually nourishing," if not literally so. In the same way, we draw analogies between items (i.e., interpret one thing in terms of another) by understanding the similarity of the relationship that exists between them: a bird's nest is like a human apartment (the relation of "home"), a dog wagging its tail is like a human's smile (the relation of "friendly behavior"). The ability to see beyond (another metaphor . . .) the juxtaposition of different elements to the relationship that exists between them (i.e., moving beyond observable features), so that it is possible to identify a similar relationship between an entirely different pair of items, is argued to be a key human trait—perhaps even unique[4]—and one that allows us to contemplate and understand the world in a more complex and abstract way than is available to many other creatures. Why, then, might this kind of reasoning lead us astray? Surely it is an extremely useful skill?

The answer is that it is, obviously a very useful ability, and, most pertinently, one that, as noted at the beginning of the chapter, plays a large role in scientific thinking. In science, we often have to deal with highly abstract concepts that would otherwise be very hard to grasp, precisely because they are so far outside our everyday experience. Metaphors are, therefore, an essential part of science.[5] One suggestion is that metaphors help us extend the boundaries of our knowledge (itself a metaphor . . .)

via a process of "catachresis"—which means to deliberately use a word or term to denote something for which, without the catachresis, there simply is no name. When we do this, we can bring into being entirely new ways of thinking, and pursue ideas that would never have occurred to us otherwise.[6]

So, this way, the structure of the atom was famously likened to the solar system, and DNA is often seen as a form of digital storage device. As we've already discussed in chapter 1, the action of natural selection is often compared to the intentions (the desires and beliefs) of humans. As the latter case also illustrated, however, this kind of reasoning can create problems when the metaphors employed are taken too literally. The same goes for the conventional view of perception discussed in the previous chapter: the suggestion that our brains "make inferences," "test hypotheses," and "present arguments" is, at base, metaphorical. As we noted, brains cannot literally do any of these things, but the misconstrual of this metaphor (or the simple failure to remain aware that a metaphor is being used) can lead one astray. No doubt all the neuroscientists referred to in the previous chapter would emphatically deny that they are arguing for a homunculus in the head, but by speaking of the brain as "inferring," "perceiving," and "asking questions," that is exactly what they are doing.

In what follows, we're going to consider a very powerful metaphor that helped shape the fields of psychology, cognitive science, and artificial intelligence for many years, and which may explain why we often get trapped into anthropocentric ways of thinking about the cognition of nonhuman animals. Specifically, we're going to consider the way in which many neuro-, cognitive, and comparative psychologists liken the brain to a computer (the "inferences" and "hypotheses" view of perception discussed in the previous chapter is obviously one aspect of this view). Indeed, some people go so far as to argue that the (human) brain is not just analogous to a computer in a strictly metaphorical sense, but that it actually is a computer that takes in input, processes it in various ways, and then produces a specific output.[7]

As with our anthropocentric tendencies, our use of the computer metaphor is so familiar and comfortable that we sometimes forget that we are dealing only with a metaphor, and that there may be other, equally interesting (and perhaps more appropriate) ways to think about brains and nervous systems and what they do. After all, given that our metaphors for the brain and mind have changed considerably over time, there's no reason to expect that, somehow, we've finally hit on the correct one, as opposed to the one that just reflects something about the times in which we live. Socrates considered the mind to be a wax tablet; John Locke, the seventeenth-century British philosopher, famously considered the mind to be "a blank slate," on which our "sense data" were written or painted; and, as the epigraph opening this chapter suggests, Freud compared the brain to a hydraulic system (with all its connotations of pressure build-ups and the need for "release"). The mind/brain has also been compared to an abbey, cathedral, aviary, theater, and warehouse, as well as a filing cabinet, clockwork mechanism, camera obscura, and phonograph, and also a railway network and telephone exchange. The use of a computer metaphor is simply the most recent in a long line of tropes that pick up on the most advanced and complex technology of the day.[8] This, in itself, should make us somewhat skeptical about claims for the computerlike nature of the brain; what should really make us wary, however, is how the computer metaphor took hold in the first place. To grasp this, we need to consider a little history.

Artificially Anthropocentric Intelligence

> Artificial intelligence is no match for natural stupidity.
> —*Anonymous*

> Chess is the *Drosophila* of artificial intelligence. However, computer chess developed much as genetics might have if the geneticists had concentrated their efforts starting in 1910 on breeding racing *Drosophila*. We would have some science, but mainly we would have very fast fruit flies.
> —*John McCarthy*

The computer metaphor first rose to prominence in the early 1950s. Prior to this, the telephone exchange served as our best metaphor for the brain. Brains were considered to be electronic switching devices that connected a stimulus to a response in the same way that a telephone operator connected one caller to another.[9] As the most prominent school of thought in psychology at the time was radical behaviorism, this analogy worked extremely well: for the most part (there were some exceptions),[10] behaviorists dealt not with internal, mental processes, but with brain-body behavior considered as a whole;[11] more specifically, their concerns were with behavior that could be controlled as a response to a stimulus, via learning. As it became clear that this stimulus-response account of the behaviorist approach couldn't provide an adequate account for all that an animal was (or wasn't) capable of learning, the idea began to gain ground that some internal processing had to mediate between a stimulus and a response.[12] At the same time that psychologists were rejecting and rethinking behaviorism, computer scientists were developing what came to be known as "artificial intelligence," and using computers to simulate cognitive processes. Psychologists began to cotton on to the idea that understanding brains and intelligence could be achieved not only via the analogy of the computer, but also by the actual use of computers to model and mimic the activities of the brain.

