

# INTRODUCTION

## *Exploring Inner Space*

*The sciences have developed in an order the reverse of what might have been expected. What was most remote from ourselves was first brought under the domain of law, and then, gradually, what was nearer: first the heavens, next the earth, then animal and vegetable life, then the human body, and last of all (as yet very imperfectly) the human mind.*

—Bertrand Russell, 1935

## A BRAVE NEW WORLD

We are in the midst of a revolution. For centuries, science has made great strides in our understanding of the external observable world. Physics revealed the motion of the planets, chemistry discovered the fundamental elements of matter, and biology has told us how to understand and treat disease. But during much of this time, there were still many unanswered questions about something perhaps even more important to us—the human mind.

What makes mind so difficult to study is that, unlike the phenomena described above, it is not something we can easily observe, measure, or manipulate. In addition, the mind is the most complex entity in the known universe. To give you a sense of this complexity, consider the following. The human brain is estimated to contain 10 billion to 100 billion individual nerve cells or neurons. Each of these neurons can have as many as 10,000 connections to other neurons. This vast web is the basis of mind and gives rise to all the equally amazing and difficult-to-understand mental phenomena, such as perception, memory, and language.

The past several decades have seen the introduction of new technologies and methodologies for studying this intriguing organ. We have learned more about the mind in the past half-century than in all the time that came before that. This period of rapid discovery has coincided with an increase in the number of different disciplines—many of them

entirely new—that study mind. Since then, a coordinated effort among the practitioners of these disciplines has come to pass. This interdisciplinary approach has since become known as cognitive science. Unlike the science that came before, which was focused on the world of external, observable phenomena, or “outer space,” this new endeavor now turns its full attention to the discovery of our fascinating mental world, or “inner space.”

## WHAT IS COGNITIVE SCIENCE?

**Cognitive science** can be roughly summed up as the scientific interdisciplinary study of the mind. Its primary methodology is the scientific method—although, as we will see, many other methodologies also contribute. A hallmark of cognitive science is its interdisciplinary approach. It results from the efforts of researchers working in a wide array of fields. These include philosophy, psychology, linguistics, **artificial intelligence (AI)**, robotics, and neuroscience. Each field brings with it a unique set of tools and perspectives. One major goal of this book is to show that when it comes to studying something as complex as the mind, no single perspective is adequate. Instead, intercommunication and cooperation among the practitioners of these disciplines tell us much more.

The term *cognitive science* refers not so much to the sum of all these disciplines but to their intersection or converging work on specific problems. In this sense, cognitive science is not a unified field of study like each of the disciplines themselves but, rather, a collaborative effort among researchers working in the various fields. The glue that holds cognitive science together is the topic of mind and, for the most part, the use of scientific methods. In the concluding chapter, we talk more about the issue of how unified cognitive science really is.

To understand what cognitive science is all about, we need to know what its theoretical perspective on the mind is. This perspective centers on the idea of **computation**, which may alternatively be called information processing. Cognitive scientists view the mind as an information processor. Information processors must both represent and transform information. That is, a mind, according to this perspective, must incorporate some form of mental representation and processes that act on and manipulate that information. We will discuss these two ideas in greater detail later in this chapter.

Cognitive science is often credited with being influenced by the rise of the computer. Computers are, of course, information processors. Think for a minute about a personal computer. It performs a variety of information-processing tasks. Information gets into the computer via input devices, such as a keyboard or modem. That information can then be stored on the computer—for example, on a hard drive or other disk. The information can then be processed using software, such as a text editor. The results of this processing may next serve as output, either to a monitor or to a printer. In like fashion, we may think of people performing similar tasks. Information is “input” into our minds through

perception—what we see or hear. It is stored in our memories and processed in the form of thought. Our thoughts can then serve as the basis of “outputs,” such as language or physical behavior.

Of course, this analogy between the human mind and computers is highly abstract. The actual physical way in which data are stored on a computer bears little resemblance to human memory formation. But both systems are characterized by computation. In fact, it is not going too far to say that cognitive scientists view the mind as a machine or mechanism whose workings they are trying to understand.

## REPRESENTATION

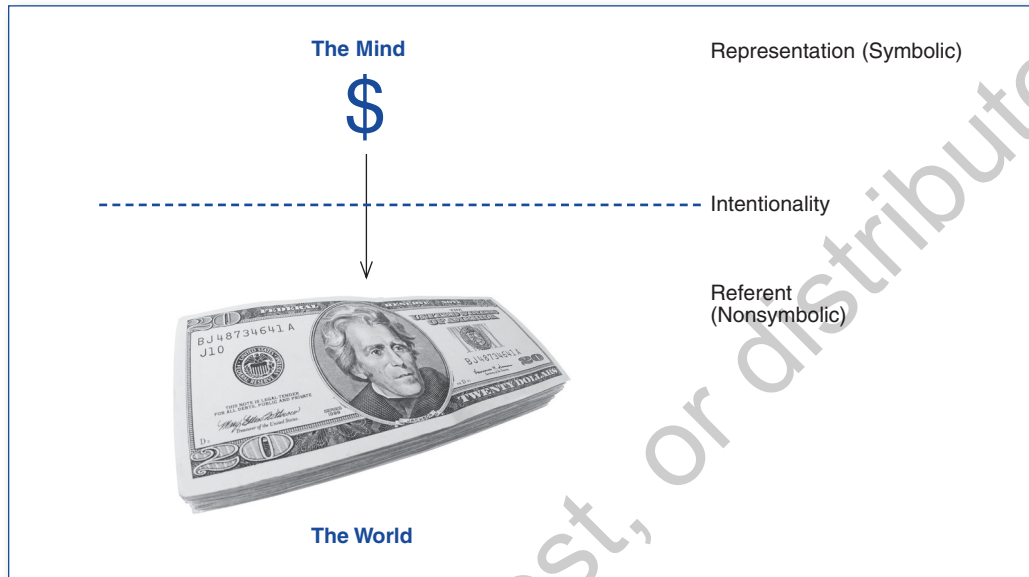
---

As mentioned before, representation is fundamental to cognitive science. But what is a representation? Briefly stated, a **representation** is something that stands for something else. Before listing the characteristics of a representation, it is helpful to describe briefly four categories of representation. A concept stands for a single entity or group of entities. Single words are good examples of concepts. The word *apple* denotes the concept of that particular type of fruit. Propositions are statements about the world and can be illustrated with sentences. The sentence “Mary has black hair” is a proposition that is itself made up of concepts. Rules are yet another form of representation that can specify the relationships between propositions. For example, the rule “If it is raining, I will bring my umbrella” makes the second proposition contingent on the first. There are also analog representations. An analogy helps us make comparisons between two similar situations. We will discuss all four of these representations in greater detail in the “Interdisciplinary Crossroads” section at the end of this chapter.

