or scientists, for that matter—predict, months in advance, the trajectories of two objects thousands of miles apart to an accuracy of inches and minutes? And do it from information that can be conveyed in a few seconds of conversation? The calculus behind this forecasting is intuitive psychology: the knowledge that I *want* to meet my friend and vice versa, and that each of us *believes* the other will be at a certain place at a certain time and *knows* a sequence of rides, hikes, and flights that will take us there. No science of mind or brain is ever likely to do better. That does not mean that the intuitive psychology of beliefs and desires is itself a science, but it suggests that scientific psychology will have to explain how a hunk of matter, such as a human being, can have beliefs and desires and how the beliefs and desires work so well.

The traditional explanation of intelligence is that human flesh is suffused with a non-material entity, the soul, usually envisioned as some kind of ghost or spirit. But the theory faces an insurmountable problem: How does the spook interact with solid matter? How does an ethereal nothing respond to flashes, pokes, and beeps and get arms and legs to move? Another problem is the overwhelming evidence that the mind is the activity of the brain. The supposedly immaterial soul, we now know, can be bisected with a knife, altered by chemicals, started or stopped by electricity, and extinguished by a sharp blow or by insufficient oxygen. Under a microscope, the brain has a breathtaking complexity of physical structure fully commensurate with the richness of the mind.

Another explanation is that mind comes from some extraordinary form of matter. Pinocchio was animated by a magical kind of wood found by Geppetto that talked, laughed, and moved on its own. Alas, no one has ever discovered such a wonder substance. At first one might think that the wonder substance is brain tissue. Darwin wrote that the brain "secretes" the mind, and recently the philosopher John Searle has argued that the physico-chemical properties of brain tissue somehow produce the mind just as breast tissue produces milk and plant tissue produces sugar. But recall that the same kinds of membranes, pores, and chemicals are found in brain tissue throughout the animal kingdom, not to mention in brain tumors and cultures in dishes. All of these globs of neural tissue have the same physico-chemical properties, but not all of them accomplish humanlike intelligence. Of course, *something* about the tissue in the human brain is necessary for our intelligence, but the physical properties are not sufficient, just as the physical properties of bricks are not sufficient to explain architecture and the physical properties of oxide particles are not sufficient to explain music. Something in the *patterning* of neural tissue is crucial.

Intelligence has often been attributed to some kind of energy flow or force field. Orbs, luminous vapors, auras, vibrations, magnetic fields, and lines of force figure prominently in spiritualism, pseudoscience, and science-fiction kitsch. The school of Gestalt psychology tried to explain visual illusions in terms of electromagnetic force fields on the surface of the brain, but the fields were never found. Occasionally the brain surface has been described as a continuous vibrating medium that supports holograms or other wave interference patterns, but that idea, too, has not panned out. The hydraulic model, with its psychic pressure building up, bursting out, or being diverted through alternative channels, lay at the center of Freud's theory and can be found in dozens of everyday metaphors: anger welling up, letting off steam, exploding under the pressure, blowing one's stack, venting one's feelings, bottling up rage. But even the hottest emotions do not literally correspond to a buildup and discharge of energy (in the physicist's sense) somewhere in the brain. In Chapter 6 I will try to persuade you that the brain does not actually *operate by* internal pressures but *contrives them* as a negotiating tactic, like a terrorist with explosives strapped to his body.

A problem with all these ideas is that even if we *did* discover some gel or vortex or vibration or orb that spoke and plotted mischief like Geppetto's log, or that, more generally, made decisions based on rational rules and pursued a goal in the face of obstacles, we would still be faced with the mystery of *how* it accomplished those feats.

No, intelligence does not come from a special kind of spirit or matter or energy but from a different commodity, *information*. Information is a correlation between two things that is produced by a lawful process (as opposed to coming about by sheer chance). We say that the rings in a stump carry information about the age of the tree because their number correlates with the tree's age (the older the tree, the more rings it has), and the correlation is not a coincidence but is caused by the way trees grow. Correlation is a mathematical and logical concept; it is not defined in terms of the stuff that the correlated entities are made of.

