1

Introduction

How is the brain related to the mind? Do our minds work like computers? Can science's new knowledge about the brain tell us anything of importance about the way our minds work? How can a three-pound mass of tiny jelly-like blobs connected by vast numbers of microscopic filaments be the basis of all our thoughts, feelings, memories, hopes, intentions, knowledge? Will all the new knowledge scientists are gaining about our brains enable them to read our minds with electronic devices?

There is a new way of thinking about the mind and the brain which takes it for granted that the human mind is inseparable from the human body. Since the evidence indicates that the center of our mental activities is the brain, the advocates of this approach try to understand the functioning of the mind on the basis of what we know about the functioning of the brain. This new scientific paradigm has led to the construction of new theories and models of the mind which are variously known as connectionist theories, or neural network models, or theories of parallel distributed processing (PDP). Although some of the ideas on which these new theories are based have been around for over a century, the detailed working out of these models began only in the 1970s.

These new ideas are called connectionist theories because they claim that our mental processes and capacities – how we perceive what is out there in the world, how our knowledge about these things is organized, how we combine all this information to draw new conclusions, how we decide what to do next in order to get what we want – can be explained on the basis of what is known about the multiple interconnections between the neurons, or nerve cells, in the brain. They are called neural network models because they present detailed computer models of how interconnected units can work together to form networks analogous to those in the brain. And they are called theories of parallel distributed processing because they claim that a variety of mental operations are carried out at the same time, in parallel, and that these operations are distributed over large numbers of units rather than occurring within individual units separately. Each of these concepts will be explained in detail shortly, but first I would like to say a little more about the relation between the mind and the brain.

How are the mind and the brain related?

The connectionist view is based on the idea that there can be different levels of explanation for talking about the same thing. The concept of levels of explanation is well known in such disciplines as physics. For example, there is a difference between the level of our ordinary talk about tables and the atomic level of description. In our ordinary way of talking, a table is a solid object that entirely fills the space it is in and can have a smooth surface with no bumps on it. On the atomic level, in contrast, there is a great deal of empty space between the atoms that make up the table, and since these atoms are constantly jiggling around, there is no clear boundary between the top of the table and the air above it.

Similarly, there are both a mental and a physical level of explanation for talking about human mental functions. When we use the word "mind," we are on the mental level of explanation. It is on this level that we talk about seeing a sunset, remembering our trip to the Grand Canyon, knowing that a canary is a bird, and knowing how to tie our shoelaces. When we use the word "brain," we are on the physical level of explanation. On this level we can talk about individual nerve cells firing when they are activated by other nerve cells, the arrangement of nerve cells into columns in certain parts of the cortex, and the fibers that link one part of the cortex with another. But when we talk about red and blue color receptors being activated in the visual cortex, or the supplementary motor cortex sending electrical impulses to the primary motor cortex so that the muscles in our hands will contract in a particular way in order to cross one end of the shoelace over the other, we are combining two different levels of explanation of the same event – the mental and the physical, the mind and the brain. "Red" and "blue" are on the mental level of explanation, "receptors" on the physical level.

In the field of perception almost all scientists use the physical level of explanation in trying to understand the mental one. At least part of the explanation of the way we see requires an understanding of how the brain processes the signals coming in from our eyes. In contrast, many scientists working in the field of the higher mental processes – cognitive scientists, who study such topics as memory, language processes and the organization of concepts – claim that the mental level of explanation can proceed independently of the physical level. Most of them agree that the mind is inseparable from the brain, but they do not believe that it is necessary to ground the mental level of explanation in the physical one. Some cognitive scientists also claim that mental explanations of cognitive processes should be the same whether the processes are taking place in a human or a computer, for example; these researchers are often called students of artificial intelligence, or AI for short.

In contrast, connectionists believe that it is helpful to make use of our knowledge about the physical workings of the brain in our explanations of cognitive processes. The connectionist level of explanation may be thought of as a third level, intermediate between the mental and the physical ones. It is not identical to the physical level because it does not talk about individual nerve cells firing and is not concerned with the physical layout of the various parts of the brain. It is not identical to the mental level because it does not talk about our concepts, say, as abstract things that could be found in any entity that is able to process information. Rather, connectionists propose a description of mental processes that takes account of the physical structure of the brain and its interactions with the environment. As we shall see throughout the book, they try to explain how our concepts are formed, how they are related, and how they are used by looking at how the neurons in the brain are connected with one another and with the surrounding environment. Connectionists' explanations of mental processes are thus based on what we know about brain processes in much the same way as physicists' explanations of the different properties of wood and glass tables are based on the atomic composition of these materials.