When Is a Turing Machine Not a Turing Machine?

The British mathematician Alan Turing, often regarded as the father of computer science, is widely credited with developing the "brain as computer" metaphor owing to his analyses of "Turing machines";[13] these were very basic devices—consisting of a read-write "head" (like that on a tape recorder) that could print, read, and erase symbols on an infinite tape of paper—that manipulated symbols in a very precise way. It is important to be aware that Turing machines do not actually exist; they are entirely abstract descriptions of a computing device that could be used to solve logical problems, via an "algorithm" (a set of rules followed in sequence).[14]

One can "build" an infinite variety of Turing machines, each of which is capable of computing a single specific sequence of numbers depending on how its read-write head interacts with the symbols on the tape (i.e., based on its specific algorithm). Building on this idea, Turing proposed that it was possible to develop a "universal" Turing machine—one that would be able to simulate the operation of any other possible Turing machine—so that, instead of being able to compute only a single

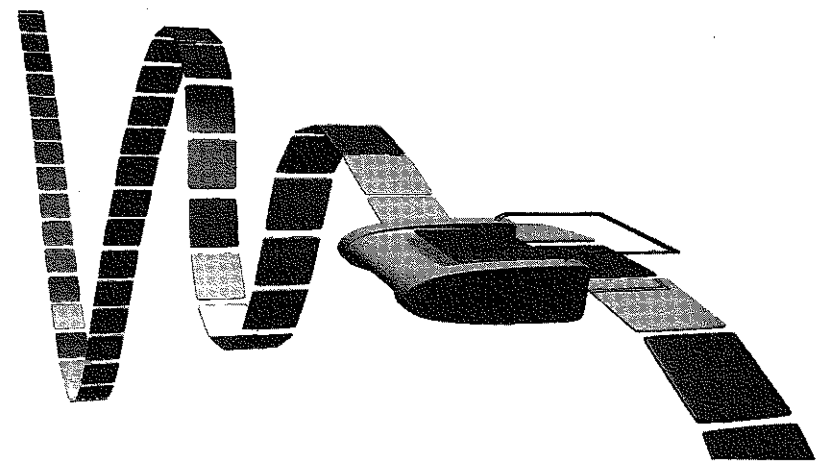

Figure 7.1. A Turing machine is an abstract device thought up by the British mathematician Alan Turing to demonstrate how numbers can be computed through the use of an algorithm (a set of rules followed in sequence).

sequence of numbers, a universal machine would be able to calculate any possible sequence of numbers, provided there was a specific Turing machine whose operations it could reproduce. Strange as it may seem, it was proved that this purely mechanical procedure—the "algorithm" used by Turing machines—could be employed to calculate the answer to any question that any other kind of computer could calculate (that is, not just mathematical questions, but any kind of question at all, provided it could be encoded by the symbols used by the Turing machine). This led to much excitement and speculation that perhaps human thought was a similar kind of algorithmic, symbol-manipulating process, and, even more excitingly, perhaps the brain was a real-world universal Turing machine.[15] This opened up the possibility for modeling human thought, language, perception, categorization—whatever process one liked—using a digital computer. Why? Because, like a universal Turing machine, and supposedly like the human brain, computers use algorithms to perform calculations (or computations).

As a result—and as computers became a reality, and not just theoretical proposals—it seemed possible that humans would be able to create a brain capable of humanlike thought using human-made silicon chips, instead of biologically evolved neurons.[16] This follows logically because the computational processes of a Turing machine do not depend on the actual materials that are used to make it: a Turing machine can be made out of anything you like, not just paper tape and magnetic heads, but anything at all, from "two kinds of pebbles and a roll of toilet paper," as Jerry Fodor once put it,[17] to "cats, mice and cheese," as the philosopher Ned Block once suggested.[18] Moreover, because the brain alone was seen as the key to understanding cognition—based on the idea discussed in the previous chapter that, in order for us to perceive and think about the world, representations of that world must be constructed to compensate for the poor quality of the information received by the sense receptors—it meant that bodies and the environment became completely irrelevant to the study of cognition. This further reinforced the idea that it was possible to create humanlike intelligence in a computer; a computer can be considered equivalent to a brain, but not to an active, moving body.