There are four crucial aspects of any representation (Hartshorne, Weiss, & Burks, 1931–1958). First, a “representation bearer” such as a human or a computer must realize a representation. Second, a representation must have content—meaning it stands for one or more objects. The thing or things in the external world that a representation stands for are called **referents**. A representation must also be “grounded.” That is, there must be some way in which the representation and its referent come to be related. Fourth, a representation must be interpretable by some interpreter, either the representation bearer or somebody else. These and other characteristics of representations are discussed next.

The fact that a representation stands for something else means it is **symbolic**. We are all familiar with symbols. We know, for instance, that the dollar symbol (\$) is used to stand for money. The symbol itself is not the money but, instead, is a surrogate that refers to its referent, which is actual money. In the case of mental representation, we say that there is some symbolic entity “in the head” that stands for real money. Figure 1.1 shows a visual representation of money. Mental representations can stand for many different types of things and are by no means limited to simple conceptual ideas such as “money.” Research

**Figure 1.1** Different aspects of the symbolic representation of money.



Source: PhotoObjects.net/Thinkstock.

suggests that there are more complex mental representations that can stand for rules—for example, knowing how to drive a car and analogies, which may enable us to solve certain problems or notice similarities (Thagard, 2000).

Human mental representations, especially linguistic ones, are said to be **semantic**, which is to say that they have meaning. Exactly what constitutes meaning and how a representation can come to be meaningful are topics of debate. According to one view, a representation's meaning is derived from the relationship between the representation and what it is about. The term that describes this relation is **intentionality**. Intentionality means “directed on an object.” Mental states and events are intentional. They refer to some actual thing or things in the world. If you think about your brother, then the thought of your brother is directed toward him—not toward your sister, a cloud, or some other object.

An important characteristic of intentionality has to do with the relationship between inputs and outputs to the world. An intentional representation must be triggered by its referent or things related to it. Consequently, activation of a representation (i.e., thinking about it) should cause behaviors or actions that are somehow related to the referent. For example, if your friend Sally told you about a cruise she took around the Caribbean last December, an image of a cruise ship would probably pop to mind. This might then cause

you to ask her if the food onboard was good. Sally's mention of the cruise was the stimulus input that activated the internal representation of the ship in your mind. Once it was activated, it caused the behavior of asking about the food. This relation between inputs and outputs is known as an **appropriate causal relation**.

Symbols can be assembled into what are called physical symbol systems, or more simply as formal systems. In a **formal system**, symbols are combined into expressions. These expressions can then be manipulated using processes. The result of a process can be a new expression. For example, in formal logic, symbols can be words like "all" or "mammals" and expressions could be statements like "all mammals nurse their young." The processes would be the rules of deduction that allow us to derive true concluding expressions from known expressions. In this instance, we could start off with the two expressions "all mammals nurse their young," "whales nurse their young," therefore "whales are mammals." More on this below where we discuss propositions and syllogisms.

According to the **physical symbol system hypothesis (PSSH)**, a formal system can allow for intelligence (Newell & Simon, 1976). Since we as humans appear to have representational and computational capacity being able to use things that stand for things, we seem to be intelligent. Beyond this, we could also infer that machines are intelligent, since they too have this capacity, although of course, this is debated.

There are several critiques that have been leveled against the PSSH (Nilsson, 2007). First, it is argued that the symbols computers use have no meaning or semantic quality. To be meaningful, symbols have to be connected to the environment in some way. People and perhaps animals seem to have meaning because we have bodies and can perceive things and act on them. This "grounds" the symbols and imbues them with semantic quality. Since machines are not embodied, they cannot acquire meaning. This issue is known as **the symbol grounding problem** and is in effect a re-expression of the concept of intentionality.

A counterargument to this is that computer systems do have the capability to **designate**. An expression can designate an object if it can affect the object itself or behave in ways that depend on the object. One could argue that a robot capable of perceiving an object like a coffee mug and able to pick it up could develop semantics toward it in the same way that a person might. Ergo, the robot could be intelligent. Also, there are examples of AI programs like expert systems that have no sensor or effector capability yet are able to produce intelligent and useful results. Some expert systems, like MYCIN, are able to more accurately diagnose certain medical disorders than are members of the Stanford University medical school (Cawsey, 1998). They can do this despite being able to see or act in the world.

## **Types of Representation**

The history of research in cognition suggests that there are numerous forms of mental representation. Paul Thagard (2000), in *Mind: Introduction to Cognitive Science*, proposes

four: concepts, propositions, rules, and analogies. Although some of these have already been alluded to and are described elsewhere in the book, they are central to many ideas in cognitive science. It is, therefore, useful to sketch out some of their major characteristics here.

A **concept** is perhaps the most basic form of mental representation. A concept is an idea that represents things we have grouped together. The concept “chair” does not refer to a specific chair, such as the one you are sitting in now, but it is more general than that. It refers to all possible chairs no matter what their colors, sizes, and shapes. Concepts need not refer to concrete items. They can stand for abstract ideas—for example, “justice” or “love.” Concepts can be related to one another in complex ways. They can be related in a hierarchical fashion, where a concept at one level of organization stands for all members of the class just below it. “Golden retrievers” belongs to the category of “dogs,” which in turn belongs to the category of “animals.” We discuss a hierarchical model of concept representation in the network approach chapter. The question of whether concepts are innate or learned is discussed in the philosophical approach chapter.

A **proposition** is a statement or assertion typically posed in the form of a simple sentence. An essential feature of a proposition is that it can be proved true or false. For instance, the statement “The moon is made out of cheese” is grammatically correct and may represent a belief that some people hold, but it is a false statement. We can apply the rules of formal logic to propositions to determine the validity of those propositions. One logical inference is called a syllogism. A syllogism consists of three propositions. The first two are premises and the last is a conclusion. Take the following syllogism:

All men like football.

John is a man.

John likes football.

Obviously, the conclusion can be wrong if either of the two premises is wrong. If it is not true that all men like football, then it might not be true that John likes football, even if he is a man. If John is not a man, then he may or may not like football, assuming all men like it. Syllogistic reasoning of this sort is the same as deductive reasoning mentioned earlier.

You may have noticed that propositions are representations that incorporate concepts. The proposition “All men like football” incorporates the concepts “men” and “football.” Propositions are more sophisticated representations than concepts because they express relationships—sometimes very complex ones—between concepts. The rules of logic are best thought of as computational processes that can be applied to propositions to determine their validity. However, logical relations between propositions may themselves be considered a separate type of representation. The evolutionary approach chapter provides an interesting account of why logical reasoning, which is difficult for many people, is easier under certain circumstances.