How are connectionist explanations different from other explanations of mental processes?

The key to the difference between connectionist explanations and those given by other cognitive scientists, some of whom like to call themselves "classical," lies in the words "parallel" and "distributed" of the expression "parallel distributed processing." Classical explanations of mental processes describe them as taking place serially, one after the other. A good example is what is often called "memory search." When you actively and consciously search for a given memory, say, where you left your credit card when you notice that it is missing, you do a serial search: "I started out at the bookstore, then I went to the record shop, then to the video rental. Since I was able to use my card in all these places, I must have left it at the last one – the video rental." This is the way conscious memory searches take place, because we can focus on only one thing at a time, so we have to do our thinking serially.

But most of the time we seem to remember things without any conscious search of our memory. You see an unusual fruit, try to think of its name, and eventually "guava" just pops into your mind without any awareness of a search. Classical descriptions of this sort of remembering are modeled on conscious memory searches, due to an implicit assumption that the two types of remembering are similar. According to the classical theory, what happens when you see this unusual fruit is that you unconsciously go through a list of all the fruits you know and check each one to see if it matches the one you are looking at. This would have to occur very fast, of course, since "guava" pops into your mind fairly quickly, but in theory there could be a series of very fast processes. In contrast, connectionists claim that the names of the fruits you know are all activated to some degree when you see the guava. The more common fruits are activated more quickly, so it takes a while for the sight of the guava in front of you to make the word "guava" come to the forefront, but eventually it does. This is called "parallel processing" because the names of all the fruits are activated at the same time, in parallel, and no serial search is needed. This is explained in detail in the central chapter of the book, Chapter 5, "What are connectionist networks?"

The other important difference is embodied in the word "distributed." In classical explanations every concept, including "guava," is stored in one particular place in the mind. "Guava" is associated with "tropical", "fruit," and its various other features, but these are separate concepts, as I explain in Chapter 4, "Theories and models of how the mind functions." Connectionist explanations, in contrast, claim that every concept is made up of many parts, so that "tropical" and "fruit" are actually part of the concept "guava" rather than just being associated with it. This idea is explained in detail in Chapter 5.

These different modes of explanation have had different degrees of success in explaining different sorts of mental functions. The degree of success is often judged by how well the mental process can be modeled by a computer program, as I explain in Chapter 4. Classical cognitive scientists have had great success in making computer models of the sort of things human beings generally find difficult, such as playing chess, solving mathematical problems, diagnosing rare diseases, or finding flaws in complicated machinery. These are processes that we often know how to describe, things we do with a good deal of conscious awareness of how we do them. Thus experts in these areas could tell classical computer scientists how they did these things, and the computer people could program their computers to do them just as well – or even better.

But when the computer scientists tried to get their digital computers to do things humans do easily – such as making sense of simple stories with obvious details left out, or looking at a two-dimensional picture and seeing individual three-dimensional objects in it, or carrying on a sensible conversation – they had a very difficult time. Part of the reason is that we don't really know how we do these things, since most of the work is done outside conscious awareness. Another part of the reason is that neither the brain nor the mind works like a digital computer.

The brain isn't programmed by anyone; it grows and develops the abilities it has. The brain does come equipped with its basic structure, but this structure is constantly being changed by our experience in the world. This is not like changing a program on a computer, where the "hardware" remains the same but the "software" changes. In the brain there is no distinction between "hardware" and "software." Every change is a change of hardware. Every time we learn something new, every time we see a new scene, hear a new sentence or tune, touch a new fabric, taste a new dessert, smell a new flower, the connections between the neurons in our brain undergo some changes, and this constitutes a change in what we know – a change in our mind.

Connectionists also use computer models, but they try to make them work in a parallel distributed way. This involves trying to make computers be more like the brain instead of assuming that the brain works like a computer. Chapter 8 describes some models of this sort.

Is the mind in any way like the Internet?

Instead of comparing the mind to a computer, it might be more useful to compare it to the Internet. I am not saying that the mind is actually very similar to the Internet, but only that there are some interesting aspects of the mind that can be appreciated more easily by comparing them with properties of the Internet.

One of the most interesting properties of the Internet is the fact that it is not hierarchical – it is not controlled at the top by anyone who tells everyone else what to do. Instead, every part of the Internet can communicate with every other part, sometimes directly and sometimes through one or a few intermediate stations.