From this point on, psychological processes—in both human and nonhuman animals—became closely identified with various kinds of "information processing." The idea was that sensory input came into the cognitive system; the cognitive system algorithmically manipulated symbols,[19] as would a Turing machine/digital computer, and then produced an output that manipulated the body. It is at this point that the clear separation of perception, cognition, and action, which we have noted in earlier chapters, began to be made, and efforts to understand the workings of the mind (and "thought" and "intelligence") came to mean efforts to identify and understand the "information processing" that occurred between sensory input and motor output. With the computational metaphor in place, it became almost inevitable that the brain would be seen as the equivalent of computer hardware, with cognitive processes operating like the brain's software: an idea that has permeated modern Western culture at all levels. In the film *The Matrix*, for example, it was possible to download computer programs directly into people's brains via a portal at the back of the head, obviating a long-drawn-out learning process and providing the recipient with expert abilities in, among other things, kung fu (again emphasizing that the body is largely irrelevant to the development of even such highly physical skills; a dangerous assumption, as we shall see in chapters 9 and 10).

Aside from the problems of neglecting bodies and environment, another problem of the development and use of the "brain as computer" metaphor is—as Andrew Wells points out in his marvelous book on the subject—that it completely misrepresents a Turing machine and also Turing's aim in developing them.[20] To get the full story, and so understand what Turing was attempting, you should really stop here and read Wells's book for yourself, but assuming you don't do that (even though you should), a brief summary will serve our purposes.

When Turing's paper was published, way back in 1936, a "computer" was not a machine, but a person. A person who computed sums. Turing's aim was to try to find a way of mechanizing this process, thereby producing a labor-saving device that could do the work of human computers. As we noted above, Turing conceived of his machine as an infinite paper tape, divided into squares on which symbols could be read and printed. This tape passed through a head that could move either to left or to right, one square at a time, and this head could both read what was written on the tape and print on it. In most books and articles in which a Turing machine is discussed, this whole kit and kaboodle is used as an analogy for the mind or for cognitive processes: inside our heads, it is argued, we have a Turing machine that receives input in symbolic form, manipulates it, and then provides an output. The tape of a Turing machine is, in essence, a model for human memory.[21]

Now, the truth of the matter is that this couldn't be further from what Turing was actually attempting to model. Remember that he was trying to conceive of a machine that could calculate sums in the same way that a human computer calculated them. How do we calculate sums? If they're long and complicated, most of us do it on a piece of paper—maybe even graph paper—using a pen or pencil. In this light, let's consider the abstract Turing machine again. The paper tape that is usually seen as internal memory was, for Turing, part of the environment. Specifically, it represented the paper on which a human computer could work out his or her sum. So the paper tape is not a model of memory in the head but a model of graph paper in the environment.[22] Equally, the "machine-head" that reads and writes the symbols does not represent the cognitive processes taking place inside a person's brain, but instead represents the person as a whole, using pen and paper to calculate sums. Wells refers to this setup as a "mini-mind" to get across the idea that these can be either complete descriptions of very simple minds or, alternatively, partial descriptions of more complicated minds (after all, there is more to a person than simply computing sums). So a Turing machine actually consists of a mini-mind that has a finite number of states (because human memory has finite limits), and an infinite tape divided into squares (because, when real people do real calculations, their ability to do so is not usually limited by their having access to only a fixed amount of paper). The combination of the state of the mini-mind and the contents of the tape is called a "configuration." The current configuration determines the moves the machine makes, what it prints, and what the succeeding configuration will be.

It couldn't be clearer from this description, then, that the symbols a Turing machine manipulates are outside the mind, and not part of it.[23] A Turing machine is, as a result, a very ecological contraption, in Gibson's sense of the word. Computing is about the relationship between a human computer and his or her environment (which consists of the paper and pencil used to do the sums). One cannot understand the behavior of a Turing machine simply by looking at the state of the mini-mind (the person, if you like), nor can one understand what the computer will do just by looking at the tape (the environment). To understand a Turing machine's behavior, one has to look at the relation between the agent and the environment. Wells uses this insight to argue that a mind is both formed and maintained by the continuous interactions between an agent and the environment.[24] Turing modeled exactly these kinds of interactions, but only in a very specific context. It was never his intention to provide a general analysis of human behavior, nor to suggest that all human cognition conformed to this specific kind of computational process. Indeed, Turing's concerns were clearly mathematical, rather than psychological. He was simply interested in what numbers it was possible to compute, as a human did, using a pencil and paper.[25]

So, if Turing's machines were never intended to be a model of the mind or of mental processes, where did our current idea of the brain as a computer come from? For the answer, we have to cross the Atlantic. The first real-world version of a Turing machine was constructed for the United States army and known as the ENIAC (Electronic Numerical Integrator and Computer). Owing to the way it was built, and the fact that it was a special-purpose Turing machine (rather than a "universal" one), the computer's entire physical hardware had to be changed and rewired

every time a new kind of calculation was needed. In the late 1940s, John von Neumann was one of several people charged with the task of making the ENIAC more convenient and useful, and it was he who designed the architecture used by all modern computers today: a central processing unit, a main memory, a set of peripherals (like keyboard and monitor), and a second memory that could be used to store information externally, like hard drives, CDs, and memory sticks. It is, therefore, to von Neumann that we owe the "brain as computer" metaphor, as it was he who helped create self-contained digital computers. In addition, it was he who specifically compared his computer architecture to that of the brain, suggesting that the central control (CPU) of his computer corresponded to the "associative" neurons of the human nervous system, and that the input and output devices were the equivalents of sensory and motor neurons, respectively.[26] This "von Neumann architecture" is one that has been used in many different kinds of artificial intelligence projects and programs, and it is this, rather than a universal Turing machine, on which our metaphors of mind are based. Our notion that Turing machines represent the basis for our current view of cognition is completely off-track.