Information itself is nothing special; it is found wherever causes leave effects. What is special is information *processing*. We can regard a piece of matter that carries information about some state of affairs as a symbol; it can "stand for" that state of affairs. But as a piece of matter, it can do other things as well—physical things, whatever that kind of matter in that kind of state can do according to the laws of physics and chemistry. Tree rings carry information about age, but they also reflect light and absorb staining material. Footprints carry information about animal motions, but they also trap water and cause eddies in the wind.

Now here is an idea. Suppose one were to build a machine with parts that are affected by the physical properties of some symbol. Some lever or electric eye or tripwire or magnet is set in motion by the pigment absorbed by a tree ring, or the water trapped by a footprint, or the light reflected by a chalk mark, or the magnetic charge in a bit of oxide. And suppose that the machine then causes something to happen in some other pile of matter. It burns new marks onto a piece of wood, or stamps impressions into nearby dirt, or charges some other bit of oxide. Nothing special has happened so far; all I have described is a chain of physical events accomplished by a pointless contraption.

Here is the special step. Imagine that we now try to interpret the newly arranged piece of matter using the scheme according to which the original piece carried information. Say we *count* the newly burned wood rings and interpret them as the age of some tree at some time, even though they were not caused by the growth of any tree. And let's say that the machine was carefully designed so that the interpretation of its new markings made sense—that is, so that they carried information about something in the world. For example, imagine a machine that scans the rings in a stump, burns one mark on a nearby plank for each ring, moves over to a smaller stump from a tree that was cut down at the same time, scans its rings, and sands off one mark in the plank for each ring. When we count the marks on the plank, we have the age of the first tree at the time that the second one was planted. We would have a kind of *rational* machine, a machine that produces true conclusions from true premises—not because of any special kind of matter or energy, or because of any part that was itself intelligent or rational. All we have is a carefully contrived chain of ordinary physical events, whose first link was a configuration of matter that carries information. Our rational machine owes its rationality to two properties glued together in the entity we call a symbol: a symbol carries information, and it causes things to happen. (Tree rings correlate with the age of the tree, and they can absorb the light beam of a scanner.)

When the caused things themselves carry information, we call the whole system an information processor, or a computer.

Now, this whole scheme might seem like an unrealizable hope. What guarantee is there that *any* collection of thingamabobs can be arranged to fall or swing or shine in just the right pattern so that when their effects are interpreted, the interpretation will make sense? (More precisely, so that it will make sense according to some prior law or relationship we find interesting; any heap of stuff can be given a contrived interpretation after the fact.) How confident can we be that some machine will make marks that actually correspond to some meaningful state of the world, like the age of a tree when another tree was planted, or the average age of the tree's offspring, or anything else, as opposed to being a meaningless pattern corresponding to nothing at all?

The guarantee comes from the work of the mathematician Alan Turing. He designed a hypothetical machine whose input symbols and output symbols could correspond, depending on the details of the machine, to any one of a vast number of sensible interpretations. The machine consists of a tape divided into squares, a read-write head that can print or read a symbol on a square and move the tape in either direction, a pointer that can point to a fixed number of tickmarks on the machine, and a set of mechanical reflexes. Each reflex is triggered by the symbol being read and the current position of the pointer, and it prints a symbol on the tape, moves the tape, and/or shifts the pointer. The machine is allowed as much tape as it needs. This design is called a Turing machine.

What can this simple machine do? It can take in symbols standing for a number or a set of numbers, and print out symbols standing for new numbers that are the corresponding value for any mathematical function that can be solved by a step-by-step sequence of operations (addition, multiplication, exponentiation, factoring, and so on—I am being imprecise to convey the importance of Turing's discovery without the technicalities). It can apply the rules of any useful logical system to derive true statements from other true statements. It can apply the rules of any grammar to derive well-formed sentences. The equivalence among Turing machines, calculable mathematical functions, logics, and grammars, led the logician Alonzo Church to conjecture that *any* well-defined recipe or set of steps that is guaranteed to produce the solution to some problem in a finite amount of time (that is, any algorithm) can be implemented on a Turing machine.