Connectionist theories claim that the brain, and therefore the mind, works this way too. The various parts of the mind – its individual networks – are all interconnected in a vast web, each part of which can communicate with any other part. In most of the book I will be describing the way the individual networks operate and how certain ones communicate with other ones, but it is important to remember that the flow of information can go in all directions. There is no "master" operator – no "DOS" – that determines where the information should go or what should be done with it when it gets there.

The mind does indeed have what is often called an "executive function," which is responsible for such activities as planning ahead. But although this function is generally in charge, it too can be overridden. My "executive function" may have determined that I will read a chapter of my physics textbook this evening, but if my external senses tell the sensory parts of my brain that the waterpipes in my kitchen have just burst, or if my internal sensors tell the proprioceptive parts of my brain – the parts that monitor internal states – that I am intensely thirsty, or if the person I love tells me that we need some time together, I may temporarily abandon my carefully thought-out study plans. And in case you are thinking that there is some "super-executive" function which decides that the flooding emergency is more important than studying physics, just think of the situation where you stop studying in order to watch a silly comedy on television. Different parts of your mind are in conflict here, and which one wins is not decided by some "super-judge"; the winner is simply the one that manages to gain the upper hand at that particular moment.

Thus the "executive function" too is just one part of the mind among many, all of which talk to each other incessantly and jointly direct our actions. This is the most important way in which the mind is like the Internet.

What about consciousness?

The idea that there is no hierarchy in the mind is very similar to a notion expounded by Daniel Dennett in his book *Consciousness Explained*. In this book Dennett demonstrates that there is no one part of the brain in which consciousness is "located," because all the various parts work together to produce our conscious experiences. The connectionist theories I describe in the present book fit in well with Dennett's ideas. However, they center on specific areas of the brain, how they are organized internally, and how they interact with other specific areas. They have not yet reached the point where they can come to grips with the interplay of processes occurring in many parts of the brain that probably underlies consciousness. Therefore I will not be discussing the topic of consciousness in this book.

It is very likely, nevertheless, that someday connectionism will have something interesting to say about consciousness, since I do not think that it is some mysterious entity that cannot in principle be explained by science. I am convinced that consciousness is embodied primarily in brain function, just like all other aspects of the mind. In fact, there are a number of possible scenarios that I can envisage if connectionists ever do try to explain consciousness. They may end up agreeing with Dennett that the concept of consciousness is not a particularly useful one for understanding the mind. They may come to the conclusion that it is such a complicated function of all the networks acting together that it is too difficult to explain. They may find out that it is grounded in a different part of the brain from the ones that hold the knowledge networks to be described in the following chapters, and so operates in a different way. But whatever may turn out to be the case for consciousness, I find that it is possible to understand a great many things about how our minds function without considering it. I shall therefore put aside the problem of consciousness in the present book.

What about the emotions?

The mind is sometimes considered to include the emotions as well as the intellect. In this book I will be discussing only the intellect, not the emotions. I will use the word "mind" rather than "intellect" because "mind" is the word that people generally use when they are speaking about such human functions as classifying things into categories, speaking, understanding speech, solving problems and the like.

It is possible that connectionist theory may turn out to apply to the emotions as well as the intellect, but it is also possible that it may not apply to the emotions. The reason for this is rooted in the structure of the brain. The part of the brain that is organized into neural networks is the cortex – the deeply folded outer part – which is primarily responsible for perception, thinking and the planning of action. Therefore it makes sense to use the properties of these neural networks to try to understand the mental functions that involve mainly the cortex.

The emotions, in contrast, involve primarily the inner structures of the brain, which are not organized in quite the same way. Thus rather different models may be needed to describe how the emotions work. Perhaps network theory will someday be able to shed light on the human emotions as well as the human intellect, but it has not yet done so. The discussions in this book will therefore be limited to the intellectual functions of the mind.

What questions does this book try to answer?

Some of the questions about our mental processes that connectionist theories offer answers to, which I discuss in this book, are listed below.

Will scientists ever be able to read our minds? If we understand the brain completely, will we be able to know what other people are thinking by looking inside their brain? In what ways are the minds of different people similar, and in what ways are they different? These questions are discussed in Chapter 2.

How is the brain put together so as to serve as the basis for our mental functions? What are the physical connections that allow us to form mental associations between the different things we know about the world? What are the physical bases of the changes in our minds that constitute learning? These are some of the questions discussed in Chapter 3.

Chapter 4 discusses the following questions: What is the difference between a theory and a model? What models were used to explain the organization of our knowledge before connectionist theories were developed? How can computers help us understand the way our minds work even though the mind doesn't work like a computer?