I mention all this here, of course, to highlight the possibility that, had people recognized the true psychological implications of Turing machines (that they reflect the ongoing mutual relationship between a "computer" and her environment, and are not a model of the mind divorced from the environment), we might have had a very different view of cognition and the brain, and a different kind of psychology might have been the result. Indeed, this is Wells's point: he explicitly shows how one can marry Gibson's ecological theory with Turing's theory of computation to provide a formal model of affordances (one that works better than the available alternatives),[27] and one that can serve as an alternative model of cognition. Space doesn't permit a detailed examination of Wells's argument here, but, in essence, affordances can be characterized and studied as the "configurations" of a Turing machine (the state of the "mini-mind" and the contents of the tape), an idea that captures the complementarity between animal and environment that is essential to Gibson's theory.[28] Just as affordances "point" both ways—toward the animal and the environment—so do the configurations of a Turing machine. The Turing machine model also gets at the issue of internal structure in the organism versus external structure in the environment, which has been the source of much of the criticism of Gibson's model. As Wells notes, there is a theoretical trade-off between internal and external structure in a Turing machine: a machine with only two possible internal states can compute numbers provided it has access to an external alphabet that is large enough. Equally, a Turing machine that has access to only a two-symbol alphabet can compute numbers provided it has a large number of possible internal states. The Turing machine model suggests, therefore, that structure in the animal will complement structure in the environment.[29] This means that one cannot simply assume a particular behavior results only from structure in the animal versus that of the environment or vice versa: it should reflect a trade-off between these two, and to discover what that is, you have to go and find out (the point we made in the previous chapter).

Finally, the idea of a universal Turing machine, once properly understood, also tends to support a Gibsonian view of the world. Unlike those Turing machines that compute only a specific sequence of numbers, and which always start on a blank tape, the universal machine works by beginning on a tape that already contains a string of symbols, which allows it to produce the output of the machine it is simulating. As the tape is actually part of the environment, a universal machine supports the notion that the information available for perception is found mainly in the environment, and not in the head.[30] Wells's combining of Turing's theory with Gibson's theory is, in a way, wonderfully subversive, because it brings together the most cognitive of all models in psychology—the Turing machine as isolated brain—and marries it to a theory that requires complete complementarity between organism and environment. Of course, in another way, it is not subversive at all, because it is merely correcting the misconception that the notion of a Turing machine supports the "brain as computer" metaphor that currently holds such sway.

Alternative Metaphors for the Brain?

Looking at Turing machines from an ecological perspective, and highlighting the differences between Turing machines and von Neuman architecture is a point well worth making because, although the computer analogy built on von Neumann architecture has been useful in a number of ways, and there is also no doubt that work in classic artificial intelligence (or, as it's often known, Good Old Fashioned AI: GOFAI)[31] has

had its successes, these have been somewhat limited, at least from our perspective here as students of cognitive evolution.

As a number of cognitive scientists and roboticists have pointed out over the years,[32] the classical AI perspective, with its emphasis on the algorithmic manipulation of symbols using a von Neumann architecture, naturally gravitated toward those aspects of cognition, like natural language, formal reasoning, planning, mathematics, and playing chess, in which the processing of abstract symbols in a logical fashion was most apparent. As a result, classic artificial intelligence also placed humans front and center, with the focus of research resting squarely on understanding some peculiarly human aspects of intelligence: none of them are very athletic—they don't require an active organism in the Gibsonian-sensorimotor sense—and none of them require any specific interaction with the environment, as opposed to seeing the environment simply as the arena in which the products of these computations are played out. Unfortunately, this rather arbitrary initial emphasis on these particular (and specialized) kinds of logical, algorithmically based tasks gained such momentum that researchers came to the conclusion that everything brains did (human and nonhuman alike) was simply a form of logical reasoning, and that they employed an algorithmic process to achieve this. I say "unfortunately" because, while this view (eventually) managed to generate a computer that could beat the world chess champion, it has, so far, failed to give us any real insight into the mechanisms that underlie the more natural forms of intelligence we've been discussing up to now: how adaptive behavior is produced in a changeable environment. In human terms, this would include things like how we recognize a face in a crowd, how we coordinate our movements and manipulate all the objects necessary to make cup of tea, or even something as apparently simple as how we manage to walk, run, and even hop over uneven ground without falling flat on our faces.