Logic is not the only system for performing operations on propositions. Rules do this as well. A **production rule** is a conditional statement of the following form: “If  $x$ , then  $y$ ,” where  $x$  and  $y$  are propositions. The “if” part of the rule is called the condition; the “then” part is called the action. If the proposition that is contained in the condition ( $x$ ) is true, then the action that is specified by the second proposition ( $y$ ) should be carried out, according to the rule. The following rules help us drive our cars:

If the light is red, then step on the brakes.

If the light is green, then step on the accelerator.

Notice that, in the first rule, the two propositions are “the light is red” and “step on the brakes.” We can also form more complex rules by linking propositions with “and” and “or” statements:

If the light is red or the light is yellow, then step on the brakes.

If the light is green and nobody is in the crosswalk, then step on the accelerator.

The *or* that links the two propositions in the first part of the rule specifies that if either proposition is true, the action should be carried out. If an *and* links these two propositions, the rule specifies that both must be true before the action can occur.

Rules bring up the question of what knowledge really is. We usually think of knowledge as factual. Indeed, a proposition such as “Candy is sweet,” if validated, does provide factual information. The proposition is then an example of declarative knowledge. **Declarative knowledge** is used to represent facts. It tells us what is and is demonstrated by verbal communication. **Procedural knowledge**, on the other hand, represents skill. It tells us how to do something and is demonstrated by action. If we say that World War II was fought during the period 1939 to 1945, we have demonstrated a fact learned in history class. If we ski down a snowy mountain slope in the winter, we have demonstrated that we possess a specific skill. It is, therefore, very important that information-processing systems have some way of representing actions if they are to help an organism or machine perform those actions. Rules are one way of representing procedural knowledge. We discuss two cognitive, rule-based systems—the atomic components of thought and SOAR (state, operator, and result) models—in the cognitive approach chapters.

Another specific type of mental representation is the **analogy**—although, as is pointed out below, the analogy can also be classified as a form of reasoning. Thinking analogically involves applying one’s familiarity with an old situation to a new situation. Suppose you had never ridden on a train before but had taken buses numerous times. You could use your understanding of bus riding to figure out how to take a ride on a train. Applying knowledge that you already possess and that is relevant to both scenarios would enable



you to accomplish this. Based on prior experience, you would already know that you have to first determine the schedule, perhaps decide between express and local service, purchase a ticket, wait in line, board, stow your luggage, find a seat, and so on.

Analogies are a useful form of representation because they allow us to generalize our learning. Not every situation in life is entirely new. We can apply what we already have learned to similar situations without having to figure out everything all over again. Several models of analogical reasoning have been proposed (Forbus, Gentner, & Law, 1995; Holyoak & Thagard, 1995).

## COMPUTATION

---

As mentioned earlier, representations are only the first key component of the cognitive science view of mental processes. Representations by themselves are of little use unless something can be done with them. Having the concept of money doesn't do much for us unless we know how to calculate a tip or can give back the correct amount of change to someone. In the cognitive science view, the mind performs computations on representations. It is, therefore, important to understand how and why these mental mechanisms operate.

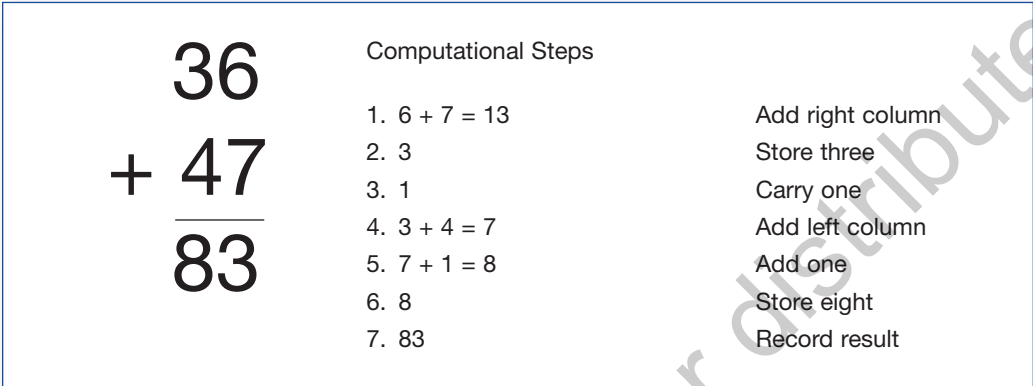
What sorts of mental operations does the mind perform? If we wanted to get details about it, the list would be endless. Take the example of mathematical ability. If there were a separate mental operation for each step in a mathematical process, we could say the mind adds, subtracts, divides, and so on. Likewise, with language, we could say that there are separate mental operations for making a noun plural, putting a verb into past tense, and so on. It is better, then, to think of mental operations as falling into broad categories. These categories can be defined by the type of operation that is performed or by the type of information acted on. An incomplete list of these operations would include sensation, perception, attention, memory, language, mathematical reasoning, logical reasoning, decision making, and problem solving. Many of these categories may incorporate virtually identical or similar subprocesses—for example, scanning, matching, sorting, and retrieving. Figure 1.2 shows the kinds of mental processes that may be involved in solving a simple addition problem.

### *The Tri-Level Hypothesis*

Any given information process can be described at several different levels. According to the tri-level hypothesis, mental or artificial information-processing events can be evaluated on at least three different levels (Marr, 1982). The highest or most abstract level of analysis is the **computational level**. At this level, one is concerned with two tasks. The first is a clear specification of what the problem is. Taking the problem as it may originally have been posed, in a vague manner perhaps, and breaking it down into its main constituents or parts



**Figure 1.2** Some of the computational steps involved in solving an addition problem.



can bring about this clarity. It means describing the problem in a precise way such that the problem can be investigated using formal methods. It is like asking, What exactly is this problem? What does this problem entail? The second task one encounters at the computational level concerns the purpose or reason for the process. The second task consists of asking, Why is this process here in the first place? Inherent in this analysis is adaptiveness—the idea that human mental processes are learned or have evolved to enable the human organism to solve a problem it faces. This is the primary explanatory perspective used in the evolutionary approach. We describe a number of cognitive processes and the putative reasons for their evolution in the evolution chapter.

Stepping down one level of abstraction, we can next inquire about the way in which an information process is carried out. To do this, we need an **algorithm**, a formal **procedure** or system that acts on informational representations. It is important to note that algorithms can be carried out regardless of a representation’s meaning; algorithms act on the form, not the meaning, of the symbols they transform. One way to think of algorithms is that they are “actions” used to manipulate and change representations. Algorithms are formal, meaning they are well defined. We know exactly what occurs at each step of an algorithm and how a particular step changes the information being acted on. A mathematical formula is a good example of an algorithm. A formula specifies how the data are to be transformed, what the steps are, and what the order of steps is. This type of description is put together at the **algorithmic level**, sometimes also called the programming level. It is equivalent to asking, What information-processing steps are being used to solve the problem? To draw an analogy with computers, the algorithmic level is like software because software contains instructions for the processing of data.