The essence of connectionist theory is discussed in Chapter 5, with emphasis on the following questions: How do we put things into categories? How do we know, for example, that the animal we are looking at is a dog and not a cat? How are all the things we know about various animals

Cambridge University Press
0521561043 - The Brain-Shaped Mind: What the Brain Can Tell Us About the Mind - Naomi Goldblum
Excerpt
More information

connected with each other, so that we can instantly produce a long list of properties of any animal, or provide the name of the animal from a description of it?

Another central problem is how these connections form in the first place. How do children learn the difference between a cat and a dog? Why do some children call all four-legged animals "dog" or "horse" at first, although they never call a bird "horse"? How do they later learn to distinguish between different types of dogs, and even between individual dogs? These are some of the questions addressed in Chapter 6.

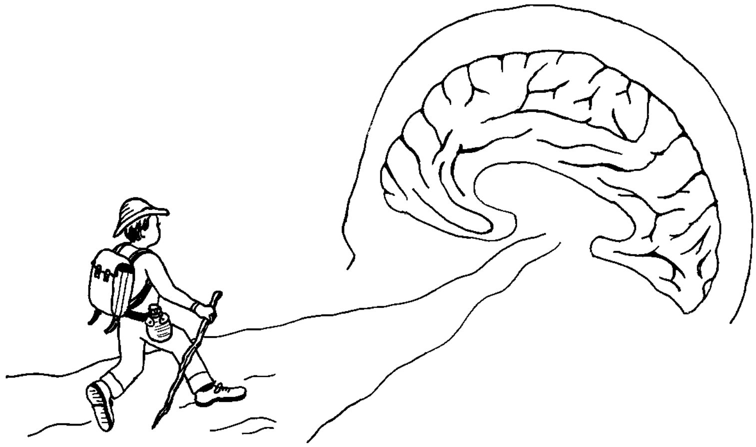
Of course, not all our mental associations are between objects and their names or physical properties. We also know that various things are related, such as mothers and fathers, or salt and pepper, or even properties of things, such as high and low, or sweet and sour. But how do we know in what way these things are related? How do we know that a wolfhound is a kind of dog rather than a kind of wolf? Why do we say "dog" and not "tiger" when we are asked for a word associated with "cat," even though cats and tigers are more closely related than cats and dogs? These questions are discussed in Chapter 7.

What sort of experimental evidence is there that the models discussed in this book actually describe what goes on in the human mind? Chapter 8 describes some of this evidence, including computer models of how children learn to talk and how they learn the past tenses of verbs.

Chapter 9 discusses the difference between things we remember for a long time and things we remember for only a little while. For example, how do we know that today we had eggs for breakfast, even though we usually have cereal, and why is it that next year we will still remember that we had cereal for breakfast most of the time this year, but we will not remember that we had eggs today? How do we remember before we go to the bank that we have to go to the bank this morning, yet instantly afterwards remember that we already performed this errand?

All the questions so far have involved the normal functioning of the mind in the intact brain. What happens if this functioning is disrupted by damage to the brain, such as that caused by strokes, head injuries or degenerative diseases like Alzheimer's? Chapter 10 discusses the different sorts of dysfunction associated with each of these causes and the prospects for regaining normal mental functioning in certain cases even if the brain damage itself cannot be repaired.

Is there any way we can use all this knowledge to improve our mental functioning in our daily lives? Can it help us study better, or teach others



more effectively, or solve our personal problems? These and other practical implications of connectionist theory are discussed in Chapter 11.

Neural network models are still fairly speculative at this point. Although the experimental evidence supports the theory that the brain and the mind work in this parallel distributed manner, the evidence that the mind works this way is rather less certain than the evidence with regard to the brain. As a result, there has been much criticism of connectionist theory, particularly by people who advocate other theories about mental functioning that are less closely tied to the workings of the brain. A sampling of these criticisms is presented in Chapter 12, together with replies that have been offered by researchers in connectionist theory.

Despite the fact that connectionism has its critics, it does reflect one of the most prevalent ways of thinking about the mind at the present time. In my discussion of what the models imply about how the mind works, I will therefore assume that they are true, and present their exciting implications by saying simply that this is what our minds do, rather than continually repeating that this is what the models say that our minds do.

In essence, although I describe a fairly simple neural network model in some detail, my discussion of this model is not meant to be an end in itself. I see it, rather, as a way of showing how our developing understanding of how the brain works can help us understand some aspects of how the mind works as well. Brain processes are only one of the forces that shape the mind – it is also shaped by input from the environment, both physical and social. Describing how these two types of forces interact requires another book entirely.