You should now begin to see the problem. Our metaphor of the brain—and hence of cognitive processes—is one that was originally derived from a heavily anthropocentric focus on a few peculiar human cognitive achievements, all of which involved abstract symbol manipulation. As we've now seen, this in itself was derived from a misreading of Turing's work on computable numbers: work that made no claims of generality as far as psychology and cognition were concerned, but dealt only with a very specific human activity (so one cannot accuse Turing of misplaced anthropocentrism; he was quite clear about the aim of his work). As Wells shows, when properly understood, the Turing machine model can be seen as a critique of current cognitive approaches, one that supports the underlying philosophy of ecological psychology[33] (and, as we shall see later, the ideas of "embodied" and "distributed" cognition). Although it is true that certain aspects of our cognition can be understood and analyzed as computational processes involving the manipulation of symbolic representations—or, as some would suggest, are best understood via this kind of analogy—you should now appreciate that this isn't quite the whole story: not for us, and certainly not for other, nonlinguistic species.

As we've noted in our consideration of both Gibson's and Turing's work, the missing ingredient in all this is the recognition that the body and the environment really do matter as far as cognition goes. After all, when brains evolved initially, they did so in animals that already possessed bodies, and long before they possessed anything that we would recognize as a brain.[34] By failing to account for this (and indeed by completely misinterpreting the nature of a Turing machine itself) the computer metaphor has generated a view of cognition as a process that has no real link to the body or the outside world, taking place purely in the brain alone.

What is worse is that we have taken this strange view of cognition—that it takes place inside the "Turing machine" of the brain and involves the disembodied, logical manipulation of internal representations—and applied it directly to other animals. The metaphor of the computer and the idea of a computational, representational mind is one that pervades studies of comparative cognition[35] (even those articles that are critical of the more anthropocentric/anthropomorphic interpretations of such studies nevertheless take the existence of von Neumann–like computational, representational processes to be axiomatic, rather than an assumption to be tested).[36] What is also interesting is that we have applied the computational model to other animals even though it doesn't adequately explain most facets of our own natural cognition (leaving aside the fact that this kind of psychological generality was never the intention of the earliest proponents of this model in the first place).[37]

By presenting an essentially disembodied view of cognition, the computer metaphor, with its input-output (stimulus-response) structure, also

suggests a very static picture of animal life. Like the computer on your desk, an animal just has to sit there until a stimulus (an input) comes along and causes it to act. As we have seen, the vast majority of animals are not passive in this sense: they seek out relevant and significant resources in the world, in an active, animate fashion. This again is a consequence of their possessing a body and not only a brain (and of possessing bodies before brains). To ignore the body and the environment when considering how animals behave in an "intelligent" fashion is, at the very least, to miss out on half the story. Indeed, we've already seen how the ears of the cricket and the eyes of the *Portia* spiders are highly relevant factors to consider when we are trying to understand the behavior of these animals in their natural environments. As we'll discover more fully in the following chapters, we are likely to gain a better understanding of the natural kinds of intelligence that we see every day (and engage in ourselves) only when we take the body as seriously as we do the brain.

Returning to the Age of Steam

> Steam is no stronger now than it was a hundred years ago,
> but it is put to better use.
> —*Ralph Waldo Emerson*

Given these problems with the "brain as computer analogy," how, then, should we think about cognitive processes? One solution, as we've seen, is the "ecological" computational approach suggested by Andrew Wells. We can also, however, consider other kinds of models besides these computer-based metaphors. Or, as Tim van Gelder, a philosopher at Melbourne University in Australia puts it, "What might cognition be, if not computation?"[38] His suggestion follows on from that of the ecological psychologists, to some degree, by recognizing the dynamic way an animal's sensory systems interact with its motor systems and how these interact with the world. As the nervous system, body, and environment are simultaneously changing and influencing each other in a continual cycle of adjustment (they are "dynamically coupled"), we should properly consider a "cognitive system" to be a single, unified system that encompasses all three elements and doesn't privilege the brain alone (especially a disembodied and autonomous one, translating abstract input relations into similarly abstract output relations).[39] Interestingly, this notion of dynamic coupling, where each change in one element of a system continually influences every other element's direction of change, can be captured in another machine-based analogy. As van Gelder suggests, a better model for how cognition works may be not a modern digital computer, but something like a Watt governor.

A Watt governor, also known as a flyball or centrifugal governor, is a device used to regulate the speed of a steam engine, regardless of changes in the workload of the engine or the fuel supply. It was named after James Watt, who designed some for use on the first steam engines (although it should be noted that Watt himself didn't invent the governor: governors of a similar design had already been in use in windmills for many years). The governor consists of two flyballs (hence the name) connected to a spindle by two flyball arms. The spindle is attached directly to the shaft of the steam engine. If the speed of the spindle increases, the flyball arms move upward owing to centrifugal force. The clever bit comes in here: the flyball arms are connected to a throttle valve that regulates the amount of steam that enters the engine. When the engine speeds up, the upward motion of the flyball arms closes the throttle valve, thereby reducing the steam input to the engine and slowing it down. Of course, as the speed of the engine falls, so, too, does the spindle, which means that the flyball arms drop. This has the effect of opening up the throttle valve, which allows more steam into the engine, which then speeds up.