What does this mean? It means that to the extent that the world other definitions of computation. For example, we need not assume that the computation is made up of a sequence of discrete steps, that the symbols must be either completely present or competely absent (as opposed to being stronger or weaker, more active or less active), that a correct answer is guaranteed in a finite amount of time, or that the truth value be "absolutely true" or "absolutely false" rather than a probability or a degree of certainty. The computational theory thus embraces an alternative kind of computer with many elements that are active to a degree corresponding to the *probability* that some statement is true or false, and in which the activity levels change smoothly to register new and roughly accurate probabilities. (As we shall see, that may be the way the brain works.) The key idea is that the answer to the question "What makes a system smart?" is not the kind of stuff it is made of or the kind of energy flowing through it, but what the parts of the machine stand for and how the patterns of changes inside it are designed to mirror truth-preserving relationships (including probabilistic and fuzzy truths).

NATURAL COMPUTATION

Why should you buy the computational theory of mind? Because it has solved millennia-old problems in philosophy, kicked off the computer revolution, posed the significant questions of neuroscience, and provided psychology with a magnificently fruitful research agenda.

Generations of thinkers have banged their heads against the problem of how mind can interact with matter. As Jerry Fodor has put it, "Self-pity can make one weep, as can onions." How can our intangible beliefs, desires, images, plans, and goals reflect the world around us and pull the levers by which we, in turn, shape the world? Descartes became the laughingstock of scientists centuries after him (unfairly) because he proposed that mind and matter were different kinds of stuff that somehow interacted in a part of the brain called the pineal gland. The philosopher Gilbert Ryle ridiculed the general idea by calling it the Doctrine of the Ghost in the Machine (a phrase that was later co-opted for book titles by the writer Arthur Koestler and the psychologist Stephen Kosslyn and for an album title by the rock group The Police). Ryle and other philosophers argued that mentalistic terms such as "beliefs," "desires," and "images" are meaningless and come from sloppy misunderstandings of language, as if someone heard the expression "for Pete's sake" and went around looking for Pete. Simpatico behaviorist psychologists claimed that these invisible entities were as unscientific as the Tooth Fairy and tried to ban them from psychology.

And then along came computers: fairy-free, fully exorcised hunks of metal that could not be explained without the full lexicon of mentalistic taboo words. "Why isn't my computer printing?" "Because the program doesn't *know* you replaced your dot-matrix printer with a laser printer. It still *thinks* it is *talking to* the dot-matrix and is *trying* to print the document by *asking* the printer to *acknowledge* its *message*. But the printer doesn't *understand* the message; it's *ignoring* it because it *expects* its input to begin with '%!' The program *refuses* to *give up control* while it *polls* the printer, so you have to *get the attention* of the *monitor* so that it can *wrest control* back from the program. Once the program *learns* what printer is connected to it, they can *communicate*." The more complex the system and the more expert the users, the more their technical conversation sounds like the plot of a soap opera.

Behaviorist philosophers would insist that this is all just loose talk. The machines aren't really understanding or trying anything, they would say; the observers are just being careless in their choice of words and are in danger of being seduced into grave conceptual errors. Now, what is wrong with this picture? The *philosophers* are accusing the *computer scientists* of fuzzy thinking? A computer is the most legalistic, persnickety, hard-nosed, unforgiving demander of precision and explicitness in the universe. From the accusation you'd think it was the befuddled computer scientists who call a philosopher when their computer stops working rather than the other way around. A better explanation is that computation has finally demystified mentalistic terms. Beliefs are inscriptions in memory, desires are goal inscriptions, thinking is computation, perceptions are inscriptions triggered by sensors, trying is executing operations triggered by a goal.

(You are objecting that we humans *feel* something when we have a belief or a desire or a perception, and a mere inscription lacks the power to create such feelings. Fair enough. But try to separate the problem of explaining intelligence from the problem of explaining conscious feelings. So far I'm trying to explain intelligence; we'll get to consciousness later in the chapter.)

The computational theory of mind also rehabilitates once and for all the infamous homunculus. A standard objection to the idea that thoughts are internal representations (an objection popular among scientists trying to show how tough-minded they are) is that a representation would require a little man in the head to look at it, and the little man would require an even littler man to look at the representations inside him, and so on, ad infinitum. But once more we have the spectacle of the theoretician insisting to the electrical engineer that if the engineer is correct his workstation must contain hordes of little elves. Talk of homunculi is indispensable in computer science. Data structures are read and interpreted and examined and recognized and revised all the time, and the subroutines that do so are unashamedly called "agents," "demons," "supervisors," "monitors," "interpreters," and "executives." Why doesn't all this homunculus talk lead to an infinite regress? Because an internal representation is not a lifelike photograph of the world, and the homunculus that "looks at it" is not a miniaturized copy of the entire system, requiring its entire intelligence. That indeed would have explained nothing. Instead, a representation is a set of symbols corresponding to *aspects* of the world, and each homunculus is required only to react in a few circumscribed ways to some of the symbols, a feat far simpler than what the system as a whole does. The intelligence of the system emerges from the activities of the not-so-intelligent mechanical demons inside it. The point, first made by Jerry Fodor in 1968, has been succinctly put by Daniel Dennett:

> Homunculi are *bogeymen* only if they duplicate *entire* the talents they are rung in to explain. . . . If one can get a team or committee of *relatively* ignorant, narrow-minded, blind homunculi to produce the intelligent behavior of the whole, this is progress. A flow chart is typically the organizational chart of a committee of homunculi (investigators, librarians, accountants, executives); each box specifies a homunculus by prescribing a function *without saying how it is accomplished* (one says, in effect: put a little man in there to do the job). If we then look closer at the individual boxes we see that the function of each is accomplished by subdividing it via another flow chart into still smaller, more stupid homunculi. Eventually this nesting of boxes within boxes lands you with homunculi so stupid (all they have to do is remember whether to say yes or no when asked) that they can be, as one says, "replaced by a machine." One *discharges* fancy homunculi from one's scheme by organizing armies of idiots to do the work.

in a chess-playing computer turns out, by a remarkable coincidence, to be identical to the battlefield events in the Six-Day War (King's knight = Moshe Dayan, Rook to c7 = Israeli army captures the Golan Heights, and so on). Would the program be "about" the Six-Day War every bit as much as it is "about" the chess game? Suppose that someday we discovered that cats are not animals after all, but lifelike robots controlled from Mars. Any inference rule that computed "If it's a cat, then it must be an animal" would be inoperative. The inferential role of our mental symbol `cat` would have changed almost beyond recognition. But surely the *meaning* of `cat` would be unchanged: you'd still be thinking "cat" when Felix the Robot slunk by. Score two points for the causal theory.

A third view is summarized by the television ad parody on *Saturday Night Live*: You're both right—it's a floor wax *and* a dessert topping. *Together* the causal and inferential roles of a symbol determine what it represents. (On this view, Swampman's thoughts would be about my mother because he has a *future*-oriented causal connection with her: he can recognize her when he meets her.) Causal and inferential roles tend to be in sync because natural selection designed both our perceptual systems and our inference modules to work accurately, most of the time, in this world. Not all philosophers agree that causation plus inference plus natural selection are enough to nail down a concept of "meaning" that would work perfectly in all worlds. ("Suppose Swampman has an identical twin on another planet . . .") But if so, one might respond, so much the worse for that concept of meaning. Meaning might make sense only relative to a device that was designed (by engineers or by natural selection) to function in a particular kind of world. In other worlds—Mars, Swampland, the Twilight Zone—all bets are off. Whether or not the causal-plus-inferential theory is completely philosopher-proof, it takes the mystery out of how a symbol in a mind or a machine can mean something.

Another sign that the computational theory of mind is on the right track is the existence of artificial intelligence: computers that perform humanlike intellectual tasks. Any discount store can sell you a computer that surpasses a human's ability to calculate, store and retrieve facts, draft drawings, check spelling, route mail, and set type. A well-stocked software house can sell you programs that play excellent chess and that recognize alphabetic characters and carefully pronounced speech. Clients with deeper pockets can buy programs that respond to questions in English about restricted topics, control robot arms that weld and spray-paint, and duplicate human expertise in hundreds of areas such as picking stocks, diagnosing diseases, prescribing drugs, and troubleshooting equipment breakdowns. In 1996 the computer Deep Blue defeated the world chess champion Gary Kasparov in one game and played him to a draw in two others before losing the match, and it is only a matter of time before a computer defeats a world champion outright. Though there are no Terminator-class robots, there are thousands of smaller-scale artificial intelligence programs in the world, including some hidden in your personal computer, car, and television set. And progress continues.

These low-key successes are worth pointing out because of the emotional debate over What Computers Will-Soon/Won't-Ever Do. One side says robots are just around the corner (showing that the mind is a computer); the other side says it will never happen (showing that it isn't). The debate seems to come right out of the pages of Christopher Cerf and Victor Navasky's *The Experts Speak*:

> Well-informed people know it is impossible to transmit the voice over wires and that were it possible to do so, the thing would be of no practical value.
>
> —Editorial, *The Boston Post*, 1865

> Fifty years hence . . . [w]e shall escape the absurdity of growing a whole chicken in order to eat the breast or wing, by growing these parts separately under a suitable medium.
>
> —Winston Churchill, 1932

> Heavier-than-air flying machines are impossible.
>
> —Lord Kelvin, pioneer in thermodynamics and electricity, 1895

> [By 1965] the deluxe open-road car will probably be 20 feet long, powered by a gas turbine engine, little brother of the jet engine.
>
> —Leo Cherne, editor-publisher of The Research Institute of America, 1955

> Man will never reach the moon, regardless of all future scientific advances.
>
> —Lee Deforest, inventor of the vacuum tube, 1957

> Nuclear powered vacuum cleaners will probably be a reality within 10 years.
>
> —Alex Lewyt, manufacturer of vacuum cleaners, 1955

The one prediction coming out of futurology that is undoubtedly correct is that in the future today's futurologists will look silly. The ultimate attainments of artificial intelligence are unknown, and will depend on countless practical vicissitudes that will be discovered only as one goes along. What is indisputable is that computing machines can be intelligent.

Scientific understanding and technological achievement are only loosely connected. For some time we have understood much about the hip and the heart, but artificial hips are commonplace while artificial hearts are elusive. The pitfalls between theory and application must be kept in mind when we look to artificial intelligence for clues about computers and minds. The proper label for the study of the mind informed by computers is not Artificial Intelligence but Natural Computation.

The computational theory of mind has quietly entrenched itself in neuroscience, the study of the physiology of the brain and nervous system. No corner of the field is untouched by the idea that information processing is the fundamental activity of the brain. Information processing is what makes neuroscientists more interested in neurons than in glial cells, even though the glia take up more room in the brain. The axon (the long output fiber) of a neuron is designed, down to the molecule, to propagate information with high fidelity across long separations, and when its electrical signal is transduced to a chemical one at the synapse (the junction between neurons), the physical format of the information changes while the information itself remains the same. And as we shall see, the tree of dendrites (input fibers) on each neuron appears to perform the basic logical and statistical operations underlying computation. Information-theoretic terms such as "signals," "codes," "representations," "transformations," and "processing" suffuse the language of neuroscience.

Information processing even defines the legitimate questions of the field. The retinal image is upside down, so how do we manage to see the world right-side up? If the visual cortex is in the back of the brain, why doesn't it feel like we are seeing in the back of our heads? How is it possible that an amputee can feel a phantom limb in the space where his real limb used to be? How can our experience of a green cube arise from neurons that are neither colored green nor in the shape of a cube? Every neuroscientist knows that these are pseudo-questions, but why? Because they are about properties of the brain that make no difference to the transmission and processing of information.

If a scientific theory is only as good as the facts it explains and the discoveries it inspires, the biggest selling point for the computational theory of mind is its impact on psychology. Skinner and other behaviorists insisted that all talk about mental events was sterile speculation; only stimulus-response connections could be studied in the lab and the field. Exactly the opposite turned out to be true. Before computational ideas were imported in the 1950s and 1960s by Newell and Simon and the psychologists George Miller and Donald Broadbent, psychology was dull, dull, dull. The psychology curriculum comprised physiological psychology, which meant reflexes, and perception, which meant beeps, and learning, which meant rats, and memory, which meant nonsense syllables, and intelligence, which meant IQ, and personality, which meant personality tests. Since then psychology has brought the questions of history's deepest thinkers into the laboratory and has made thousands of discoveries, on every aspect of the mind, that could not have been dreamed of a few decades ago.

The blossoming came from a central agenda for psychology set by the computational theory: discovering the form of mental representations (the symbol inscriptions used by the mind) and the processes (the demons) that access them. Plato said that we are trapped inside a cave and know the world only through the shadows it casts on the wall. The skull is our cave, and mental representations are the shadows. The information in an internal representation is all that we can know about the world. Consider, as an analogy, how *external* representations work. My bank statement lists each deposit as a single sum. If I deposited several checks and some cash, I cannot verify whether a particular check was among them; that information was obliterated in the representation. What's more, the *form* of a representation determines what can easily be inferred from it, because the symbols and their arrangement are the only things a homunculus stupid enough to be replaced by a machine can respond to. Our representation of numbers is valuable because addition can be performed on the numbers with a few dronelike operations: looking up entries in the addition table and carrying digits. Roman numerals have not survived, except as labels or decorations, because addition operations are far more complicated with them, and multiplication and division operations are practically impossible.

Pinning down mental representations is the route to rigor in psychology. Many explanations of behavior have an airy-fairy feel to them because they explain psychological phenomena in terms of other, equally mysterious psychological phenomena. Why do people have more trouble with this task than with that one? Because the first one is "more difficult." Why do people generalize a fact about one object to another object? Because the objects are "similar." Why do people notice this event but not that one? Because the first event is "more salient." These explanations are scams. Difficulty, similarity, and salience are in the mind of the beholder, which is what we should be trying to explain. A computer finds it more difficult to remember the gist of *Little Red Riding Hood* than to remember a twenty-digit number; you find it more difficult to remember the number than the gist. You find two crumpled balls of newspaper to be similar, even though their shapes are completely different, and find two people's faces to be different, though their shapes are almost the same. Migrating birds that navigate at night by the stars in the sky find the positions of the constellations at different times of night quite salient; to a typical person, they are barely noticeable.

But if we hop down to the level of representations, we find a firmer sort of entity, which can be rigorously counted and matched. If a theory of psychology is any good, it should predict that the representations required by the "difficult" task contain more symbols (count 'em) or trigger a longer chain of demons than those of the "easy" task. It should predict that the representations of two "similar" things have more shared symbols and fewer nonshared symbols than the representations of "dissimilar" things. The "salient" entities should have different representations from their neighbors; the "nonsalient" entities should have the same ones.

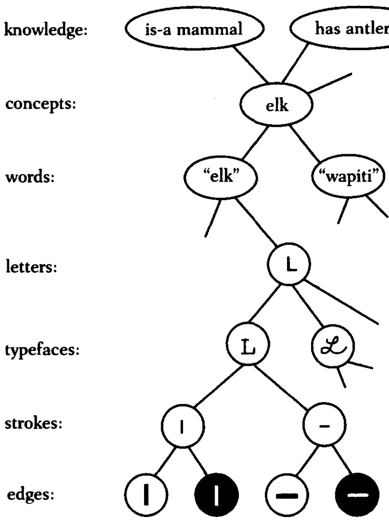
Research in cognitive psychology has tried to triangulate on the mind's internal representations by measuring people's reports, reaction times, and errors as they remember, solve problems, recognize objects, and generalize from experience. The way people generalize is perhaps the most telltale sign that the mind uses mental representations, and lots of them.

Suppose it takes a while for you to learn to read a fancy new typeface, festooned with curlicues. You have practiced with some words and are now as quick as you are for any other typeface. Now you see a familiar word that was not in your practice set—say, *elk*. Do you have to relearn that the word is a noun? Do you have to relearn how to pronounce it? Relearn that the referent is an animal? What the referent looks like? That it has mass and breathes and suckles its young? Surely not. But this banal talent of yours tells a story. Your knowledge about the word *elk* could not have been connected directly to the physical shapes of printed letters. If it had, then when new letters were introduced, your knowledge would have no connection to them and would be unavailable until you learned the connections anew. In reality, your knowledge must have been connected to a node, a number, an address in memory, or an entry in a mental dictionary representing the abstract word *elk*, and that entry must be neutral with respect to how it is printed or pronounced. When you learned the new typeface, you created a new visual trigger for the letters of the alphabet, which in turn triggered the old *elk* entry, and everything hooked up to the entry was instantly available, without your having to reconnect, piece by piece, everything you know about elks to the new way of printing *elk*. This is how we know that your mind contains mental representations specific to abstract entries for words, not just the shapes of the words when they are printed.

These leaps, and the inventory of internal representations they hint at, are the hallmark of human cognition. If you learned that *wapiti* was another name for an elk, you could take all the facts connected to the word *elk* and instantly transfer them to *wapiti*, without having to solder new connections to the word one at a time. Of course, only your zoological knowledge would transfer; you would not expect *wapiti* to be *pronounced* like *elk*. That suggests you have a level of representation specific to the concepts behind the words, not just the words themselves. Your knowledge of facts about elks hangs off the concept; the words *elk* and *wapiti* also hang off the concept; and the spelling *e-l-k* and pronunciation [ɛlk] hang off the word *elk*.

We have moved upward from the typeface; now let's move downward. If you had learned the typeface as black ink on white paper, you wouldn't have to relearn it for white ink on red paper. This unmasks a representation for visual edges. Any color abutting any other color is seen as an edge; edges define strokes; an arrangement of strokes makes up an alphanumeric character.

The various mental representations connected with a concept like an elk can be shown in a single diagram, sometimes called a semantic network, knowledge representation, or propositional database.



This is a fragment of the immense multimedia dictionary, encyclopedia, and how-to manual we keep in our heads. We find these layers upon layers of representations everywhere we look in the mind. Say I asked you to print the word *elk* in any typeface you wanted, but with your left hand (if you are a righty), or by writing it in the sand with your toe, or by tracing it with a penlight held in your teeth. The printing would be messy but recognizable. You might have to practice to get the motions to be smoother, but you would not have to relearn the strokes composing each letter, let alone the alphabet or the spelling of every English word. This transfer of skill must tap into a level of representation for motor control that specifies a geometric trajectory, not the muscle contractions or limb movements that accomplish it. The trajectory would be translated into actual motions by lower-level control programs for each appendage.

Or recall Sally escaping from the burning building earlier in this chapter. Her desire must have been couched as the abstract representation flee-from-danger. It could not have been couched as run-from-smoke, because the desire could have been triggered by signs other than smoke (and sometimes smoke would not trigger it), and her flight could have been accomplished by many kinds of action, not just running. Yet her behavioral response was put together for the first time there and then. Sally must be modular: one part of her assesses danger, another decides whether to flee, yet another figures out how to flee.

The combinatorics of mentalese, and of other representations composed of parts, explain the inexhaustible repertoire of human thought and action. A few elements and a few rules that combine them can generate an unfathomably vast number of different representations, because the number of possible representations grows exponentially with their size. Language is an obvious example. Say you have ten choices for the word to begin a sentence, ten choices for the second word (yielding a hundred two-word beginnings), ten choices for the third word (yielding a thousand three-word beginnings), and so on. (Ten is in fact the approximate geometric mean of the number of word choices available at each point in assembling a grammatical and sensible sentence.) A little arithmetic shows that the number of sentences of twenty words or less (not an unusual length) is about 10^{20}: a one followed by twenty zeros, or a hundred million trillion, or a hundred times the number of seconds since the birth of the universe. I bring up the example to impress you not with the vastness of language but with the vastness of thought. Language, after all, is not scat-singing: every sentence expresses a distinct idea. (There are no truly synonymous sentences.) So in addition to whatever ineffable thoughts people might have, they can entertain something like a hundred million trillion different effable thoughts.

The combinatorial immensity of thinkable structures is found in many spheres of human activity. The young John Stuart Mill was alarmed to discover that the finite number of musical notes, together with the maximum practical length of a musical piece, meant that the world would soon run out of melodies. At the time he sank into this melancholy, Brahms, Tchaikovsky, Rachmaninoff, and Stravinsky had not yet been born, to say nothing of the entire genres of ragtime, jazz, Broadway musicals, electric blues, country and western, rock and roll, samba, reggae, and punk. We are unlikely to have a melody shortage anytime soon because music is combinatorial: if each note of a melody can be selected from, say, eight notes on average, there are 64 pairs of notes, 512 motifs of three notes, 4,096 phrases of four notes, and so on, multiplying out to trillions and trillions of musical pieces.