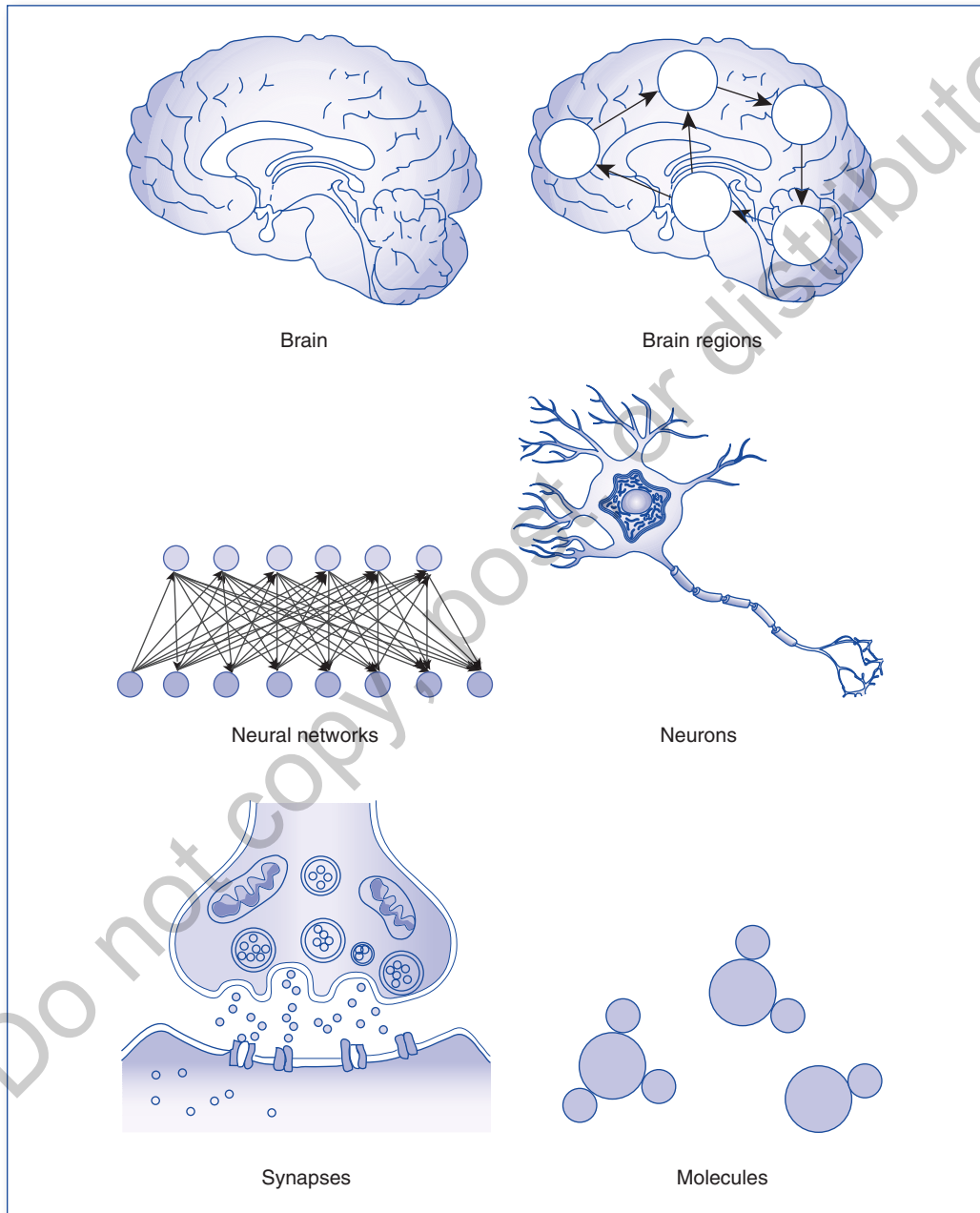
The most specific and concrete type of description is formulated at the **implementational level**. Here we ask, What is the information processor made of? What types of physical or material changes underlie changes in the processing of the information? This level is sometimes referred to as the hardware level, since in computer parlance, the hardware is the physical “stuff” the computer is made of. This would include its various parts—a monitor, hard drive, keyboard, and mouse. At a smaller scale, computer hardware consists of circuits and even the flow of electrons through the circuits. The hardware in human or animal cognition is the brain and, on a smaller scale, the neurons and activities of those neurons.

At this point, one might wonder, Why do we even need an algorithmic or formal level of analysis? Why not just map the physical processes at the implementational level onto a computational description of the problem or, alternatively, onto the behaviors or actions of the organism or device? This seems simpler, and we need not resort to the idea of information and representation. The reason is that the algorithmic level tells us how a particular system performs a computation. Not all computational systems solve a problem in the same way. Computers and humans can both perform addition but do so in drastically different fashions. This is true at the implementational level, obviously, but understanding the difference formally tells us much about alternative problem-solving approaches. It also gives us insights into how these systems might compute solutions to other novel problems that we might not understand.

This partitioning of the analysis of information-processing events into three levels has been criticized as being fundamentally simplistic since each level can, in turn, be further subdivided into levels (Churchland, Koch, & Sejnowski, 1990). Figure 1.3 depicts one possible organization of the many structural levels of analysis in the nervous system. Starting at the top, we might consider the brain as one organizational unit, brain regions as corresponding to another organizational unit one step down in spatial scale, then neural networks, individual neurons, and so on. Similarly, we could divide algorithmic steps into different substeps and problems into subproblems. To compound all this, it is not entirely clear how to map one level of analysis onto another. We may be able to specify clearly how an algorithm executes but be at a loss to say exactly where or how this is achieved with respect to the nervous system.

### ***Differing Views of Representation and Computation***

Before finishing our discussion of computation, it is important to differentiate between several different conceptions of what it is. So far, we have been talking about computation as being based on the formal systems notion. In this view, a computer is a **formal symbol manipulator**. Let's break this definition down into its component parts. A system is formal if it is syntactic or rule governed. In general, we use the word **syntax** to refer to the set of rules that govern any symbol system. The rules of language and mathematics are formal

**Figure 1.3** Structural levels of analysis in the nervous system.

systems because they specify which types of allowable changes can be made to symbols. Formal systems also operate on representations independent of the content of those representations. In other words, a process can be applied to a symbol regardless of its meaning or semantic content. A symbol, as we have already indicated, is a form of representation and can assume a wide variety of forms. Manipulation here implies that computation is an active, embodied process that takes place over time. That is, manipulations are actions, they occur physically in some type of computing device, and they take some time to occur (i.e., they don't happen instantaneously).