As a consequence of this constant adjustment of the spindle, flyball arms, and throttle valve, the engine maintains a constant speed through smooth and swift adjustment, despite fluctuations in the steam pressure and workload. It should be apparent that, despite the way I've described it above, it is very difficult to identify a discrete sequence of events in a flyball governor because everything is happening continuously and smoothly, all at the same time: the angle of the flyball arm determines the speed of the engine, but, of course, it is the speed of the engine that determines the angle of the flyball arms: the angle of the flyball arms and engine speed are both determining, and determined by, each other. The Watt governor therefore solves the problem of constant engine speed in an entirely noncomputational, nonrepresentational way.

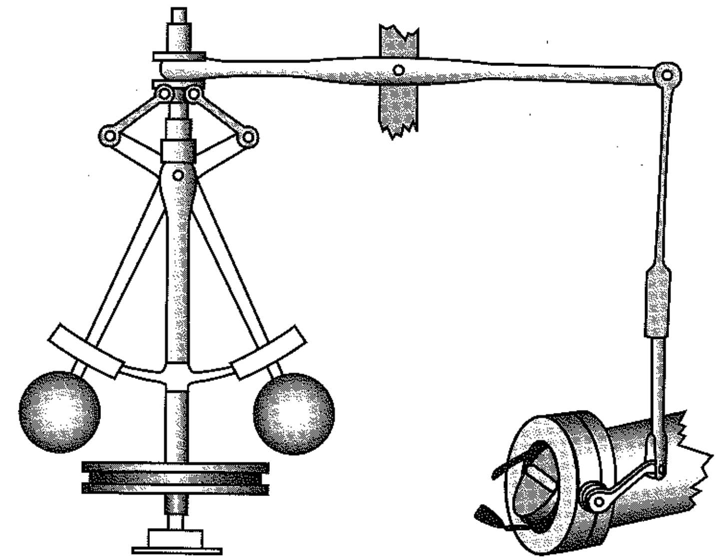

Figure 7.2. A Watt governor for a steam engine. The flyball arms are connected to a throttle valve that controls the amount of steam entering the engine and a spindle that is connected to the drive shaft of the engine.

Of course, as van Gelder (1995) argues, one could, in principle, come up with a computational, representational governor that would do much the same job. Van Gelder himself produced an example of just such a computational algorithm:

1. Measure the speed of the engine.
2. Compare this speed against the desired speed.
3. If no discrepancy was detected, then return to step 1. Otherwise:
   a. Measure current steam pressure.
   b. Calculate desired alteration in steam pressure.
   c. Calculate necessary throttle valve adjustment.
4. Make throttle valve adjustment.
5. Return to step 1.

What we need to remember here, however, is that this computational solution, while fine in principle, isn't the one that actually solved the problem of controlling engine speed, at least partly because the technology needed to implement such a computational approach wasn't available at the time. However, it is also noteworthy that the governor that was invented did its job superbly well. It wasn't just an inferior and primitive fix, the best that could be done in the absence of computer technology. It is also true that the computational solution is a lot more complicated in terms of the parts needed and operations performed. There have to be devices that can measure the relevant parameters, as well as devices to implement the response needed. A computational governor is therefore likely to be more expensive to build and run than the Watt governor, and there are more parts that can go wrong. This is relevant from our evolutionary perspective, since evolution is a thrifty process and tends toward the cheapest possible route to solve a problem effectively. A computational solution is not the only possible way to solve the problem of variable engine speed, as the Watt governor demonstrates, and we should take this lesson to heart: just because one can very easily come up with computational solutions to problems, including those of animal cognition, we should not be misled into thinking these are the only solutions possible; other, potentially cheaper, equally effective, solutions may be there for the asking. The noncomputational solution to the problem of engine speed is in no way inferior to the computational one; it is merely different, but it does the job as well, if not better, and at a lower cost.

Still, even if we agree that the Watt governor is just fine, and there is no necessity to replace it with a fully computational algorithmic device, one could still make a case that the Watt governor itself is, in fact, using representations and is, therefore, a computational device. One could argue that the angle of the flyball arms does, in fact, "represent" the speed of the engine because the angle of the flyball arms is correlated to engine speed. One could, in principle, use this angle to stand in for how fast the engine is running.[40] This does, however, miss a very important point about how the governor does its job: although there is indeed a correlation between angle arm and engine speed, the angle of the arms is at all times determining the amount of steam that can enter the engine, and hence at all times the speed of the engine depends on the arm angle, just as much as arm angle depends on engine speed. To argue that one "represents" the other is massively oversimplistic, and it also fails to capture the fact that the system is dynamic and in a constant state of flux. Consequently, if the governor is not strictly representational, then it can't be computational (if we stick with the definition that computation involves the rule-determined manipulation of representations). Without representations, and because of the mutually determining nature of each element in the governor, one cannot identify any discrete algorithmic steps in the operation of the governor, and there is a sense in which the system simply cannot be considered computational (but we will revisit this below in slightly different terms). The cut-out-and-keep message here, then, is that cognition need not be—either by definition or by logical inference—a purely computational process. It also suggests strongly that flexible, intelligent systems need not be separated into "hardware" and "software" components (the sticky "wetware" of the brain and its cognitive processes)—they are one and the same. In other words, although the computer metaphor has been, and perhaps still is, useful in helping to predict and explain certain aspects of (human) psychology, we shouldn't make the mistake of thinking that this means that natural cognition really is computational and therefore that the brain really is some kind of biological computer.

