MENTAL LEAPS

ANALOGY IN CREATIVE THOUGHT

KEITH J. HOLYOAK AND PAUL THAGARD

Mental Leaps

Analogy in Creative Thought

Keith J. Holyoak and Paul Thagard

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Third printing, 1999
First MIT Press paperback edition, 1996

This book was set in Bembo by DEKR Corporation, Woburn, Massachusetts, and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Holyoak, Keith James, 1950–
Mental leaps : analogy in creative thought / Keith J. Holyoak and Paul Thagard.
p. cm.
Includes bibliographical references and index.
ISBN 0-262-58144-2 (PB)
1. Analogy. 2. Creative thinking. I. Thagard, Paul. II. Title.
BD190.H64 1995
169—dc20 94-22734
CIP

Contents

Preface

Our interests in analogy can be traced back to 1975 when we were in graduate school. Keith Holyoak was studying cognitive psychology at Stanford University, where he was working on the representation of semantic relations in human memory (a topic that he would connect a few years later with the use of analogy in problem solving). Meanwhile Paul Thagard was a graduate student in philosophy at the University of Toronto, where he became fascinated with the creative use of analogy in the work of scientists such as Darwin. After completing our doctoral degrees, we each arrived at the University of Michigan, where Thagard taught philosophy at the Dearborn campus and Holyoak began to work on analogy in problem solving. Richard Nisbett noticed that the psychologist and the philosopher were in different ways dealing with similar issues, and he introduced us in 1980. For the next few years we met regularly with him and John Holland to discuss analogy, along with many other kinds of reasoning. This four-way collaboration culminated in our book *Induction: Processes of Inference, Learning, and Discovery,* which appeared in 1986.

By then, Holyoak had moved to UCLA and Thagard to Princeton University, from which he went to the University of Waterloo in 1992. Nonetheless, we continued a collaboration focused directly on analogy. Our joint research, in which many students and colleagues also participated, produced a series of computer models of human use of analogy, coupled with extensive experimental investigations that provided guidance for the construction of our simulations.

After publishing many technical papers on analogy, we decided to take a broader look at what we and other researchers had learned about the topic over the previous fifteen years or so. Like *Induction,* the book that has resulted is genuinely collaborative; both of us have worked on every chapter. Analogy by its very nature freely oversteps the traditional boundaries between knowledge domains, making it possible to use ideas from one domain to achieve insights in another. But although a number of fine books have reviewed what is known about such specialized topics as the use of analogy by children, by scientists, by poets, and by political leaders, no single book has provided an integrated treatment of analogical thinking across many different domains. This book presents a general theory of analogical thinking, illustrated by applications drawn from the widest possible spectrum. We have not attempted the encyclopedic task of surveying all the valuable research on analogy that has been performed by researchers in many different fields, but have instead tried to show how our own theory can provide a unified explanation of the diverse operations and applications of analogy.

Analogy is a mental tool that everyone uses to some degree. Understanding how we draw analogies is important for people interested in the evolution of thinking in animals and in its development in children; for those whose focus is on either creative thinking or errors of everyday reasoning; for those concerned with how decisions are made in law, business, and politics; and for those striving to improve education. In order to make our discussion as accessible as possible, we have tried to keep technical matters to a minimum. For ease of reading, we have kept the text free of references and footnotes, but extensive reference notes keyed to pages in the text can be found at the end of the book. Throughout the book we emphasize the basic principles that govern the use of analogy. Our major aim in the book is to show how the richness and diversity of analogy in thought can be understood in terms of these principles. Of course, much still remains to be learned. As well as providing a progress report on the scientific analysis of analogical thinking, we hope this book provides some pointers for those who pursue the investigation further.

Acknowledgments

Along the long road that led to completion of this book, we have accumulated many debts. Each author has been involved in a long series of collaborative research projects on analogy, involving many students and colleagues. Keith Holyoak's work on this topic began in 1978 in collaboration with Mary Gick. Their project investigated the use of analogy in problem solving, which became a continuing focus of subsequent research. Later work in Holyoak's laboratories was conducted in collaboration with Dorrit Billman, Richard Catrambone, Ellen Junn, and Kyunghee Koh (at the University of Michigan), Miriam Bassok (at the University of Pittsburgh), and Bruce Burns, Paul Downing, John Hummel, Trent Lange, Eric Melz, Laura Novick, Nina Robin, Bobbie Spellman, Jeff Thompson, Charles Wharton, and Tom Wickens (at UCLA). Bobbie Spellman's contributions to the book were especially wide-ranging, including not only relevant experiments but also discussions of the role of analogy in political decision making, suggestions and legal research on analogical arguments from the law, and critical comments on the entire first draft. Paul Thagard also was aided immensely by a strong group of collaborators who worked with him at Princeton, including Dawn Cohen, David Gochfeld, Susan Hardy, and especially Greg Nelson. At Waterloo, he has similarly benefited from working with Allison Barnes, Lori Buchanan, John Ching, Roy Fleck, Steve Joordens, and Cameron Shelley.

We also learned from many careful reviews and discussions of early drafts of the book. Patricia Cheng was always ready on short notice to help Keith Holyoak grapple with conceptual issues and to give rapid feedback on sections of the manuscript. In addition to reviewing a draft of the book, Graeme Halford provided many stimulating hours of dis-

We are grateful to the respective publishers for permission to include material published elsewhere:

Thagard, P. (1992) Analogy, explanation and education. *Journal of Research in Science Teaching* 29:537–44. John Wiley and Sons, Inc. (ch. 8)

Thagard, P. (1993). The greatest analogies in the history of science. *Canadian Artificial Intelligence* 31, (Winter), 14–20. (ch. 8)

Thagard, P., and E. Millgram (in press). Inference to the best plan: A coherence theory of decision. In *Goal-directed learning,* edited by A. Ram and D. B. Leake, Cambridge, Mass.: MIT Press. (ch. 6)

Our families provided us with some of the examples in this book, and a whole lot more. The book is dedicated to them.

1

First Steps

A Bird's Backyard

We usually think we see the world the way it actually is. This outlook is comforting and easy to maintain as we follow the rhythms of everyday life, dealing with familiar people and objects in routine situations. Understanding the world around us usually seems simple and effortless. But in fact our conception of the world is in large part a matter of our own creation. Our sense of direct understanding is an illusion, because the apparent simplicity of everyday comprehension arises from the subtlety and complexity of the human mind.

We all have moments when the illusion is shattered. Confronted with unfamiliar or surprising situations that do not readily fit into known patterns, we no longer feel we are seeing the world as it actually is. Instead, everything seems disordered and confused. When a parent is confronted by an unexpected crisis in dealing with a child, when a business finds its old strategies failing in the face of new competitors, or when a country tries to fathom the direction in which its new leader is heading—on such occasions the familiar patterns are broken.

It is then that we become aware of a conscious struggle to see the world as something we can understand. The mind must work to build new patterns. Often it is pressed to work quickly. Novelty can signal danger, a new problem that demands a solution before it is too late. If a primitive hunter is surprised by an unfamiliar large animal in the forest, or the leader of a modern democracy is confronted by the rise of a threatening dictator abroad, action cannot wait for the gradual accumulation of knowledge over hundreds of similar occurrences. In such situations we want a rapid understanding of the situation that brings with it some idea about what to do. Even a rough response will be more useful than a slow, in-depth analysis that requires many more observations than it is possible to make. A good way to proceed is to try to understand the novel challenge in terms of what is already known, even if making the connection requires a mental leap.

When we are young, before most of the familiar patterns of everyday life have been learned, such challenges are the rule rather than the exception. Knowledge of the world awaits construction. For a child the pool of known situations is still small, and novelty is the norm; so much has to be understood in terms of so little. But already the fundamental thought processes that guide the creation of understanding are hard at work. Perhaps it would be more accurate to say they are hard at play, since for a child (as for a scientist), understanding the world often becomes a game driven by natural curiosity.

Consider the following discussion between a mother and her four-year-old son, Neil, who was considering the deep issue of what a bird might use for a chair. Neil suggested, reasonably enough it would seem, that a tree could be a bird's chair. A bird might sit on a tree branch. His mother said that was so and added that a bird could sit on its nest as well, which is also its house. The conversation went on to other topics. But several minutes later, the child had second thoughts about what a tree is to a bird: "The tree is not the bird's chair—it's the bird's backyard!"

In this conversation Neil makes a mental leap, exploring connections between two very different domains. He is trying to understand the relatively unfamiliar world of creatures of the air in terms of the familiar patterns of everyday human households. This small example conveys what we mean by reasoning by *analogy,* or *analogical thinking*. The child's everyday world is the *source analog*: a known domain that the child already understands in terms of familiar patterns, such as people sitting on chairs and houses that open onto backyards. The bird's world is the *target analog*—a relatively unfamiliar domain that the child is trying to understand. Analogical thinking is not "logical" in the sense of a logical deduction—there is no reason why birds and people should necessarily organize their habitats in comparable ways. Yet the analogy is certainly not haphazard. In a loose sense, there is indeed some sort of logic—call it *analogic*—that constrains the way the child uses analogy to try to understand the target domain by seeing it in terms of the source domain.

If we look carefully at Neil's conversation with his mother, we can see the first steps in the use of analogy. As is often the case with children, Neil seems to begin with a question: What might a bird use for a chair?

But where does this question come from? Notice that the question contains the seeds of an analogical answer: Neil is looking for something in the world of a bird that corresponds to something in the world of a person. Already the basic connection has been made between the target and source domains. This connection is not at all random. In fact, using knowledge about people as the source analog for understanding non-human domains is so commonplace that it has a name, *personification.* Personification means to treat something that is not a person as if it were one. Such acts of imaginative creation lead into the realm of myth and metaphor, where death becomes not a mere physical process but the name of someone who walks the earth laboring as the Grim Reaper, and the moon becomes a mother who watches over us as we sleep. But personification begins in the minds of children, who from an early age try to understand animals, and sometimes plants and inanimate objects, in terms of people. A bird moves of its own volition, eats and drinks, and nurtures its young as they grow. These basic similarities suggest that a bird is sufficiently personlike to have other properties in common with people, such as needing a place to sit. But of course the child knows perfectly well that birds are not generally found sitting on chairs at the dinner table. So the answer to the question "What do birds sit on?" is not that they sit on chairs but that they must sit on something in their own aerial world that *corresponds* to a person's chair.

The child's initial answer had little or nothing to do with obvious physical resemblances. A tree, after all, does not really *look* like a chair. But it does serve a similar *function*: a bird can alight on a tree to rest, just as a person can sit on a chair to rest. The goal that motivated the child's question—the search for something that functions for a bird the way a chair functions for a person—also provides a basic criterion for identifying a plausible answer.

Yet, most surprisingly, Neil does not simply rest contented with what might seem to be a perfectly sensible answer to his question. Instead, he goes on to examine a different question, apparently triggered by his own answer to his initial query. What is a tree to a bird, really? A tree is a bird's chair in that a bird can sit on it. But a tree is a bird's backyard in that the bird's nest corresponds to a person's house, and the tree is located immediately outside of the nest, just as a backyard is immediately outside of the house. Neil concludes that the latter interpretation is somehow preferable. Even though he began with the question of what is a bird's chair, he arrived at an answer to the question of what is a bird's backyard.

This simple example of a four-year-old's analogical thinking gives rise to many surprises and puzzles.

- Preschool children, without any formal training, have a natural capacity to reason by analogy.
- The source and target analogs might never have been explicitly associated before. That is, we have no reason to believe that anyone told Neil that the world of a bird is supposed to be understood in terms of the world of a person.
- Analogical thinking is guided by the goal that triggers the analogy yet can actually give rise to a new goal that changes the way the analogy is used.
- Children evaluate their analogies on the basis of internal criteria that go beyond simple reward and punishment by adults. Neil spontaneously evaluated and revised his interpretation of the tree as a chair, not because his mother disputed this initial suggestion (to the contrary, she voiced her approval), but simply because she introduced a new element into the situation, the role of the nest as a house.
- Part of the evaluation is based on a kind of competition among alternative interpretations. Neil acts as if the tree can be potentially interpreted as either the bird's chair or its backyard, but not both. He gives up the "chair" interpretation not because he finds anything directly wrong with it, but because the rival "backyard" interpretation seems somehow to be more satisfying.

Perhaps the deepest puzzle raised by the example is to explain what led Neil to decide that a tree is not a bird's chair but rather its backyard. After all, by any literal criterion a tree is neither a chair nor a backyard, so why should one interpretation be preferred to the other? The answer takes us to the heart of analogical thinking, which involves establishing a *mapping,* or systematic set of correspondences, between the elements of the source and the target analog. In our example, Neil entertains two alternative possible mappings between the source analog (child's world) and the target analog (bird's world):

Source		*Target*	*Source*		*Target*
person	↔	**bird**	**person**	↔	**bird**
chair	↔	**tree**	**chair**	↔	**?**
			house	↔	**nest**
			backyard	↔	**tree**

We can see at a glance that the second mapping provides one more correspondence. In addition, the second mapping is based on a greater number of interconnected facts that relate the two domains to each other. Roughly, the first mapping captures the fact that a bird can sit on just one element in the source domain (and vice versa). It is for this reason that when Neil decides that the tree plays the role of a backyard, he no longer accepts it as playing the role of a chair.

Third, the exploration of the analogy is guided by the person's goals in using it, which provide the *purpose* for considering the analogy at all. Thus Neil initially was directed by his entering desire to understand what birds might sit on. But it was apparent that his broader goal was a rather playful quest to understand the bird's habitat in terms of his own familiar dwelling. As we will see in other examples, people generally tend to select especially familiar situations to serve as source analogs.

These three kinds of constraints—similarity, structure, and purpose—do not operate like rigid rules dictating the interpretation of analogies. Instead they function more like the diverse pressures that guide an architect engaged in creative design, with some forces converging, others in opposition, and their constant interplay pressing toward some satisfying compromise. One of our main goals in this book is to show how the human mind creatively uses these constraints together in thinking by analogy.

From Socrates to Velcro and Vietnam

Analogy is not just child's play. There is a common thread connecting the Greek philosopher Socrates, the twentieth-century inventor of Velcro fasteners, and the political leaders of the 1960s who were responsible for the American military debacle in Vietnam during the 1960s. The same thread connects the most famous kite flyer in history, family therapists, and marriage negotiators in New Guinea. All of these, and many others we will also describe in this book, have put analogy to work. The uses they have made of analogy are diverse: generating new ideas, producing inventions, making decisions about war and peace, communicating with other people. Socrates explained his philosophical work by comparing himself to a midwife. Georges de Mestral invented Velcro after noticing the way burrs would stick on his dog's fur. President Lyndon Johnson became convinced that Vietnam was a "domino" in Southeast Asia, the fall of which would cause its neighbors to also topple into the grasp of world communism. Benjamin Franklin conducted his kite experiment to test an analogy he drew between lightning and electricity. Family therapists sometimes try to help troubled families by describing other families with similar relationships and problems. In New Guinea, telling a neighbor a story about betel nuts can be a way to between the king and the thief is designed to lure David into an indirect acknowledgment of his own guilt. Similarly, the fables of Aesop (circa 600 B.C.) are stories about animals endowed with human intellect and foibles, whose activities can be mapped onto human situations. Recall the classic story about a fox who cannot reach the grapes he covets and finally pronounces them sour as he gives up trying to reach them. This tale has obvious human analogs, such as the job seeker who fails to obtain a desired position and in the aftermath tells people the job would have been boring anyhow. Indeed, we understand the expression "sour grapes" as the name of this type of situation. For those who remember Aesop, this term is a shorthand for the entire source analog. For those who do not, the phrase will seem more arbitrary, a metaphor cut off from its roots.

It may seem strange that people should speak in parables and metaphors. Why not just come out and say what we mean? It turns out there are often advantages to indirectness. Just as an architect finds it easier to first construct a scale model before going on to build an actual building, it may be easier to express an idea in a familiar source domain before trying to use it to build understanding of the target. The source provides a convenient and tractable model for constructing an explanation or a solution that can then be transferred to the more perplexing target domain.

In some cases it is too dangerous to speak directly. Nathan would have taken a far greater risk if he had simply walked up to David and denounced him as a murderous adulterer; by using his parable, he let the king denounce himself. When living in a repressive society lacking freedom of speech, a political satirist may sometimes reach a broader audience by constructing analogies instead of direct polemics. In the 1960s, when communism still prevailed in what was then the Soviet Union, the Russian novelist Aleksandr Solzhenitsyn wrote the novel *The Cancer Ward*. The book was ostensibly about life in a Soviet cancer ward in the spring of 1955, and the struggle of the patients to maintain their humanity and dignity in the face of the feared and debilitating disease. At the same time, the novel provided a transparent analog to the repression that had faced the citizens of that society under Stalin and their hope for some sort of social and political recovery. In some cultures, such delicate social negotiations as arranging a marriage are carried on by constructing stories that are understood as analogs. Then if anything goes wrong and negotiations collapse, the story can simply be dropped; nothing was ever actually said about a marriage, so no one has lost face if no marriage comes about.

And yet, compelling as these purposes of analogy are, something more fundamental also seems to be at work. There is something inherently pleasurable about finding a mesh between two superficially unrelated situations. Some basic human joy is triggered by the discovery of unexpected connections. To denounce Stalinism indirectly by describing life in a cancer ward is not simply to adopt an expedient disguise; the novel is far more compelling than any literal-minded political diatribe could be. To see one thing *as if* it were another creates a tension between two perspectives: the thing as itself and the thing as something else. To resolve this tension by finding an integrated interpretation is a satisfying achievement. Consider the opening lines of Dante's *Divine Comedy*:

Midway upon the journey of our life,
I found myself within a forest dark,
For the right road was lost.

We understand from this passage that the person's life is like a journey along a road, headed toward some destination. Like a traveler who has become lost in a wood, at midlife the person's life has unexpectedly lost its purpose. George Lakoff and Mark Turner, who have analyzed such poetic metaphors in terms of mappings between analogous domains, note that Dante has built on the fundamental conception of "life as a journey," which permeates everyday thinking. To understand the metaphor we must understand a life to be not only a succession of everyday human activities but also an organized chain of actions directed toward some ultimate goal that provides a criterion by which progress can be measured. The same basic conception of life as a journey helps us to easily understand the meaning of an indefinite number of commonplace expressions, such as "Her career was derailed" or "He lost his bearings after his wife left him." We generally do not feel we are performing a mental synthesis when we understand such conventional utterances; but the possibility of novel variations of the underlying analogical theme keeps the potential for creativity alive.

Ripples in the Air

Another fundamental purpose of analogy is to gain understanding that goes beyond the information we receive from our senses. We hear thunder and see lightning, for example, but how could any amount of sensory experience give rise to the idea that sound and light are transmitted by invisible waves? Indeed, how can we come up with any answers at all to even more basic questions, such as how the human race came into being, or what happens to us after death? To answer such questions our minds must somehow form ideas that go beyond anything we can directly experience.

Analogy provides one powerful mechanism for forming such concepts. Consider, for example, the conceptions of creation, life, and death that form the basis of many people's religious beliefs. The Christian Lord's Prayer begins, "Our Father who art in Heaven." That simple phrase conveys the idea that God is related to people as a father is to his children: as progenitor, protector, and disciplinarian. He exists in a place, Heaven, that is analogous to the earth we know but invisible to us. Many believe that the journey of life ends with a final passage to some other place—Heaven or its less attractive counterpart. God is understood to be more powerful than any human father, and Heaven more spacious and attractive than earth; God and Heaven are not visible to mortal eyes. But despite these differences, the basic conception of the divine sphere is built on an analogy to the social and physical world in which we live. The result is that God and Heaven are not vague abstractions; rather, these concepts are grounded by the analogy to direct experience. Analogy manages to give rise to ideas that take us beyond sensory experience while still maintaining conceptual links to it.

Religion and science are often taken to be entirely different enterprises, but analogy plays an important role in both. Scientific theories, like systems of religious belief, often propose unobservable things like subatomic particles, black holes, and gravity waves. Scientific theories, unlike religions, must eventually be evaluated in relation to observable evidence. For example, even if black holes cannot be observed directly, the astronomical theory that posits them does make potentially testable predictions about, for example, the motion of visible celestial objects. In contrast, faith that "whatever happens is the will of God" may provide us with spiritual comfort, but certainly not with testable predictions. But these differences in how religious and scientific concepts are used should not obscure the fact that the same mental mechanism, analogical thinking, may be involved in the origins of both. (We discuss theological analogies in chapter 7 and scientific analogies in chapter 8.)

The role of analogy in science can be traced back at least two thousand years. The first recorded use of analogy to develop an enduring scientific theory produced the idea that sound is propagated in the form of waves. Although we obviously hear sound, we do not have any way of directly perceiving the nature of its transmission. During the reign of the emperor Augustus, a Roman architect and engineer named Vitruvius wrote a set of ten short books on architecture. In these texts he laid out what amounted to the authoritative standards to be followed in extending Roman civilization: how to select a site for a city, how to build the walls, where to situate the temple, how to transport water by aqueducts. In his fifth book he describes how theaters should be built, drawing heavily on the legacy of the Greeks. In his discussion of the acoustic requirements for a theater, Vitruvius describes the nature of sound by analogy to water waves:

> Voice is a flowing breath of air, perceptible to the hearing by contact. It moves in an endless number of circular rounds, like the innumerably increasing circular waves which appear when a stone is thrown into smooth water, and which keep spreading indefinitely from the centre unless interrupted by narrow limits, or by some obstruction which prevents such waves from reaching their end in due formation. When they are interrupted by obstructions, the first waves, flowing back, break up the formation of those that follow.
>
> In the same manner the voice executes its movements in concentric circles; but while in the case of water the circles move horizontally on a plane surface, the voice not only proceeds horizontally, but also ascends vertically by regular stages. Therefore, as in the case of the waves formed in the water, so it is in the case of the voice: the first wave, when there is no obstruction to interrupt it, does not break up the second or the following waves, but they all reach the ears of the lowest and highest spectators without an echo.
>
> Hence the ancient architects, following in the footsteps of nature, perfected the ascending rows of seats in theatres from their investigations of the ascending voice.

Notice that Vitruvius not only points out some of the important ways in which sound and water waves behave similarly, such as the parallels between echoes and water waves striking an obstacle, but also some of their differences. He observes that whereas water waves spread out horizontally along a surface, sound waves move in three dimensions. And although too obvious for him to mention, it was doubtless apparent to Vitruvius that sound waves are not wet!

It turned out that the analogy between water and sound waves was fruitful and could centuries later be treated mathematically, moving well beyond Vitruvius's simple qualitative description. One important consequence was that the concept of wave, which now embraced both visible, wet, water waves moving along a plane and invisible, dry, sound

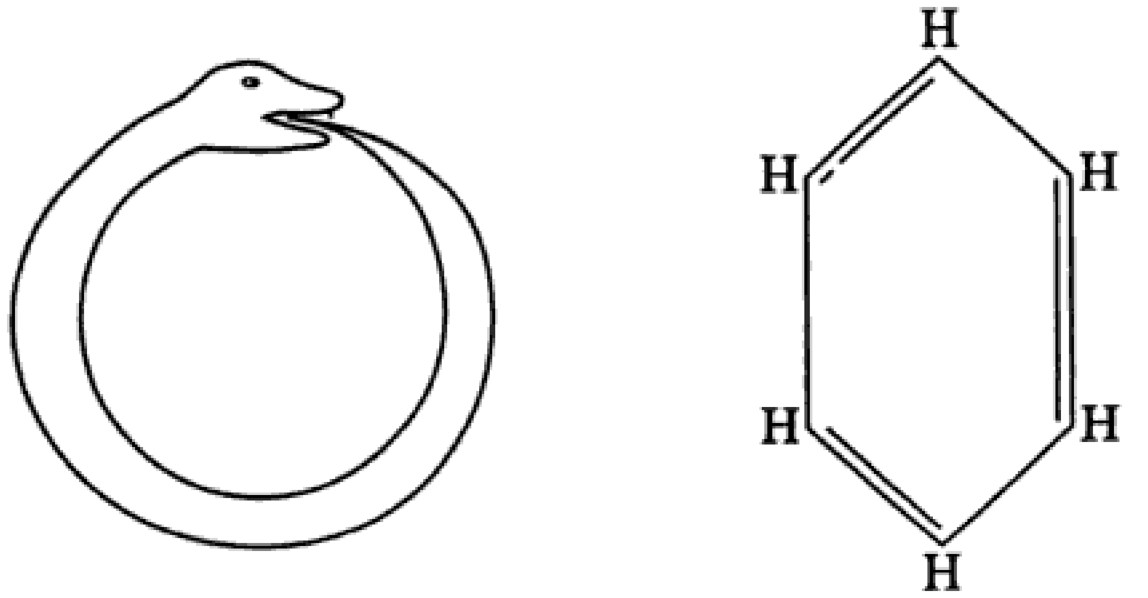

Figure 1.1
The visual analog of a snake biting its own tail contributed to Kekulé's insight that the molecular structure of benzene is circular.

he was led to the hypothesis that the carbon atoms in benzene are arranged in a ring by a reverie in which he saw a snake biting its own tail. This example, like a number of the other great scientific analogies we will describe in chapter 8, illustrates how visual representations can contribute to creative thinking using analogy.

George Polya, who wrote extensively on techniques to foster creative problem solving, went so far as to claim that "Analogy pervades all our thinking, our everyday speech and our trivial conclusions as well as artistic ways of expression and the highest scientific achievements." Although we do not believe that analogy is the only cognitive mechanism involved in creative thinking, it does play an important role. It has often been suggested that creativity is based on some mental mechanism for combining and recombining ideas in novel ways, where the recognition of viable new combinations depends in part on a kind of aesthetic judgment that the juxtaposed ideas fit well together. The mathematician Jules Henri Poincaré described his own creative process on a sleepless night when "ideas rose in crowds; I felt them collide until pairs interlocked, so to speak, making a stable combination." And among the "stable combinations," he argued, "the most fertile will often be those formed of elements drawn from domains which are far apart."

Arthur Koestler developed this idea further, proposing that creative thinking depends on "bisociation," the interlocking of two domains of knowledge previously seen as unrelated or even incompatible. He believed that bisociative thinking is the basis not only for major creative leaps but also for everyday appreciation of humor and metaphor. Figure 1.2 presents an example of how analogy relates to humor. The cartoon in general and analogy in particular but left them without any satisfactory answer:

> Their accounts are highly valuable as descriptions of the phenomena to be explained. They are even useful as the beginnings of an explanation, for they indicate where to start looking in more detail at the psychological mechanisms involved.—How does bisociation . . . actually work, and how are novel analogies recognized? . . . The more recent theories, referring to everyday abilities and expertise, raise further questions about underlying mechanisms.—How is it that people can notice things they were not even looking for? How can people recognize that two different things (two letters 'a', or two apples) fall into the same class? How is it possible for tacit knowledge to be acquired without being explicitly taught, and how can it aid creation?

These are among the questions we will be addressing in this book. In the years since Koestler wrote about bisociative thinking, the scientific discipline of cognitive science has greatly increased understanding of all of the cognitive processes that give rise to human intelligence—perceiving the world, forming and retrieving memories, and expressing thoughts in language. Cognitive science is based on the conviction that answers to questions like those raised by Boden will require integrating the insights of many different approaches to cognition. Our focus will be on what cognitive science has learned about analogical thinking. We will be presenting a very general theory of how a mind can use analogy as a way to extend knowledge in everyday and creative thinking. We call it a *multiconstraint theory,* because we understand analogy in terms of three distinct but interrelated types of constraints, involving similarity, structure, and purpose. This theory is based not only on our own work but on that of many other cognitive scientists. We will be making use of insights provided by cognitive, comparative, developmental, and social psychologists; by philosophers, linguists, and anthropologists; and by researchers in artificial intelligence.

The use of analogy typically involves several steps. Often a problem solver will select a source analog by retrieving information about it from memory (*selection*), map the source to the target and thereby generate inferences about the target (*mapping*), evaluate and adapt these inferences to take account of unique aspects of the target (*evaluation),* and finally learn something more general from the success or failure of the analogy (*learning*). Our theory is intended to apply to all four of these steps. We will see that the constraints of similarity, structure, and purpose apply to each step, but with varying degrees of relative importance.

One of the major ways in which cognitive scientists validate their theories is to see if these theories can guide the construction of computer programs that can show hints of humanlike intelligence. The successes and shortcomings of such simulations of cognitive processes can tell us a great deal about the extent to which our theories are really adequate. A computer simulation is judged successful to the extent that it behaves in ways that mimic human thinking, with its weaknesses as well as its strengths. We are not looking for a superhuman computerized analogical thinker, but rather for simulations that succeed where people tend to succeed and fail where people tend to fail, and that succeed and fail in the same ways as people do. A successful simulation is intended to be a kind of analog of the analogist.

We eventually describe some computer simulations based on our multiconstraint theory, programs that actually find and use analogies to build explanations and solve problems. We also describe other computer models that make use of analogies, a topic that in artificial intelligence is often called *case-based reasoning* (that is, reasoning from past cases when trying to solve a novel problem). However, our focus will not be on the details of particular computer simulations, but rather on the general underlying principles that begin to answer some of Koestler's and Boden's questions about analogical thinking. By understanding how we think using analogy, we can gain some insights into the creative process as well as into the ways we can be informed and misled by analogies in everyday life.

Overview

We develop our theory further in chapter 2 by describing the nature of the constraints of similarity, structure, and purpose. We argue that the human ability to find analogical correspondences is intimately linked to the evolutionary development of the capacity for explicit, systematic thinking. In chapter 3 we look more closely at the evolutionary origins of analogical thought by considering the use of analogies by animals, particularly the chimpanzee. Chapter 4 describes the results of recent research on the development of analogical thinking in children. By examining in some detail when different varieties of analogical thinking emerge, both in evolution and in human development, we hope to gain a better understanding of the constraints that govern the use of analogy by human adults. In chapter 5 we consider psychological studies of analogical thinking in adults that provide a more detailed picture of how analogy operates to make inferences and solve problems.

Chapter 6 discusses how analogies are used to make decisions about courses of action, especially in politics and law. The role of analogy in developing explanations is the focus of chapter 7. We describe how analogy has been used to help address some of the deepest questions that have confronted humanity, such as the nature of God and human minds. Chapter 8 moves on to discuss the explanatory use of analogies in the development of science and technology. We also consider the educational uses and misuses of explanatory analogies. A major concern in chapters 6 through 8 is the normative question of how analogies can be used most effectively, through mindfulness of how the constraints of structure and purpose contribute to the coherence of decisions and explanations. Chapter 9 considers the role of analogy in literature and in oral cultural traditions, highlighting communicative functions that go beyond explanation and problem solving. We discuss several overlapping cultural purposes of analogy: promoting social cohesion, allowing indirect communication, evoking amusement and other emotions, and helping with social and psychological problems. Chapter 9 includes a discussion of the relation between analogy and metaphor.

In chapter 10 we fulfill our promise to provide a more detailed account of how the constraints of similarity, structure, and purpose guide analogy through the stages of selection, mapping, evaluation, and learning. We describe computer programs developed by ourselves and others that suggest mechanisms by which the complex processes of analogical thinking may be performed by computers and, if the simulations provide successful analogies for the mind, by humans. Finally, we summarize our discussion and point to what seem to us to be the major areas in which understanding of analogical thinking needs to be extended and deepened.

2

Breaking Loose

Funes in Chains

To see a tree as a bird's backyard, to imagine an echo as a wave rebounding off the shore—these must be counted among the intellectual achievements unique to the human species. What is it about the human mind that allows us to think about analogies, to treat situations that we know to be different as if they were somehow the same? When do these capacities develop in children? To what extent are they anticipated in the mental capacities found in other species? In this chapter we prepare to answer these questions by describing the basic requirements for abstract thinking, the capacities that a mind must have to be able to break loose from sensory experience and create complex models of the world.

The essential requirement for analogical thinking is the ability to look at specific situations and somehow pull out abstract patterns that may also be found in superficially different situations. The reason sound is analogous to water waves is that sound exhibits a characteristic pattern of behavior that corresponds to the behavior of water waves: propagating across space with diminishing intensity, passing around small barriers, rebounding off of large barriers, and so on. The surface elements are very different, but the underlying pattern of relations among the elements is similar. Such abstract patterns of similar relations could not be detected by an organism whose every experience was inextricably tied to vivid sensations of sight, sound, smell, taste, and touch. It is difficult to conceive what the world would seem like to such a creature, but the writer Jorge Luis Borges conveys the flavor in his story *Funes the Memorious*. The fictional character Ireneo Funes is described as having a prodigious memory for every detail of his everyday experiences, for he "remembered not only every leaf of every tree of every wood, but also every one of the times he had perceived or imagined it." But his memory for specific details carried a heavy cost: he was incapable of abstract thought. "Not only was it difficult for him to comprehend that the generic symbol "dog" embraces so many unlike individuals of diverse size and form; it bothered him that the dog at three fourteen (seen from the side) should have the same name as the dog at three fifteen (seen from the front)." For Funes, every situation was perceived in such excruciating detail that any cross-situational commonalities were totally obscured. His mind was shackled by the power of specific sensory experience, unable to break free to see abstract patterns of similarity that we normally take for granted. Instead of enhancing his thinking power, Funes's amazing memory abilities seemed to be cognitive chains, binding him to the particular.

As Funes's limitations of thought suggest, the capacity to see analogies has much in common with the capacity to form and use concepts that represent categories of objects, events, and situations. Funes was unable to see how different events are related to one another, a capacity required for everyday categorization. Consider what it takes to see that Hercules the Great Dane and Fifi the Chihuahua are both members of the category *dog*. This judgment is based on the recognition of some important similarities between two individual objects that are nonetheless seen as distinct from one another. Hercules and Fifi are two very different creatures, only one of which has a growl that is likely to discourage intruders. Despite such differences we know that the two have fundamental biological similarities that allow us to make inferences about their hidden properties. For example, suppose you were told that Hercules has an internal structure called an "omentum" that you had never before heard of. You would then be quite likely to infer that Fifi has an omentum too. The reason is that we believe that internal biological structures are generally common across all members of the same species; so if one dog has an omentum, they probably all do. Our knowledge of biological categories encourages the inference that this internal structure will also be found in animals such as cats and other mammals that are similar to dogs. (As a matter of fact, you have one yourself: an omentum is the thin membrane that holds the intestinal organs in place.) Biological concepts such as *dog* provide powerful tools for thinking, because they guide inferences that go beyond direct observation.

From Reacting to Thinking

The ability to form concepts and think about them, which is present in humans and to a lesser degree in other mammalian species, marked a consciousness and therefore more readily verbalized by a thinker who has acquired language. When we are conscious of an explicit representation it is because it is active in our working memory—the memory that holds the information we are thinking about at any given moment. Working memory is limited in the amount of information it can handle at once, unlike our long-term memory that can hold an indefinitely large store of knowledge accumulated over a lifetime. Using explicit knowledge often requires noticeable mental effort, whereas using implicit knowledge is generally unconscious and relatively effortless.

If using explicit knowledge is hard work and sometimes more error-prone than using implicit knowledge, why bother? Explicit thinking was a radical advance in intelligence, because it brought with it great flexibility in the way knowledge can be used. Instead of simply reacting to events in the here and now, we can form representations of events that can be stored in long-term memory and later retrieved and used in a different context. Explicit representations can serve as the inputs to procedures that analyze them, reorganize them, and recombine them to form new ideas. And most importantly for us, explicit representations can be compared to one another to discover systematic similarities and differences between them. Some of these similarities and differences may be far removed from the direct products of perception. Analogical thinking, applied to explicit representations of knowledge, can find important commonalities between situations that on the surface appear quite different. Such abstract similarities often provide a basis for intelligent transfer of knowledge—that is, using prior knowledge in new situations—in order to achieve goals. As a result, the ability to use analogy conveys an evolutionary advantage. And the chains of Funes are loosened.

Similarity

To understand explicit representations of knowledge we need to consider how concepts can be formed. To create concepts one needs to be able to detect similarities between situations despite their differences. The very notion of similarity poses many problems, as the philosopher Nelson Goodman has forcefully argued. After all, any two things have indefinitely many common properties, most of which are likely to be trivial; for example, this book and your shoes are both over a million miles from Venus. The hard problem for cognitive science is to explain why particular similarities are used to form concepts, and how concepts in turn alter judgments of similarity. We will not address this deep problem here in any general way, but we cannot avoid examining the basic forms of similarity that guide our making of analogies.

When we recognize that two objects are both apples, this judgment is aided by many shared sensory features, such as their having the same or similar color, shape, taste, and smell. For many purposes it is sufficient simply to react to similar objects in the same way on the basis of their global similarity. If one apple proves tasty, we are likely to expect another similar one also to be good to eat, without necessarily pausing to reflect on exactly which sensory features determine how apples taste. However, the human capacity for analogy depends on more than simply reacting to objects on the basis of some fuzzy sense of overall similarity based on sensory features. In addition, analogy requires concepts that go well beyond immediate perception. We can think about individuals such as Hercules and Fifi, and of categories such as *dog* and *person*. We can think about actions such as *chasing* and more static relations such as *being larger than*. Our store of such concepts forms a vast *semantic network*. Figure 2.1 depicts a fragment of a semantic network that shows a few of the interconnections between everyday concepts stored in our long-term memory system. Many concepts form a chain of increasing generality, ranging from specific individuals through increasingly general categories (e.g., Hercules → Great Dane → dog → mammal → animal → living thing → physical object. Each concept is a *superordinate* of the more specific concepts in this chain. For example, "animal" is a superordinate of "dog." Besides chains of superordinates, the semantic network is interconnected by a number of other very common relations, such as *part of* (e.g., a tail is part of a dog), as well as by more specialized relations, such as *chases* (e.g., dogs chase cats). The semantic connections between concepts provide important building blocks for seeing analogies.

Concepts are needed to understand sophisticated analogies, but analogical thinking also provides part of the answer to the question of how concepts are formed. The first person who noticed that sound behaves something like water waves presumably did not already conceive of *wave* as a category so general as to include both water and sound waves. But seeing the analogy may have paved the way for forming such a category. And once a more abstract concept of wave was established, it played a role in the further extension to light waves. Thus analogy helps to form new and more abstract concepts, which in turn help to see even more remote analogies, which in turn help to form yet more abstract concepts. We will refer to the representations of complex con-

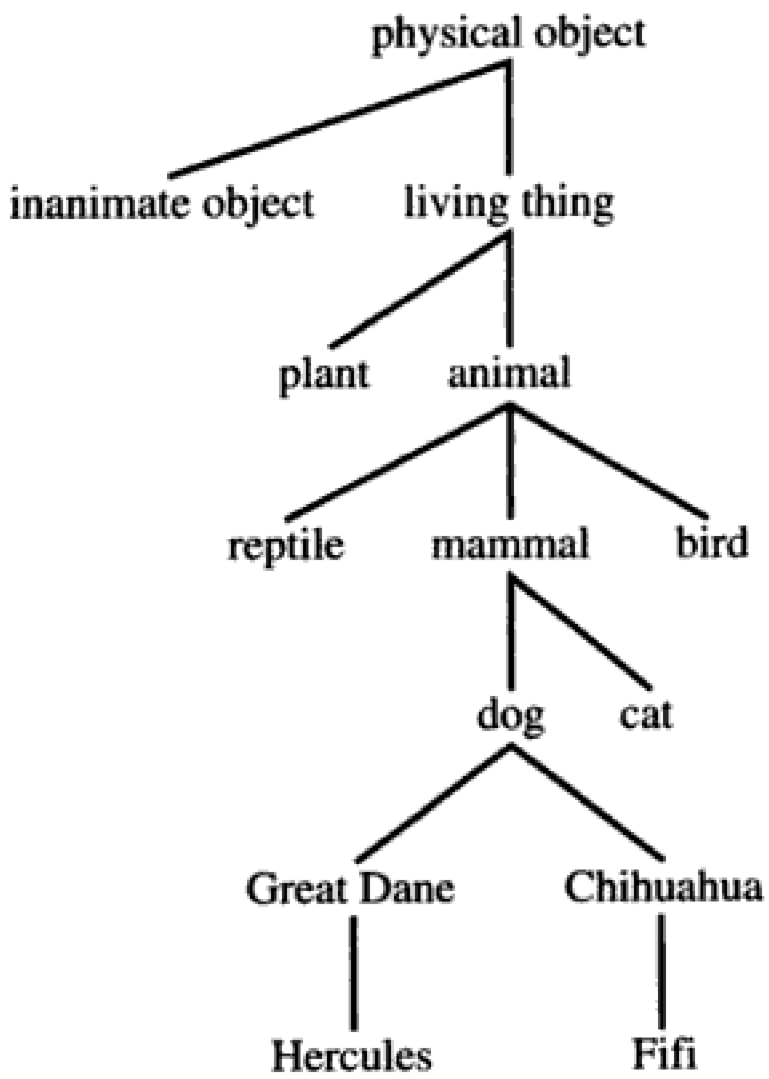

Figure 2.1
Fragment of a semantic network, illustrating superordinate chains. The lines indicate, for example, that a dog is a kind of mammal.

cepts such as *wave,* which convey patterns of relations among constituent elements, as *schemas*. For example, the schema for *wave* will represent the fact that wave motion involves the systematic propagation of a form across space, a form that bends around barriers, and so on. This symbiotic interaction between analogies based on small numbers (often pairs) of cases and general schemas for concepts that embrace indefinitely many cases provides direction for many aspects of thinking, including (as we will see in chapter 9) the ability to understand metaphor.

Structure

Although the individual concepts in a person's semantic network are important for thought, the full power of human thinking depends on its capacity to combine concepts to create more complex structures. For example, we can combine representations for Hercules the Great Dane and for the concept of brown to form the proposition that Hercules is brown. This proposition is expressed by the sentence "Hercules is brown," in which the predicate "brown" is applied to the subject "Hercules." Such a proposition can be true, or false if Hercules is some other color; and we can believe it or doubt it depending on the evidence

This example can be construed as the basis for an analogy of the very simplest form: we might say that Hercules maps to the Sahara by virtue of a corresponding attribute of brownness, indicated by having the predicate "brown" in the expression of both propositions. That is, we have the mappings

brown ↔ brown

Hercules ↔ Sahara.

This analogy is not very interesting or informative, but it does show how the simplest sort of structure can give rise to analogical connections. A slightly more creative connection would arise if we started with a representation of the Sahara Desert as tan, requiring us to make a slightly greater semantic leap with the mapping

brown ↔ tan.

If all we know about Hercules is that he is a brown dog, any color attributed to the Sahara Desert would be more likely to support a mapping of Hercules and Sahara than would other sorts of properties, such as its being dry or sandy. Brown and tan are linked in our semantic network by both being kinds of colors; they also have a more immediate perceptual link, since we know from sensory experience that they look alike. Thus similarity of concepts comes in degrees.

Notice that in this example we are mapping just one pair of objects considered in isolation from any other objects, which can be done on the basis of the semantic similarity between the attributes that apply to each object in the pair. We will call this type of mapping that takes one pair of objects at a time an *attribute mapping.* Even with such simple mappings based on propositions that concern single objects, we see a way out of the trap whereby Funes was overpowered by the uniqueness of every perceptual experience. With propositional thought, one can still appreciate the richness and diversity of experience yet nonetheless be able to extract and make explicit the similarities that connect distinct situations to one another.

Propositions can be more complex than the examples we have so far considered. Greater structure and power for abstract thought is provided by concepts that have more than one slot to fill, such as *chase,* which requires slots both for the chaser and the chased. Predicates with multiple slots are called *relations* for the obvious reason that they relate the multiple fillers of their slots. The proposition that Hercules chases Fifi can be expressed as chase (Hercules, Fifi).

Here the chaser slot is filled by Hercules, and the chased slot is filled by Fifi, in line with the general structure

chase (<chaser>, <chased>).

Imagine a creature that could form representations of Hercules, Fifi, and chasing, but that could not organize them into propositions structured by slots. Such a creature could not distinguish the proposition that Hercules chases Fifi from the proposition that Fifi chases Hercules. All it would have is an unstructured mental stew of the representations of Fifi, Hercules, and chasing. Fortunately, human thinking uses relations that impose structure on the fillers of their slots, enabling us to have the distinct thought

chase (Fifi, Hercules).

That is, we have the capacity to form new propositions by rearranging the fillers of different slots in a concept.

Having propositions with multiple slots has important psychological implications for our capacity to detect and express similarities between situations. For example, with the two-slot relation *larger than* we can express the propositions:

larger-than (Sahara, Hercules)

larger-than (truck, apple).

From these two propositions we can conclude that the Sahara is related to Hercules in a way that parallels how a truck is related to an apple—in both cases, the former is larger than the latter. Expressed as an analogy, this means that Sahara maps to truck and Hercules maps to apple by virtue of their corresponding roles as slot fillers of the relation *larger than*. We have the mappings:

larger-than ↔ larger-than

Sahara ↔ truck

Hercules ↔ apple.

The two propositions correspond by virtue of the semantic similarity (here, identity) between the relation in each proposition. Although the relation is mapped on the basis of semantic similarity, the other correspondences are not based on any direct similarity between the Sahara and the truck or between Hercules and the apple. These objects are connected only by virtue of filling the same slot in the structure of the or more slots to be filled by propositions have been termed *higher-order relations* by cognitive psychologist Dedre Gentner. In contrast, less abstract relations with slots filled by objects, such as *chase,* are called *first-order* relations. Gentner has emphasized the powerful role of higher-order relations in analogical mapping, as sets of propositions interrelated by higher-order relations can be used to help identify correspondences between two analogous structures. In the movie *The Fugitive,* a detective chases the hero, who runs away:

chase (detective, hero) *name*: chase-2

run (hero) *name*: run-2

cause (chase-2, run-2) *name*: cause-2.

Mapping this structure to the preceding one requires putting in correspondence Hercules and detective, Fifi and hero, and the two causal relations between chasing and running.

Ideally, mappings of all sorts should satisfy the two fundamental structural properties of being *one-to-one* and *structurally consistent.* A mapping is one-to-one if every element in the source (object, attribute, or relation) corresponds to a unique element in the target and vice versa. Mapping Hercules to both the detective and the hero would violate this constraint. A mapping is structurally consistent if whenever two relations are mapped their slot fillers are also mapped. In the above example, mapping *chase* to *chase* requires mapping Hercules to detective, because they fill the first slot in the relations. Mapping Hercules to hero or Fifi to detective would be a violation of structural consistency. In our multiconstraint theory, one-to-one mapping and structural consistency are viewed as *soft* constraints, rather than *hard* ones that a mapping must satisfy to contribute to an analogy. That is, satisfying these constraints is highly desirable but not essential. When they are both satisfied, the mapping is an *isomorphism*.

Often an analogy does not simply involve putting known fillers into correspondence, but in addition requires finding new fillers to fit into the whole pattern. If you were told that the plot of *The Fugitive* is like that of the book *Les Misérables,* in which a police inspector pursues a convict, you could infer that it is the convict who runs away. You would merely be filling in the question mark in this structure:

chase (inspector, convict) *name*: chase-3

run (?) *name*: run-3

cause (chase-3, run-3) *name*: cause-3.

The basic device for generating inferences by analogy is called "copying with substitution," because it essentially consists of simply copying over propositions known to be true of the source to become inferences about the target, at the same time substituting target elements for source elements in accord with established mappings (such as **hero ↔ convict**).

Inferences made by analogy using copying with substitution are never guaranteed to be true. The point to remember is that analogy is a source of plausible conjectures, not guaranteed conclusions. Our claim is simply that *if* two domains are in fact isomorphic in terms of how the objects in each are related to one another, then inferences derived from the source will have some plausibility for the target. In actual practice, the best one can do is select a source analog with sufficient correspondences to the target to justify a tentative mapping, use these correspondences to generate inferences about the target, and then check whether these inferences actually hold up when the target domain is directly investigated.

Indeed, an interesting trade-off emerges between the completeness of an analogy and its usefulness in generating inferences. The more complete the initial correspondences are between source and target, the more confident you can be that the two are in fact isomorphic. But unless you know more about one analog than the other—in other words, unless the initial correspondences between source and target propositions are incomplete—the mapping will not allow any new inferences to be made. A complete isomorphism has nothing to be filled in, leaving no possibility for creative leaps. Incompleteness may well weaken confidence in the overall mapping, but it also provides the opportunity for using the source to generate a plausible (but fallible) inference about the target.

Sensitivity to the structural constraint of isomorphism, coupled with the ability to think explicitly about higher-order relations, provides the basis for moving to a deeper level of analogical thinking. So far the source and target analogs used in our examples have always included similar or identical first-order relations (even though the objects being mapped, such as Hercules and detective, were often dissimilar). However, if the analogs provide a richly interconnected system of relations, it is possible to map them even in the absence of similar first-order relations. For example, suppose the source analog is one of Aesop's fables, the "sour grapes" story in which a fox wants and fails to reach grapes and then says the grapes were sour anyway. Hungering for the grapes caused the fox to try to reach them, and failing to reach them caused him to say that they were sour. We can recognize the analogy between the sour grapes story and the case of a person who tries and fails to get a promotion and in the aftermath concludes the job would have been boring anyway. Desiring the job caused the person to apply for it, and being turned down caused the person to say the job was boring. In this analogy there is little similarity between the corresponding objects, such as grapes and job, and only a modest degree of similarity between the corresponding first-order relations, such as *hunger for* and *desire*. The major similarities involve higher-order relations, most notably *cause*. In addition, the objects and relations are highly interconnected, and we can see that each element in the source maps consistently and uniquely to an element in the target. We will use the term *system mapping* to refer to mappings based on similar higher-order relations coupled with a high degree of one-to-one mapping and structural consistency. In a system mapping like the analogy between the "sour grapes" fable and the case of the disgruntled job seeker, the simpler levels of attribute and relational mapping play secondary roles. Of course many analogies (such as that between the plots of *The Fugitive* and *Les Misérables*) involve similarities at multiple levels of abstraction and hence are based on a mixture of attribute, relational, and system mapping. A mind able to use attribute, relational, and system mappings and which can form analogical inferences by substituting as well as copying is free of the chains that bound Funes to sensory experience.

Models

The causal relations in the sour grapes and job-seeker stories show them to be isomorphic at that level. Isomorphism is not simply an esoteric idea used to describe how analogies work. In fact, it is fundamental to the way we understand the world. Consider an ordinary thermometer, such as one that might hang outside a house. How is it actually related to the temperature of the air? After all, the column of red mercury certainly does not look like air. It is not even the case that, say, one centimeter of mercury is equivalent to ten degrees: thermometers can come in a range of different sizes, so there is no general equivalence between any particular length of the mercury column and any particular temperature. The answer is simply that the relations between different heights of the mercury in the tube are isomorphic to the relations between different change should similarly have something in common with what occurs in the brain when an image of a cat running turns into an image of a cat resting.

Analogy takes us a step beyond ordinary mental models. A mental model is a representation of some part of the environment. For example, our knowledge of water provides us with a kind of internal model of how it moves. Similarly, our knowledge of sound provides us with a kind of model of how sound is transmitted through the air. Each of these mental models links an internal representation to external reality. But when we consider the analogy between water waves and sound propagation, we are trying to build an isomorphism between two internal models. Implicitly, we are acting as if our model of water waves can be used to modify and improve our model of sound. The final validation of the attempt must be to examine whether by using the analogy we can better understand how sound behaves in the external world. The basic analogical comparison, however, is not between an internal model and the world but between two different internal models. We will see in chapter 4 that analogical thinking poses special challenges for very young children, because a source analog is both a model of part of the world and the basis for improving another model of a different part of the world, the target.

The idea that analogy is based on finding isomorphisms takes us beyond one of the most influential notions about how prior knowledge is able to influence how we understand new situations. Early in the twentieth century, the psychologist Edward Thorndike argued that transfer of learning depends on the two situations' involving identical elements. Exactly what could count as an identical element was never made very clear, leading to considerable confusion over the years. The idea has often been interpreted, however, as meaning that transfer is only possible when the source and target situations share very specific perceptual properties, such as color, shape, and size. On this interpretation, without such shared properties there is no basis at all for linking the two situations. The view that transfer requires highly specific identical elements has had considerable influence on educational practice, suggesting that what students learn in school will only be useful in tasks that are very similar to those used in teaching the skill. Particularly in our complex modern world, in which most people have to adapt to new job requirements over the course of their careers, the requirement of specific identical elements leads to a rather pessimistic conclusion about the potential usefulness in later life of formal education. But two analogs can be isomorphic even though they do not share any identical (or even particularly similar) elements. In such cases, what is identical across the two structures is not their objects or predicates but rather the overall system of correspondences—the pattern of relations.

We can therefore imagine how transfer might be effected from a source to a target in the absence of identical elements. In later chapters we will explore whether and how this potential is actually realized when people face novel problems. Of course, identical elements can be useful when available even if transfer does not require them. In fact, transfer by identical elements is closely related to the constraint of similarity that we have discussed—the pressure to map elements that are perceptually or semantically similar. And semantic similarity can be based not just on similarity of objects and first-order relations but also on similarity of higher-order relations. Higher-order similarity of relations between relations does not require any identical elements at the perceptual level. The most creative use of analogies depends on both noticing higher-order similarities and being able to map isomorphic systems of relations. These constraints make it possible to map elements that are highly dissimilar, perhaps drawn from very different knowledge domains. We can map elements despite the fact that they are dissimilar in many ways, based largely on the constraint of structure. Because the elements to be mapped will have many differences that must be ignored, this kind of use of analogy is difficult. Nonetheless, it provides the possibility of such creative leaps as recognition of the parallels between motion of sound and motion of water.

Purpose and Cause

Why do people use analogies? The sound/water example is a case of the explanatory use of analogy in which a source is used to explain or develop a new hypothesis about the target. Analogies are also useful for problem solving and planning, when the emphasis is on deciding what to do to accomplish some practical goal, such as finishing an algebra assignment or arranging a vacation. In politics, law, and everyday life, analogies are often used to construct arguments intended to persuade others to adopt a particular course of action. Literary analogies and metaphors can be used to evoke emotional responses, as when Shakespeare compared his lover to a summer's day. In later chapters we will explore all these uses of analogy in much greater detail.

In real-life uses of analogy, the purpose and context of the thinker play an important role in constraining how analogy is used. We would

Causal relations are among the earliest that children are able to recognize, at least at an implicit level. Even six-month-old infants can perceive and react to basic physical causal connections, such as the launching of a billiard ball into motion when it is struck by another. By age two or so, children are able to explicitly represent and talk about particular physical causes, such as cutting and melting, and psychological causes, such as their brother's being mean. We will see in chapter 4 that tasks involving analogies based on well-understood causal relations provide evidence that preschool children are already capable of relational mapping.

However, to think explicitly about causal relations in general, as opposed to particular causal relations, it is necessary to make use of higher-order relations based on causal concepts. It is not enough to be able to represent specific causal regularities, such as the fact that a knife can cut bread; we must also be able to represent more general and abstract regularities, such as the fact that a person can use an instrument to cause a change in an object. A number of concepts are commonly used by older children and adults to describe how one situation is causally connected to another. These higher-order relations include cause, explain, imply, entail, presuppose, facilitate, hinder, and prevent. A thinker who is capable of forming such higher-order relations will be able to use explicit knowledge about causal relations to guide analogical mapping at the system level. Such a thinker can recognize that a secretary intervening quickly to correct the boss's error can be viewed as analogous to a person catching a ball. Despite the lack of similarities between the objects and actions involved in the two situations, they are analogous by virtue of systematic correspondences between patterns of higher-order causal relations.

System mappings bring with them a deeper role for the constraint of structure, along with sensitivity to more abstract similarities of higher-order relations; but they also provide a stronger basis for the constraint of purpose. Once higher-order causal relations can be explicitly represented, they provide a much more direct way of evaluating whether an analogy serves to achieve one's goals. In later chapters we will see that in politics, law, philosophy, science, education, and literature, the same question arises: How can a thinker use analogies productively? Of course, we offer no guarantees, but our general answer is: Work with analogies based on system mappings that can be explicitly evaluated with respect to purpose as well as similarity and structure. In later chapters we will show how this advice can be put into practice.

We can now see how each of our three general constraints on analogy can be used to support an analogy. First, the mapping between

MENTAL LEAPS

ANALOGY IN CREATIVE THOUGHT

Keith J. Holyoak and Paul Thagard

"At once a contribution to the continuing scholarly debate on the nature of analogy and an accessible interdisciplinary overview for the general reader of this central process of thinking and communication."

David Lorimer, *Times Higher Education Supplement*

Analogy—recalling familiar past situations to deal with novel ones—can provide invaluable creative insights, but it can also lead to dangerous errors. In *Mental Leaps,* two leading cognitive scientists show how analogy works and how it can be used most effectively. Keith Holyoak and Paul Thagard provide a unified, comprehensive account of the diverse operations and applications of analogy, including problem solving, decision making, explanation, and communication. They present their own theory of analogy, considering its implications for cognitive science in general, and survey examples from many other domains.

Keith J. Holyoak is Professor of Psychology at the University of California, Los Angeles. Paul Thagard is Professor of Philosophy and Adjunct Professor of Psychology and Computer Science at the University of Waterloo.

A Bradford Book

The MIT Press
Massachusetts Institute of Technology
Cambridge, Massachusetts 02142
HOLNP 0-262-58144-2

90000
9 780262 581448