But this is not the only conception of what computation is. The connectionist or network approach to computation differs from the classical formal systems approach of cognitive science in several ways. In the classical view, knowledge is represented locally—in the form of symbols. In the connectionist view, knowledge is represented as a pattern of activation or weights that is distributed throughout a network and so is more global than a single symbol. Processing style is also different in this approach. The classical view has processing occurring in discrete serial stages, whereas in connectionism, processing occurs in parallel through the simultaneous activation of many nodes or elements in the network. Some cognitive scientists downplay these differences, arguing that information processing occurs in both systems and that the tri-level hypothesis can be applied equally to both (Dawson, 1998). We further compare and contrast the classical and connectionist views at the beginning of the network approach chapter.

What happens to representations once they get established? In some versions of the symbolic and connectionist approaches, representations remain forever fixed and unchanging. For example, when you learned about the concept of a car, a particular symbol standing for “car” was formed. The physical form of this symbol could then be matched against perceptual inputs to allow you to recognize a car when you see one or used in thought processes to enable you to reason about cars when you need to. Such processes can be simplified if the symbol stays the same over time. This, in fact, is exactly how a computer deals with representation. Each of the numbers or letters of the alphabet has a unique and unchanging ASCII (American Standard Code for Information Interchange) code. Similarly, in artificial neural network simulations of mind, the networks are flexible and plastic during learning, but once learning is complete, they must remain the same. If not, the representations will be rewritten when new things are learned.

A more realistic view of representation comes from the **dynamical perspective** in psychology (Friedenberg, 2009). According to this view, the mind is constantly changing as it adapts to new information. A representation formed when we first learn a concept is altered each time we think about that concept or experience information that is in some way related to it. For example, let's say a child sees a car for the first time and it is red. The child may then think that all cars are red. The next time he or she sees a car, it is blue. The child's concept would then be changed to accommodate the new color. Even after we are very familiar with a concept, the context in which we experience it is always different. You may hear a friend

discussing how fast one type of car is, or you may find yourself driving a car you have never driven before. In each of these cases, the internal representation is modified to account for the new experience. So it is unlikely that biological representations of the sort we see in humans will ever stay the same. We talk more about the dynamical systems approach and how it can unify differing perspectives on representation and computation in the concluding chapter.

Computation as conceptualized in cognitive science is believed to be neural and multi-cellular. In other words, it is the result of the flow of an electrical signal across neuron cell membranes. Bray (2011), however, argues for an alternate form of computation, one that occurs within a cell rather than between them. He believes that single-celled organisms, such as amoebas, need to compute in order to carry out their complex behavioral repertoire of hunting, fleeing, and responding to environmental cues. To do this, they have evolved a molecular form of representation, where protein molecules can stand for referents and chemical processes that act on these molecules are the equivalent of symbolic transformations. Bray believes that individual cells even contain molecular circuits that are capable of performing the logical operations that are the basis of complex information processing. This new paradigm is not without its critics. For one, the machinery of molecular computation may not be capable of duplicating the complex sorts of computing we see in brains. Also, computation of this sort may be possible only in motile single-celled organisms that need to move and so face greater adaptive challenges. If this were the case, then neurons in brains may not be able to perform such feats.

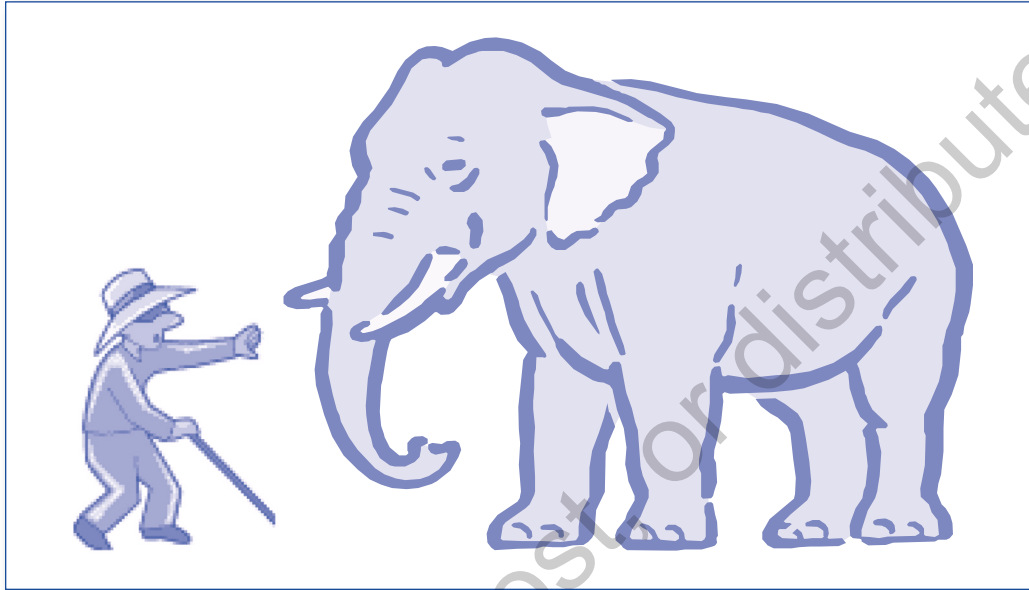
## THE INTERDISCIPLINARY PERSPECTIVE

---

There is an old fable about five blind men who stumble on an elephant (see Figure 1.4). Not knowing what it is, they start to feel the animal. One man feels only the elephant's tusk and thinks he is feeling a giant carrot. A second man feels the ears and believes that the object is a big fan. The third feels the trunk and proclaims that it is a pestle, while a fourth touching only the leg believes that it is a mortar. The fifth man, touching the tail, has yet another opinion: He believes it to be a rope. Obviously, all five men are wrong in their conclusions because each has examined only one aspect of the elephant. If the five men had gotten together and shared their findings, they may easily have pieced together what kind of creature it was. This story serves as a nice metaphor for cognitive science. We can think of the elephant as the mind and the blind men as researchers in different disciplines in cognitive science. Each individual discipline may make great strides in understanding its particular subject matter but, if it cannot compare its results to those of other related disciplines, may miss out on understanding the real nature of what is being investigated.

The key, then, to figuring out something as mysterious and complex as mind is communication and cooperation among disciplines. This is what is meant when one talks about cognitive science—not the sum of each of the disciplines or approaches but, rather, their

**Figure 1.4** If you were the blind man, would you know it is an elephant?



interaction. Recent years have seen an increase in this cooperation. A number of major universities have established interdisciplinary cognitive science centers, where researchers in such diverse areas as philosophy, neuroscience, and cognitive psychology are encouraged to work together on common problems. Each area can then contribute its unique strength to the phenomenon under study. The philosophers can pose broad questions and hypotheses, the neuroscientists can measure physiological performance and brain activity, while the cognitive psychologists can design and carry out experiments. The consequent exchange of results and ideas then leads to fruitful synergies between these disciplines, accelerating progress with respect to finding solutions to the problem and yielding insights into other research questions.

We have alluded to some of the different approaches in cognitive science. Because this book is about explaining each approach and its major theoretical contributions, it is worth describing each now in terms of its perspective, history, and methodology. In the following sections, we will also provide a brief preview of the issues addressed by each approach.

### ***The Philosophical Approach***

Philosophy is the oldest of all the disciplines in cognitive science. It traces its roots back to the ancient Greeks. Philosophers have been active throughout much of

recorded human history, attempting to formulate and answer basic questions about the universe. This approach is free to study virtually any sort of important question on virtually any subject, ranging from the nature of existence to the acquisition of knowledge to politics, ethics, and beauty. Philosophers of mind narrow their focus to specific problems concerning the nature and characteristics of mind. They might ask questions such as, What is mind? How do we come to know things? How is mental knowledge organized?

The primary method of philosophical inquiry is reasoning, both deductive and inductive. **Deductive reasoning** involves the application of the rules of logic to statements about the world. Given an initial set of statements assumed to be true, philosophers can derive other statements that logically must be correct. For example, if the statement “College students study 3 hours every night” is true and the statement “Mary is a college student” is true, we can conclude that “Mary will study 3 hours every night.” Philosophers also engage in **inductive reasoning**. They make observations about specific instances in the world, notice commonalities among them, and draw conclusions. An example of inductive reasoning would be the following: “Whiskers the cat has four legs,” “Scruffy the cat has four legs,” and, therefore, “All cats have four legs.” However, philosophers do not use a systematic form of induction known as the scientific method. That is employed within the other cognitive science disciplines.

In Chapter 2, we summarize several of the fundamental issues facing the philosophers of mind. With respect to the mind–body problem, philosophers wrangle over what exactly a mind is. Is the mind something physical like a rock or a chair, or is it nonphysical? Can minds exist only in brains, or can they emerge from the operation of other complex entities, such as computers? The knowledge acquisition problem deals with how we come to know things. Is knowledge a product of one’s genetic endowment, or does it arise through one’s interaction with the environment? How much does each of these factors contribute to any given mental ability? We also look into one of the most fascinating and enigmatic mysteries of mind—that of consciousness. In this case, we can ask, What is consciousness? Are we really conscious at all?

### Interdisciplinary Crossroads: Science and Philosophy

Philosophers have often been accused of doing “armchair” philosophy, meaning they do nothing but sit back in comfortable chairs and reflect about the world. The critique is that they ignore empirical evidence, information based on observation of the world, usually under controlled circumstances as is the case in the sciences. This may be changing, however, with the development of a new trend known as experimental philosophy.



**Experimental philosophy** utilizes empirical methods, typically in the form of surveys that assess people's understanding of constructed scenarios, to help answer philosophical questions. Followers of the movement call themselves "X-Philes" and have even adopted a burning armchair and accompanying song as their motto, a video of which is viewable on YouTube.

One of the more prominent X-Philers, Joshua Knobe (2003), presented participants with a hypothetical situation. In the situation, a vice president of a company approaches the chairman of the board and asks about starting a new program. The program will increase profits, but it will also harm the environment. In one version of the survey, the chairman responds that he or she doesn't care at all about harming the environment but just wants to make as much profit as possible and so wants to begin the new program.

When participants were asked whether the chairman harmed the environment intentionally, 82% responded "yes." In a different version of the scenario, everything was identical except that "harming" was now replaced with "helping." The program had the side effect of helping the environment but the chairman again stated that he or she didn't care at all about helping the environment and wanted to go ahead with it. This time when asked if the chairman intentionally helped the environment, only 23% responded "yes."

Logically, there should be no difference in perceived intentionality between both conditions. In each case, there was a desire to make profit and a failure to care one way or the other about the environment. But clearly, the way we see intentionality involves moral evaluations, the way people judge what is good or bad. This conclusion, what is now called the "Knobe effect" could not necessarily have been determined *a priori* using pure reasoning. The term *a priori* relates to what can be known through an understanding of how certain things work rather than by observation.

Experimental philosophy has addressed a wide variety of issues. These include cultural differences in how people know or merely believe (Weinberg, Nichols, & Stich, 2001), whether a person can be morally responsible if his or her actions are determined (Nichols & Knobe, 2007), the ways in which people understand consciousness (Huebner, 2010) and even, how philosophical intuitions are related to personality traits (Feltz & Cokely, 2009).

Of course, this new field has been critiqued. Much of experimental philosophy has looked at intuitions, and Williamson (2008) argues that philosophical evidence should not just rely on people's intuitions. Another obvious issue is that experimental philosophy is not philosophy at all, but science. In other words, if it is adopting the use of survey methodology, statistics, and so on, then it really becomes psychology or a branch of the social sciences rather than philosophy. Researchers have noted that some of the science is sloppy, using low and unrepresentative sample sizes.

## ***The Psychological Approach***

Compared with philosophy, psychology is a relatively young discipline. It can be considered old though, particularly when compared with some of the more recent newcomers to the cognitive science scene—for example, AI and robotics. Psychology as a science arose in the late 19th century and was the first discipline in which the scientific method was applied exclusively to the study of mental phenomena. Early psychologists established experimental laboratories that would enable them to catalog mental ideas and investigate various mental capacities, such as vision and memory. Psychologists apply the scientific method to both mind and behavior. That is, they attempt to understand not just internal mental phenomena, such as thoughts, but also the external behaviors that these internal phenomena can give rise to.

The scientific method is a way of getting hold of valid knowledge about the world. One starts with a hypothesis or idea about how the world works and then designs an experiment to see if the hypothesis has validity. In an experiment, one essentially makes observations under a set of controlled conditions. The resulting data then either support or fail to support the hypothesis. This procedure, employed within psychology, and cognitive science in general, is described more fully at the start of Chapter 3.

The field of psychology is broad and encompasses many subdisciplines, each one having its unique theoretical orientations. Each discipline has a different take on what mind is. The earliest psychologists—that is, the voluntarists and structuralists—viewed the mind as a kind of test tube in which chemical reactions between mental elements took place. In contrast, functionalism viewed mind not according to its constituent parts but, rather, according to what its operations were—what it could do. The Gestaltists again went back to a vision of mind as composed of parts but emphasized that it was the combination and interaction of the parts, which gave rise to new wholes, that was important. Psychoanalytic psychology conceives of mind as a collection of differing and competing minds, while behaviorism sees it as something that maps stimuli onto behaviors.

## ***The Cognitive Approach***

Starting in the 1960s, a new form of psychology arrived on the scene. Known as cognitive psychology, it came into being, in part, as a backlash against the behaviorist movement and its profound emphasis on behavior. Cognitive psychologists placed renewed emphasis on the study of internal mental operations. They adopted the computer as a metaphor for mind and described mental functioning in terms of representation and computation. They believed that the mind, like a computer, could be understood in terms of information processing.

The cognitive approach was also better able to explain phenomena such as language acquisition for which behaviorists did not have good accounts. At around the same time, new technologies that allowed better measurement of mental activity were being developed. This promoted a movement away from the behaviorist's emphasis on external observable behaviors and toward the cognitive scientist's emphasis on internal functions, as these could, for the first time, be observed with reasonable precision.

Inherent in the cognitive approach is the idea of modularity. Modules are functionally independent units that receive inputs from other modules, perform a specific processing task, and pass the results of their computation onto yet additional modules. The influence of the modular approach can be seen in the use of process models or flow diagrams. These depict a given mental activity via the use of boxes and arrows, where boxes depict modules and arrows the flow of information among them. The techniques used in this approach are the experimental method and computational modeling. Computational modeling involves carrying out a formal (typically software-based) implementation of a proposed cognitive process. Researchers can run the modeling process so as to simulate how the process might operate in a human mind. They can then alter various parameters of the model or change its structure in an effort to achieve results as close as possible to those obtained in human experiments. This use of modeling and comparison with experimental data is a unique characteristic of cognitive psychology and is also used in the AI and artificial network approaches.

Cognitive psychologists have studied a wide variety of mental processes, including pattern recognition, attention, memory, imagery, and problem solving. Theoretical accounts and processing models for each of these are given in Chapters 4 and 5. Language is within the purview of cognitive psychology, but because the approach to language is also multidisciplinary, we describe it separately in Chapter 9.

## ***The Neuroscience Approach***

Brain anatomy and physiology have been studied for quite some time. It is only recently, however, that we have seen tremendous advances in our understanding of the brain, especially in terms of how neuronal processes can account for cognitive phenomena. The study of the brain and endocrine system and how these account for mental states and behavior is called neuroscience. The attempt to explain cognitive processes in terms of underlying brain mechanisms is known as cognitive neuroscience.

Neuroscience, first and foremost, provides a description of mental events at the implementational level. It attempts to describe the biological "hardware" on which mental "software" supposedly runs. However, as discussed above, there are many levels of scale when it comes to describing the brain, and it is not always clear which level provides the best explanation for any given cognitive process. Neuroscientists, however, investigate at

each of these levels. They study the cell biology of individual neurons and of neuron-to-neuron synaptic transmission, the patterns of activity in local cell populations, and the interrelations of larger brain areas.

A reason for many of the recent developments in neuroscience is, again, the development of new technologies. Neuroscientists employ a wide variety of methods to measure the performance of the brain at work. These include positron emission tomography, computerized axial tomography, and **magnetic resonance imaging**. Studies that use these procedures have participants perform a cognitive task, and the brain activity that is concurrent with the performance of the task is recorded. For example, a participant may be asked to form a visual image of a word that appears on a computer screen. The researchers can then determine which parts of the brain became active during imagery and in what order. Neuroscientists use other techniques as well. They study brain-damaged patients and the effects of lesions in laboratory animals, and they use single- and multiple-cell recording techniques.

## **The Network Approach**

The network approach is at least partially derived from neuroscience. In this perspective, mind is seen as a collection of computing units. These units are connected to one another and mutually influence one another's activity via their connections, although each of the units is believed to perform a relatively simple computation—for example, a neuron can either “fire” by initiating an action potential or not “fire” and fail to initiate an action potential. In these networks, the connectivity of the units can give rise to representational and computational complexity.

Chapter 7, which outlines the network approach, has two parts. The first involves the construction of artificial neural networks. Most artificial neural networks are computer software simulations that have been designed to mimic the way actual brain networks operate. They attempt to simulate the functioning of neural cell populations. Artificial neural networks that can perform arithmetic, learn concepts, and read out loud now exist. A wide variety of network architectures has developed over the past 30 years.

The second part of the network chapter is more theoretical and focuses on knowledge representation—on how meaningful information may be mentally coded and processed. In semantic networks, nodes standing for concepts are connected to one another in such a way that activation of one node causes activation of other related nodes. Semantic networks have been constructed to explain how conceptual information in memory is organized and recalled. They are often used to predict and explain data obtained from experiments with human participants in cognitive psychology.

In the chapter on networks, we will finish off with a discussion of network science. This is a new interdisciplinary field, much like cognitive science. However, in this field, researchers focus on the structure and function of networks. The term *networks* is meant

in a very broad sense here to include not just artificial or natural neural networks but telephone and wireless networks, electrical power networks, and social networks. Surprisingly, we will see that there are commonalities among these different networks and that they share some organizational and operational features. We will examine these features and apply them particularly to the brain and psychology.

### ***The Evolutionary Approach***

The theory of natural selection proposed by Charles Darwin in 1859 revolutionized our way of thinking about biology. Natural selection holds that adaptive features enable the animals that possess them to survive and pass these features on to future generations. The environment, in this view, is seen as selecting from among a variety of traits those that serve a functional purpose.

The evolutionary approach can be considered in a quite general way and used to explain phenomena outside biology. The field of evolutionary psychology applies selection theory to account for human mental processes. It attempts to elucidate the selection forces that acted on our ancestors and how those forces gave rise to the cognitive structures we now possess. Evolutionary psychologists also adopt a modular approach to mind. In this case, the modules correspond with “favored” cognitive capacities that were used by ancestors successful at solving certain problems. Evolutionary theories have been proposed to account for experimental results across a wide range of capacities, from categorization to memory to logical and probabilistic reasoning, language, and cognitive differences between the sexes. They also have been proposed to account for how we reason about money—a new field of study known as behavioral economics.

Also in this chapter, we examine comparative cognition. This is the study of animal intelligence. We look at the cognitive capacities of a number of different species and discuss some of the problems that arise in comparing animals with one another and with humans.

### ***The Linguistic Approach***

Linguistics is an area that focuses exclusively on the domain of language. It is concerned with all questions concerning language ability, such as What is language? How do we acquire language? What parts of the brain underlie language use? As we have seen, language is a topic studied within other disciplines—for example, cognitive psychology and neuroscience. Because so many different researchers in different disciplines have taken on the problem of language, we consider it here as a separate discipline, united more by topic than by perspective or methodology.

Part of the difficulty in studying language is the fact that language itself is so complex. Much research has been devoted to understanding its nature. This work looks at the properties all languages share, the elements of language, and how those elements are used during communication. Other foci of linguistic investigation center on primate language use, language acquisition, deficits in language acquisition caused by early sensory deprivation or brain damage, the relationship between language and thought, and the development of speech recognition systems.

Linguistics, perhaps more than any other perspective discussed here, adopts a very eclectic methodological approach. Language researchers employ experiments and computer models, study brain-damaged patients, track how language ability changes during development, and compare diverse languages.

### ***The Emotion Approach***

As you may have surmised, humans don't just think—we also feel. Our conscious experience consists of emotions, such as happiness, sadness, and anger. Recent work in cognitive psychology and other fields has produced a wealth of data on emotions and how they influence thoughts. In Chapter 10, we start out by examining what emotions are and how they differ from feelings and moods. We examine several different theories of emotion and describe how they influence perception, attention, memory, and decision making. Following this, we look at the neuroscience underlying emotions and the role that evolutionary forces played in their formation. AI investigators have formulated models of how computers can “compute” and display emotional behavior. There are even robots capable of interacting with people in an emotionally meaningful way.

### ***The Social Approach***

Cognition happens inside individuals, but those individuals are strongly influenced by their social environment. The field of social cognition explores how people make sense of both themselves and others. We will see that thinking about people often differs from thinking about objects and that different parts of our brains are used when thinking socially. Neuroscience has revealed that in laboratory animals there are specialized cells, called mirror neurons, that are active both when an animal performs some action and when it watches another animal or person perform that same action. Later in Chapter 11, we introduce the concept of a theory of mind. This is an ability to understand and appreciate other people's states of mind. This capacity may be lacking in people suffering from autism. We conclude by summarizing work on specific social cognitive phenomena: attitudes, impressions, attributions, stereotypes, and prejudice.

## ***The Artificial Intelligence Approach***

Researchers have been building devices that attempt to mimic human and animal function for many centuries. But it is only in the past few decades that computer scientists have seriously attempted to build devices that mimic complex thought processes. This area is now known as artificial intelligence. Researchers in AI are concerned with getting computers to perform tasks that have heretofore required human intelligence. As such, they construct programs to do the sorts of things that require complex reasoning on our part. AI programs have been developed that can diagnose medical disorders, use language, and play chess.

AI also gives us insights into the function of human mental operations. Designing a computer program that can visually recognize an object often proves useful in understanding how we may perform the same task ourselves. An even more exciting outcome of AI research is that someday we may be able to create an artificial person who will possess all or many of the features that we consider uniquely human, such as consciousness, the ability to make decisions, and so on (Friedenberg, 2008).

The methods employed in the AI perspective include the development and testing of computer algorithms, their comparison with empirical data or performance standards, and their subsequent modification. Researchers have employed a wide range of approaches. An early attempt at getting computers to reason involved the application of logical rules to propositional statements. Later, other techniques were used. Chapters 12 and 13 give detailed descriptions of these techniques.

## ***The Robotics Approach***

Finally, we consider robotics. Robotics may be considered a familial relation to AI and has appeared on the scene as a formal discipline only recently. Whereas AI workers build devices that “think,” robotics researchers build machines that must also “act.” Investigators in this field build autonomous or semiautonomous mechanical devices that have been designed to perform a physical task in a real-world environment. Examples of things that robots can do presently include navigating around a cluttered room, welding or manipulating parts on an assembly line, and defusing bombs.

The robotics approach has much to contribute to cognitive science and to theories of mind. Robots, like people and animals, must demonstrate successful goal-oriented behaviors under complex, changing, and uncertain environmental conditions. Robotics, therefore, helps us think about the kinds of minds that underlie and produce such behaviors.

In Chapter 13, we outline different paradigms in robotics. Some of these approaches differ radically from one another. The **hierarchical paradigm** offers a “top-down” perspective, according to which a robot is programmed with knowledge about the world. The robot then uses this model or internal representation to guide its actions. The reactive



paradigm, on the other hand, is “bottom up.” Robots that use this architecture respond in a simple way to environmental stimuli: They react reflexively to a stimulus input, and there is little in the way of intervening knowledge.

### ***Integrating Approaches***

Many of the approaches we have just listed inform one another. For instance, the fields of AI and robotics are in some cases inseparable. AI programmers often write computer programs that serve as the “brains” of robots, telling them what to do and providing them with instructions on how to perform various tasks like object recognition and manipulation. In recent years, the cognitive and neuroscience approaches have come closer together, with cognitive psychologists providing information-processing models of specific brain processes. For example, we have seen cognitive models of hippocampal function that specify how this brain region stores memories based on our understanding of the neural substrates. In an even more interdisciplinary leap, we have seen the new area of social cognitive neuroscience in which social contexts are attached to cognitive-neural models. These are discussed in Chapter 11. In the concluding chapter (Chapter 14), we provide a more comprehensive summary of the ways in which these different approaches are and can be integrated.

### **SUMMING UP: A REVIEW OF CHAPTER 1**

1. Cognitive science is the scientific interdisciplinary study of mind and sees contributions from multiple fields, including philosophy and its newest offshoot, experimental philosophy, psychology, cognitive psychology, neuroscience, the connectionist or network approach, evolution, linguistics, the scientific study of emotions and social behavior, AI, robotics, and more.
2. Mind can be considered an information processor. At least some mental operations bear similarity to the way information is processed in a computer.
3. Information processing requires that some aspect of the world be represented and then operated on or computed. A representation is symbolic if it stands for something else. The thing a symbol stands for in the world is called its referent. The fact that symbols are “about” these things is called intentionality.
4. A formal system is made of symbols—collections of symbols that form expressions and processes that act on those expressions to form new expressions. Formal logic is an example of a formal system.

5. According to the physical symbol system hypothesis, formal systems can be said to be intelligent implying that computers may be intelligent. A problem for this is the symbol grounding problem, which states that symbols in people are grounded because we have bodies, can perceive objects, and can act on them.
6. Examples of representations are concepts, propositions, rules, and analogies. Representations are realized by an information bearer, have content, are grounded, and need to be interpreted.
7. Computations are processes that act on or transform representations. According to the tri-level hypothesis, there seem to be at least three levels of computation: computational, algorithmic, and implementational.
8. There are several different schools of thought that differ in the way they view representation and computation. In the classical cognitive science view, representations are fixed symbols and information processing is serial. In the connectionist view, representations are distributed and processing is parallel. According to the dynamical view, representations are constantly changing, being altered with each new experience.

### EXPLORE MORE

Log on to the student study site at <http://study.sagepub.com/friedenberg3e> for electronic flash cards, review quizzes, and a list of web resources to aid you in further exploring the field of cognitive science.

### SUGGESTED READINGS

- Friedenberg, J. (2009). *Dynamical psychology: Complexity, self-organization, and mind*. Charlotte, NC: Emergent.
- Harnish, R. M. (2002). *Minds, brains, computers: An introduction to the foundations of cognitive science*. Malden, MA: Blackwell.
- Sobel, C. P. (2001). *The cognitive sciences: An interdisciplinary approach*. Mountain View, CA: Mayfield.
- Stainton, R. J. (2006). *Contemporary debates in cognitive science*. Malden, MA: Blackwell.