Timing Is (Almost) Everything

> Observe due measure, for right timing is in all things the most important factor.
> —*Hesiod*

Of course, cognitive systems are no more like Watt governors than a brain is like a computer—not literally. As we've noted, both are merely metaphors. The computer metaphor has, however, been taken both very literally and very seriously, and has promoted a very particular view of cognition that has been widely and wholeheartedly adopted by many researchers.[41] Looking to the Watt governor as an alternative metaphor is useful not because there is any suggestion that cognitive systems actually work in this way, but because cognitive systems may be better understood as "dynamical systems" where inputs, internal processes, and outputs—or, to put it more concrete terms, the environment, the brain, and the body's actions—are coupled like the spindle, angle arm, and throttle valve of the governor. Dynamical systems present us with more useful means for understanding and thinking about physically embodied, environmentally embedded organisms than do standard computational models.

I want to expand on this point a little further because, as our discussion of Turing machines above makes clear, if we define a dynamical system as one that shows state-dependent change (i.e., the future state of the system depends causally on the current state of the system), then computational systems are, by definition, dynamical systems.[42] In a Turing machine, the future state of the tape depends causally on the current state of the head and what is currently written on the tape, and this represents the coupling of the mini-mind with the tape environment. Looked at in this way, computational systems can be seen as a specific subset of the kind of dynamical systems that includes the Watt governor.[43] This inclusive definition means we can account for all cognitive processes using a dynamical systems approach (potentially anyway—we are actually nowhere near doing so), without being forced into a situation where we're trying to explain how two very different sorts of processes—computational and dynamical—came into being, and how they fit together.

So far, so good. But if computational systems are dynamical systems, what, then, is the real difference between these kinds of systems and a Watt governor? Michael Wheeler identifies at least two factors that seem to be key in differentiating between them.[44] First, computational systems, by definition, involve the use of representations: to do their job they must access, manipulate, and transform symbols. As we noted above, if one felt really strongly about it, one could make a case for a representational version of a Watt governor, but we also showed that representations were not essential to get the job done (that was the whole point of the example). So that's the first difference: a computational dynamical system absolutely requires representations, whereas a noncomputational dynamical system does not.

The second difference is more important, and, much like good comedy, it's about timing. In a computational system, time is reduced to the mere sequencing of events; in a Turing machine, things have to happen in the right order, but the time it takes for transitions between states to occur is not dealt with at all, and there is no specific theoretical reason why things should happen in a specified amount of time. Similarly, the

amount of time the machine should remain in a given state is not considered, because, in a Turing machine, this would serve no specific function. Time simply doesn't matter. Of course, in the real world, one could argue that computational events have to occur swiftly enough to enable the problem to be solved in good time, but outside of that, time has no role to play.

As we saw in the Watt governor example, this isn't true of noncomputational dynamical systems. Instead, they exhibit "richly temporal phenomena."[45] This means simply that the actual rates and rhythms that characterize a particular process play an important and central role in getting the job done. This could be the way that the underlying physical processes of the brain work (how long it takes for a neurotransmitter, like nitric oxide or glutamate, to diffuse through the brain, for example, or how long it takes for such neurotransmitters to modulate neuronal activity), which in turn could affect the specific durations or rates of change in other physiological processes. Similar intrinsic rhythms in the body may also be important, as will other aspects of bodily dynamics that relate to, for example, the mechanical properties of muscle, which dictate where and how fast an animal can move. These bodily processes may, in turn, need to be synchronized precisely with temporal processes occurring outside of the animal in the environment.

This issue of timing is very clear in our Watt governor example, where the coupling of the different parts, and the specific rhythm and timing they displayed, were crucial to its success in controlling engine speed. Interestingly, when more sophisticated governors were developed, they showed behavior that was much less effective than that of the earlier models (which is rather counterintuitive): the new "improved" models "hunted" for a steady speed, continually speeding up and slowing down, rather than smoothly maintaining a steady state. This was because superior manufacture of the component parts meant that they generated less friction, and this, in turn, meant that speed adjustments were effected much more quickly. Greater friction in the older models meant that any changes in engine speed took longer to feed though the system, and this intrinsic quality helped the governor to perform the job at hand more effectively. Of course, friction and heat are features of computational systems as well (this is why your computer has a fan built into it), but the point is that, in a computational system, these are merely problems to be overcome by engineers, not an integral part of the computational process.

Thinking in terms of dynamical systems with a "rich temporality" also provides us with a new way of viewing the "failure" of evolved knowledge in the face of environmental change. Returning to our digger wasps, we can see their routine—preparing a chamber, inserting a bee into it, and laying their eggs—as a dynamical interaction of the wasp's internal state (of readiness for egg laying), its actions in the world (hunting and chamber preparation), and its environment (the presence of the chamber, the proximity of the bee to the entrance). The "failure" of the wasp to begin its "routine" in the middle is a failure only if we assume there is an underlying algorithm being followed. If, instead, we consider that the wasp's brain and body are making continual adjustments to an environment that is continually being changed by the presence of the wasp (and so changing the wasp's state at the same time as the wasp changes the state of the environment), we're less inclined to see a failure and more aware of the fact that we're watching a dynamically coupled system in action.

One must be somewhat cautious in adopting a more dynamical approach, however. In particular, the philosopher of cognitive science Andy Clark has noted that, because dynamical systems approaches are concerned with the state of a system as a whole—so-called total state explanations—we can potentially lose as much as we gain from adopting this approach over the computational approach. Clark's argument is that a dynamical approach obscures the "intelligence-based" route to evolutionary success that characterizes living cognitive systems, as compared to the other kinds of physical dynamical systems that exist in the world, such as river flow systems.[46]

As we noted in chapter 5, brains evolved in order to allow animals greater control over their environments and their destinies. Although we have spent a lot of time in this chapter putting the brain in its place, it would be foolish to suggest that brains don't matter. Brains are crucial as a location of behaviorally relevant activity, and this, as Clark notes, must mean that brain-involving dynamical systems are very different from other kinds of dynamic physical systems.[47] Brain-based systems achieve the kinds of behavioral flexibility that we're interested in precisely because the brain is able to alter the "information flow" through the system cheaply and in a wide variety of ways. If we deal only with the overall

state of the cognitive system, then those aspects of how information flow is specifically channeled and directed by the brain get lost. If, however, we are mindful of this possibility, and we consider the inner flow of information within the brain as seriously as we do the overall state of the system, we can generate what Clark calls a "powerful and interesting hybrid: a kind of dynamical computationalism." By this, he means we could combine the "standard" computational and information-processing concepts with the coupling and richly temporal phenomena of truly dynamical systems.

His suggestion is that, rather than treating computational systems as fundamentally different from noncomputational ones as described above, we should attempt to combine the two so that the conventional computational approach is given a new dynamical dimension. His argument, then, is to take Wheeler's idea of computational systems as a specific subset of dynamical systems but to try to erode the distinction between them by allowing richly temporal phenomena to transform the standard computational approach. This may well be a productive way forward: as the complexity of sensory, motor, and physiological systems increases, and more complex behavior is possible, then, as we mentioned earlier, one would predict that the brain would have to be more strongly involved in altering information flow through the brain-body system in order to provide the kinds of temporal coordination needed to permit temporally rich adaptive behavior to emerge.

With this caveat in place, a dynamical systems approach, with its emphasis on rates, rhythms, and synchrony, is preferable because it is one that, by definition, naturally gives body and world their due when it comes to cognitive processes because, as Wheeler makes clear, these nonneural components will also act as pacemakers and rhythm-setters in causally important ways, in conjunction with those taking place in the brain.[48] Even better, perhaps, a dynamical systems approach treats the brain as an integral part of the body, and not as the all-powerful highly privileged computer that "tells" the body what to do. Like nonneural bodily processes, the neural activity of the brain has its own intrinsic rhythms and undergoes change at different rates. These, in turn, must be synchronized with the events that are happening in the body and the environment to produce effective behavior. The standard computational model, which keeps perception, action, and cognition as separate, independent processes, and (implicitly) assumes these need occur only in sequence, but not in real time, is both fundamentally "disembodied" (because cognition does not depend on any aspect of an animal's intrinsic physicality) and "disembedded" (because the environment plays no intrinsic role in helping to regulate the cognitive system, but is merely the "stage" on which the products of a disembodied cognitive process are played out). We want a richly rhythmic time-dependent view that accords with the lives of real, richly rhythmic time-dependent animals, and so we shall continue to pursue a dynamical approach in the next chapter.

CHAPTER 8

There Is No Such Thing as a Naked Brain

You've got the brain of a four-year-old boy, and I bet he was glad to be rid of it.
—*Groucho Marx*

We can discover more about the dynamical approach to animal cognition and behavior by moving away from the more abstract systems of the last chapter, and taking a look at real brains, and the ways in which they are coupled to the environment. Walter Freeman, a neurophysiologist at Berkeley, has spent the last thirty or so years performing intricate and meticulous experiments on smell, vision, touch, and hearing in rabbits (mainly) and has worked out a model of learning based on the kind of dynamic coupling between brain and environment suggested by the dynamic systems approach.[1] Before we can go into detail about Freeman's work, however, we first need to cover a little more ground on the theory behind dynamical systems, so that we can more fully appreciate Freeman's views on how brains, bodies, and the environment fit together.

Mathematically speaking, a dynamical system consists of a number of "state variables" (e.g., the engine speed and flyball arm angle in the Watt governor) that specify the state of the system at a given time, along with a set of equations that describe how those variables change over time. There can also be certain values that specify quantities that can change the state of the system, but aren't themselves changed as a result: these are called the parameters of the system. Putting everything in these terms allows us to think of a dynamical system as a form of graph—a multidimensional "phase space"—where the number of dimensions is set by the number of state variables of the system. In such a phase space, each possible state of the system (all the possible combinations of all the state

BEYOND THE BRAIN

How Body and Environment Shape Animal and Human Minds

Louise Barrett

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD