Mental Leaps: Analogy in Creative Thought

Keith J. Holyoak and Paul Thagard

The MIT Press

From The MIT Press

MITCogNet

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 0-262-08233-0 (HB), 0-262-58144-2 (PB)
1. Analogy. 2. Creative thinking. I. Thagard, Paul. II. Title.
BD190.H64 1995
169—dc20
94-22734
CIP

3

The Analogical Ape

Sameness

The first steps toward analogical thinking require the recognition that two different things can be treated as the same. But Funes was right: the dog at three fifteen has surely changed from the dog at three fourteen. A little older, a new body position—why *should* the two be lumped together under a single shared name? Nonetheless, most of us would cheerfully accept that it was the same dog. Nor do we stop there, for we consider Hercules the Great Dane and Fifi the Chihuahua to be somehow the same, too. At least, two dogs are much more the same than are a dog and a grapefruit.

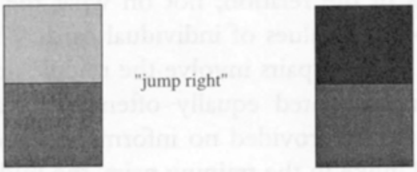
There is nothing specially human about such reactions to perceived sameness. As far as we know, all other vertebrates (and probably many invertebrates as well) are capable of recognizing the general physical similarity of objects. Of course, nonhuman animals are not able to state their views directly on such matters; in fact, except for mammals there is no evidence to suggest that animals have explicit knowledge of similarity relations. But all vertebrates have implicit knowledge of similarity and can make use of it to react adaptively to their environments. If we observe carefully how animals transfer learned behaviors from one situation to another, we can see how they divide up the world. If a bird snacks once on a noxious monarch butterfly, it will likely avoid any further encounters with insects of that species as well as with harmless viceroy butterflies, a species that has evolved to mimic the appearance of the inedible monarch. Meanwhile the bird will continue to ingest other types of butterflies that look less similar to the one that offended its palate. The bird's pattern of prey selection defines the range of butterflies that it implicitly treats as the same with respect to its goal of getting acceptable meals.

A bird is thus able to react in the same way to objects that share perceptual properties. Furthermore, the perceptual basis for the response may be quite subtle. For example, pigeons can be trained to peck a key for food in response to photographs, taken from different angles, of different people in a wide range of poses. In order to recognize that a photograph includes a person (or a tree or a bird or an example of various other natural classes that pigeons can learn), the pigeon must be attending to complex combinations of features. Even for reactions based on physical similarity, psychological "sameness" can be quite far removed from literal physical identity.

To understand the origins of analogy, however, we have to move beyond implicit reactions to similarities between objects to the evolutionary precursors of explicit thought. Many difficulties arise in discussing the evolution of thinking, because the evidence available is so scant. Thought does not leave fossils behind. Evolution did not follow a simple linear path, and there is no simple rank ordering of animals from less to more intelligent. Various species of birds, for example, display navigational abilities that far exceed those of unaided humans. However, our concern here is not with the evolution of all forms of intelligence, but rather just those forms most related to analogy. After a very brief look at the abilities of pigeons and rats in responding to relational similarities, we will focus on primate species—monkeys, chimpanzees, and humans. Fossil evidence clearly indicates that these species are related and can be ordered by the length of time since each branched off from a common ancestral species. Humans are more closely related to chimpanzees than to monkeys. As we will see, studies of thinking by primates suggest at least a rough sketch of the evolutionary origins of explicit thought.

Analogy depends on sensitivity to relations between objects, and therefore our focus will be on behaviors that appear to reveal such sensitivity. What types of animals can respond to similarity between relations in a way that goes beyond direct similarity of objects? The evidence indicates strongly that all mammals have such capabilities, and birds may as well (although the evidence is more equivocal), whereas fish probably do not. Even in mammals, however, careful tests are required to be sure that the animal is really responding to a relation and not simply to attributes of the individual objects. In 1954, Lawrence and DeRivera performed a classic experiment with rats that demonstrated these animals can indeed respond to similarities between relations. The animals were trained and then tested on a "jumping stand." The rat was placed on the stand facing a pair of gray cards, one above the other. If

A. Training Pairs



B. Transfer Pairs

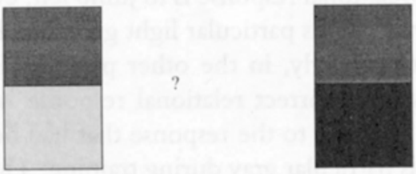


Figure 3.1
Examples of pairs of gray cards used in training (*A*) and in transfer tests (*B*) to investigate relational processing by rats in an experiment by Lawrence and DeRivera (1954).

the top card was lighter than the bottom card, the rat was rewarded with food if it jumped to the right; if the top card was darker than the bottom one the rat was rewarded if it jumped to the left. In other words, what the animal needed to do to get food depended on the brightness relation between the two cards.

The experiment included a transfer test to demonstrate that the rats were really responding to the brightness relation rather than to the brightness values of the individual cards. Each card was one of seven shades of gray. During training the bottom card was always the middle gray, and each of the other shades was used as the top card on different trials. Figure 3.1A illustrates two of the training pairs, one in which the top card is lighter than the middle gray (so the rat needs to go right), and one in which the top card is darker than the middle gray (so the rat needs to go left). The crucial thing to notice is how the absolute values of brightness were associated with reward during training. The light gray always signaled "go right"; the dark gray always signaled "go left"; and the middle gray was ambiguous, as it was always present regardless of whether a jump to the right or the left was rewarded.

On the transfer test, Lawrence and DeRivera presented the rats with new pairs of grays, combinations they had never seen during training.

The most interesting test pairs were ones for which the correct reponse could only be made on the basis of the relation, not on what the rat might have learned about the absolute values of individual cards. Two examples are shown in figure 3.1B. Both pairs involve the middle gray; as we saw, this shade had been associated equally often with each response during training, so by itself it provided no information about which way the rat should jump. Unlike in the training pairs, the middle gray is now the top card, rather than the bottom card. In one pair the light gray appears below the middle gray. Since the top card is darker than the bottom card, the correct relational response is to jump left, even though during training the presence of this particular light gray card had always been a cue to jump right. Similarly, in the other pair the dark gray is below the middle gray, so the correct relational response is to jump right. Again, this is exactly opposite to the response that had been associated with the presence of this particular gray during training. These transfer pairs thus pitted the relational response based on relative brightness against the response associated with the level of absolute brightness. The results revealed that on 74 percent of such trials, the rats jumped in the direction cued by the brightness relation.

Rats do not always respond on the basis of the relation. Other similar experiments have demonstrated that rats can respond both to absolute and to relative brightnesses. What is most important, however, is that rats and other mammals are clearly able to perceive physical relations between two objects and sometimes use these relations as the basis for action. But although rats and other mammals below the level of primates can react to relations between objects, their capacity for relational processing appears to fall short of true relational (or even attribute) mapping. Rats can only respond to a limited number of basic perceptual relations, such as relative brightness or size. Although the animals can react to relations, we lack evidence that they can think about them explicitly. That is, although a rat in Lawrence and DeRivera's experiment could perceive that the top card was lighter than the bottom card and react by jumping right, we cannot assume it was explicitly thinking about *the fact that* the top card was lighter, or even about the fact that the top card was a certain shade of gray. Nonetheless, the rat's accomplishment appears to provide a step toward the capacity for relational mapping. In recognizing that different pairs of cards exhibiting the same brightness relation require the same response, the animal is in some fashion responding to sameness of relations rather than only sameness of objects.

Adult humans can readily understand sameness of both objects and relations, and in English (as well as in all other languages, as far as we know) the same word is used for these different varieties of sameness. To understand the evolution of more abstract types of sameness, we need to make more fine-grained distinctions than ordinary language provides. We will call direct physical similarity of objects *O-sameness,* and we will call similarity of the relations between objects *R-sameness.* For simplicity we will generally ignore the obvious fact that similarity of both objects and relations is a matter of degree; for now it will suffice to divide the scale crudely into the binary values of "same" for high similarity and "different" for low similarity.

Just because an animal is able to respond to O-sameness or R-sameness does not necessarily mean that the animal can explicitly think about these concepts, as we do. For example, a bird may respond the same way to a monarch and a viceroy butterfly without being able to explicitly represent the fact that they are O-same as one another. What would be gained if an animal could explicitly represent O-sameness? Such an animal could not only treat two objects as the same, but it could start to think about the fact that the objects are the same and use this knowledge as the basis for action.

A task that can be used to assess whether animals can perceive O-sameness is illustrated in figure 3.2. As depicted in figure 3.2A, the animal is first shown an object, here an apple, which is called the sample. Then the animal is offered a choice between two objects, one that is O-same as the sample (another apple) and one that is O-different (a shoe, for example). The positions of the two alternatives (left and right) are varied across trials; hence O-sameness is defined in terms of sameness of shape rather than position. If the animal selects the object that is O-same as the sample (regardless of its position), a reward is given. (Alternatively, the animal might receive a reward only if it selects the object that is O-different.) This task is called "match-to-sample" for the obvious reason that it requires the animal to match the choice alternatives to the original sample and to select the alternative that matches the sample appropriately. There may or may not be a delay of a few seconds between presentation of the sample and the alternatives. Delay versions of the task test the animal's ability to maintain a representation of the sample in working memory. Here we concentrate on the simplest case, in which the sample and the alternatives are presented simultaneously.

Even pigeons can learn to respond correctly in the basic match-to-sample task. But as in the case of rats learning to respond to pairs of gray

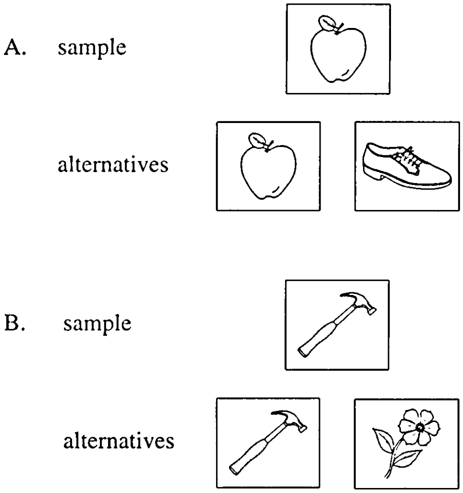


Figure 3.2
(*A*) An example of a match-to-sample problem. (*B*) A transfer problem based on a new set of objects.

cards, there is ambiguity about whether animals are in fact reacting to a relation rather than to the specific objects being related. The simplest possibility is that they might just learn a specific conditional rule, such as

If sample is apple, then pick apple.

The more interesting possibility is that they are learning to respond on the basis of general O-sameness:

If sample is O-same as alternative, then pick that alternative.

The way to test whether the animal is learning a rule based on O-sameness is to see how it behaves on a transfer test, such as that depicted in figure 3.2B. Here the sample is a hammer, and the alternatives for matching are another hammer and a flower. If all the animal had learned was "if apple, pick apple," it would have no basis for responding on this generalized transfer test, since none of the objects have any particular resemblance to apples at all.

On the other hand, suppose an animal is capable of attribute mapping and hence can in at least a crude way think about, rather than simply react to, the category or basic shape of an object. Such an animal could represent the first sample as something like apple (apple-1)

and the apple alternative as

apple (apple-2)

and then react to the O-sameness of the sample and the alternative revealed by their shared attribute. Then on the transfer test the sample would be represented as

hammer (hammer-1)

and the hammer alternative as

hammer (hammer-2).

An animal that had learned to respond on the basis of O-sameness of category or shape would then choose the hammer alternative, because it is O-same as the sample, just as the apple alternative had been O-same as its sample.

Pigeons generally perform poorly on such generalized transfer tests. But for primates, such as monkeys and chimpanzees, the evidence is clear: without any special training or further reward, the animal will transfer what it has learned about matching apples to apples to the new case of matching hammers to hammers. We can therefore be quite certain that primates can react to O-sameness of objects, even though the objects in each pair have no O-sameness to the objects in the other pair. Moreover, if the primate had learned to match an apple to something that was not an apple, then on a transfer test it will match a hammer to something that is not a hammer. Thus the animal is also able to react to the relation of O-difference.

What distinguishes a monkey's performance from that of a pigeon? A pigeon can react to the global similarity of two objects, such as one apple and another. However, it seems to lack the capacity to think explicitly about the physical attributes of objects, as appears to be required in order to perceive a relation that makes the similarity of one apple to another somehow the "same" as the similarity of one hammer to another. In contrast, the primate has evolved to be able to think about attributes of objects and to perceive the relation between sameness of one set of attributes and sameness of another set—that is, O-sameness. The primate can therefore learn to react to O-sameness, rather than to the particular objects that are the same. The result is a major extension in the breadth of transfer across situations.

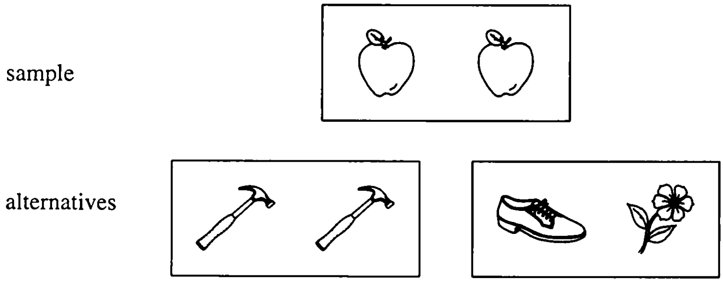


Figure 3.3
An example of a pairwise match-to-sample problem in which the sample pair and the correct alternative are each based on O-same objects.

What Sarah Thinks

Nonetheless, even an intelligent chimpanzee turns out to be surprisingly limited in its ability to respond to O-sameness. Figure 3.3 illustrates what appears to be a simple extension of the generalized match-to-sample problem that primates are able to perform so well. The only difference is that the sample and the alternatives are now not single objects but rather pairs of objects. In figure 3.3, the sample is a pair of apples, and the alternatives are either a pair of hammers or a shoe and a flower. Which would you choose? No doubt it is obvious to you that the pair of hammers is the same as the sample in a way in which the shoe-flower combination is not. At first glance, the problem appears to have the same logical form as the generalized transfer test depicted in figure 3.2 (i.e., transfer from the problem in figure 3.2A to that in figure 3.2B), so you might expect any intelligent primate would also be able to succeed on the pairwise match-to-sample task in figure 3.3.

If so, you will be disappointed in the nearest relatives to our own species. No nonhuman primate—not even an intelligent chimpanzee—has ever consistently been able to succeed on the pairwise match-to-sample test without some very special training that we will describe in a moment. No animal other than a human or a chimpanzee has ever succeeded under any circumstances.

Evidently, the apparently innocent step from the task depicted in figure 3.2 to that depicted in figure 3.3 requires another evolutionary leap. Let us look at the two versions of the matching task more carefully. In the single-object version, generalized transfer requires reacting to O-sameness—if the animal has learned to choose the apple alternative that is O-same as the apple in the sample (figure 3.2A), then it will also choose the hammer alternative that is O-same as the hammer in the sample (figure 3.2B). However, although this performance may require explicit representations of the attributes of objects, it does not require an explicit representation of the relation of O-sameness. That is, although the animal is using O-sameness, it does not know that it is using it. It may be using O-sameness implicitly, without having constructed an explicit representation.

The trouble is that an implicit reaction to O-sameness is not sufficient for success on the pairwise version of the task, because this task does not allow an immediate reaction to O-sameness. Instead, it is necessary first to explicitly represent the fact that the sample consists of one object that is O-same as another, for example,

O-same (apple-1, apple-2).

This representation must be "held in mind" while the animal processes the two sample pairs, which could be represented as

O-same (hammer-1, hammer-2),

O-different (shoe-1, flower-1).

Only after forming such explicit representations could the animal compare the relation in each alternative to that in the sample and select the alternative that is described by the same relation as that used to represent the sample. In other words, the pairwise task cannot be solved by reacting to O-sameness of two objects; rather, it requires the ability to react to R-sameness of two relations. To perceive R-sameness, it is necessary for the animal to have an explicit representation of the relation of O-sameness, which can link pairs of apples, pairs of hammers, pairs of frogs, and so on.

It turns out that although chimpanzees do not ordinarily form explicit representations of O-sameness, they can learn to do so. The first nonhuman animal to solve the pairwise match-to-sample task was an African-born chimpanzee (*Pan troglodyte*) named Sarah. Sarah's life was very different from that of any other chimpanzee who lived before her. She spent nineteen of her first twenty years in a laboratory directed by psychologist David Premack, attending a kind of school five days a week since she was about five years old. Much of her school time was spent studying a form of artificial "language." The "words" of this language were colored plastic tokens in various shapes, sizes, and textures, which Sarah was trained to put together into ordered strings to represent propositions, such as "apple is red," "blue is on yellow," "round shape of apple," and so on. Over the course of training, the strings and corresponding propositions increased in complexity, including, for example, "red on yellow, if then, Sarah take chocolate," and "Sarah take apple in red dish, banana in blue dish." These examples are given in English but with the word order that Sarah learned.

One of Sarah's words is especially important to our story: she learned to build propositions with a token for "same." She could use "same" to relate two objects of the same type, as in "apple same apple," and also to relate two strings that both expressed the same proposition, as in "apple is red, same, red color of apple." If Sarah had in fact acquired an explicit representation of the concept of O-sameness, then perhaps she would be able to succeed in the pairwise match-to-sample task illustrated in figure 3.3.

And succeed she did. Not only did she select the alternative that matched the sample by also exemplifying O-sameness when given a problem like that depicted in figure 3.3, but she also solved problems in which the sample and the favored alternative both exemplified O-difference. Figure 3.4 shows a problem of the latter sort. The sample consists of a bottle and a bell; the alternatives are identical to those in the problem shown in figure 3.3. For the problem in figure 3.4, Sarah would select the shoe-flower pair rather than the pair of hammers. Furthermore, she did not need any special reward in order to make these choices. She simply expressed her preference for whichever alternative exhibited the same relation as that exhibited by the sample.

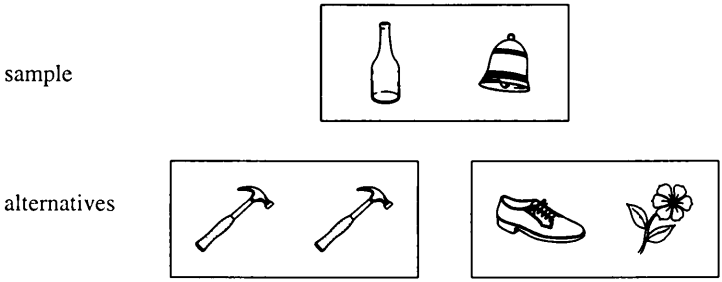


Figure 3.4
An example of a pairwise match-to-sample problem in which the sample pair and the correct alternative are each based on O-different objects.

How did Sarah solve the pairwise match-to-sample problems? The task can be viewed as an analogy problem that can be solved by relational mapping, with the sample serving as the source analog and each alternative serving as a possible target analog. Figure 3.5A illustrates how a sample consisting of two apples maps consistently to an alternative consisting of two hammers, but not to a sample consisting of a shoe and a flower. The two O-same relations are identical, producing a mapping by pairs between the first apple and the first hammer and between the second apple and the second hammer. The hammer-hammer alternative is thus analogous to the apple-apple sample, whereas the shoe-flower alternative is not (because the relation in the latter alternative differs from that in the sample). Figure 3.5B illustrates how the preference will reverse when the sample consists of two different objects, a bell and a bottle. The sample would now be represented by

O-different (bell-1, bottle-1),

and the alternatives would again be

O-same (hammer-1, hammer-2),

O-different (shoe-1, flower-1).

Now it is the shoe-flower pair that exhibits the same relation, O-different, as does the sample, whereas the hammer-hammer pair exhibits a

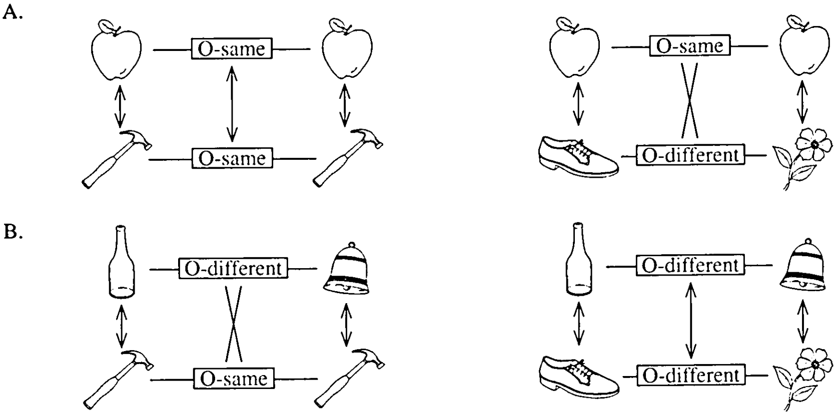


Figure 3.5
Relational mapping in the pairwise match-to-sample task. (*A*) Only the first alternative (*left*) has the same relation, O-same, as the sample. (*B*) Only the second alternative (*right*) has the same relation, O-different, as the sample.

relation different from that of the sample. Relational mapping will therefore favor selection of the shoe-flower pair.

Notice that relational mapping of the sort depicted in figure 3.5 amounts to an implicit response to R-sameness between the relations in the sample and in the preferred alternative. That is, the similarity constraint on relational mapping—the tendency to map similar relations to each other—leads to a preference for the alternative with a relation that is R-same as the relation in the sample. However, we need not assume that Sarah actually formed an explicit concept of R-sameness. As the diagrams in figure 3.5 make clear, it is possible to respond on the basis of the similarity between relations without explicitly representing the higher-order relation of R-sameness. We will see later that the apparent absence of such explicit higher-order relations from the conceptual repertoire of the chimpanzee places bounds on their analogical ability.

The obvious question, given her superior performance on the pairwise version of the matching task, is what made Sarah special? Was she simply smarter than other chimpanzees that failed the pairwise task? This does not seem to be the case. Sarah was certainly bright, but by many measures some of the other chimpanzees that lacked her special training—and failed the pairwise matching task—were also bright. The best evidence that her specialized experience was crucial in some way is provided by further tests that Premack performed using other chimpanzees who were considerably younger than Sarah. Two young language-trained chimpanzees succeeded on the pairwise matching task, whereas four otherwise comparable animals that did not receive language training failed. It is therefore clear that something about the special training with symbol manipulation was responsible. Moreover, the impact of the training on the animals' reasoning abilities was quite selective. For example, the language-trained animals (Sarah included) were no better than the others in tasks that required making inferences about the spatial locations of hidden objects.

Is it necessary for an animal to have a language to solve pairwise match-to-sample tasks? Premack himself has disavowed any claim that what Sarah and the others learned was really comparable to human language. What exactly ought to be considered a language is a highly controversial issue; fortunately, it is not an issue that is relevant here. What is clear is that the training did encourage Sarah and her younger fellow students to use explicit propositional representations to control their actions. And of special importance, the animals were taught to make responses on the basis of an explicit concept of sameness. It appears that Sarah was the first nonhuman animal ever to have acquired an explicit relational concept by learning.

What kinds of knowledge must an animal already have in order for language training to generate an explicit concept of O-sameness? This is a difficult question to answer. However, later work provided evidence that even infant chimps will react in some fashion to R-sameness, without any special training at all. David Oden, Roger Thompson, and Premack adopted a technique that is often used with preverbal human infants to test what they see as the same or different. The technique is based on the fact that chimpanzees, like human infants, get bored more quickly when an experience seems much the same as one they have already undergone recently. Four infant champanzees were presented with a pair of objects, such as two pieces of garden hose (O-same), or a plastic block and a metal bracelet (O-different). The animal being tested was first allowed five minutes to familiarize itself with the sample pair of objects by handling it or otherwise interacting with it. After a fifteen-second interval, the animal received a second pair of objects. These were always different objects from those given on the familiarization trial, but the relation between the objects in the new pair was varied. Examples would be a pair of plastic chain links (O-same), or a bottle cap and a strip of wood (O-different). The experimenters then measured how long the animal handled the new pair of objects over a further five-minute period. The entire procedure was repeated twelve times in each of three four-hour sessions, with a week separating each session.

The results of this experiment are shown in figure 3.6. The graph plots the difference between the average time the animal handled the first and second pair of objects. The higher the bar, the sooner the chimpanzee became bored with the second pair—in other words, the more it seemed like the "same thing" as the first pair. As you can see, the results show that the chimpanzees found the second pair more boring when it was R-same as the first pair. The effect appears more robust when both pairs were O-same, but the basic pattern also held when both pairs were O-different. It is as if the infant chimpanzees found it boring to deal with O-sameness repeatedly or with O-difference repeatedly.

Because infant chimpanzees without any training of the sort Sarah received are able to react to R-sameness, you might suppose that with a bit of direct training they could go on to solve the pairwise matching task. But these animals could not consistently select the alternative that

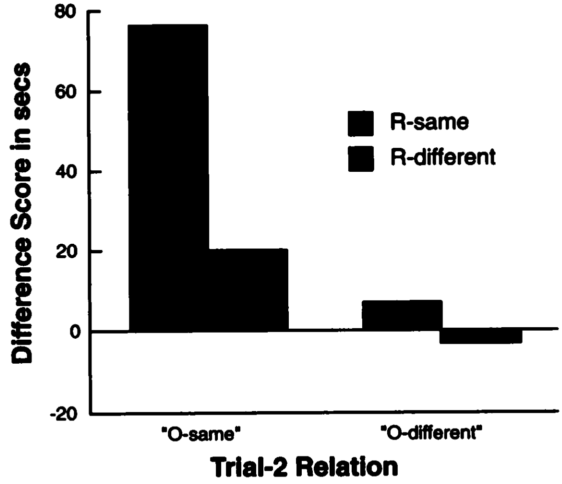


Figure 3.6
Trial 1 minus trial 2 differences in handling times for O-same and O-different pairs that are either familiar (R-same) or novel (R-different). Adapted with permission from Oden et al. (1990).

exhibited the same relation as the sample, even when they were repeatedly rewarded when they chose correctly. (Other evidence suggests, however, that such training can produce success if it is extended for one or more years.) It thus appears that chimpanzees react to a relationship that they are unable to readily use. Their reaction to R-sameness seems to be spontaneous, acquired early in life, and almost certainly not learned from experience. Furthermore, the ability to react to R-sameness sets chimpanzees apart from less intellectually advanced primates. Further experiments with monkeys, using the same procedure that shows chimpanzees react to R-sameness, have failed to find any evidence that monkeys see any similarity between what makes one pair of objects the same and what makes a different pair of objects the same. Human infants, on the other hand, also react to R-sameness, at least by the age of seven months. So chimpanzees, like humans but unlike monkeys, apparently have the inborn capacity to react to R-sameness.

While infant chimpanzees and humans apparently react to R-sameness, do they really perceive this higher-order relation? Not necessarily. It may be possible to explain the infants' pattern of handling pairs of objects solely in terms of implicit reactions based on O-sameness and O-difference. All we need to assume is that these infants implicitly perceive O-sameness and O-difference of objects and that they find repeated perception of either one of these relations to be boring. Although we can describe this pattern as "reacting to R-sameness," it is not clear that the infants are actually even perceiving R-sameness, far less that they are explicitly thinking about it as a concept. Nonetheless, their innate ability to perceive O-sameness and O-difference is undoubtedly important in giving them the potential to acquire explicit concepts of O-sameness and O-difference and then to perceive R-sameness of such relations.

How Thought Evolved

It is worth pausing to reflect on what has happened over the course of the evolution of relational mapping, culminating in Sarah's solution of the pairwise match-to-sample task. To put it in more general terms than apples and hammers, we will use letters to stand for objects of a certain recognizable kind. For example, any example of an apple might be designated by "A," and "A_1" would stand for some particular apple. An animal capable of simply reacting to global physical similarity can match A's to A's and B's to B's and respond accordingly. But although such an animal can respond to all A's in one fashion and to all B's in some other fashion, it does not necessarily perceive any commonality between the bases for these separate judgments. Such an animal is roughly at the mental level of a pigeon: in a match-to-sample task where the sample is an A, it can learn "if A, then pick A," but it will be at a loss as to how to respond if then given a transfer task in which the sample is a B instead of an A.

To be able to generalize to new objects, the animal would need to have the ability to perceive the common basis for matching objects according to a shared category or shape, namely, O-sameness. To perceive O-sameness the animal needs an explicit representation of the attributes of objects. With this additional mental machinery, which is available to primates, the animal will be able to transfer what it has learned about picking an A to match an A—to select the alternative that is O-same as the sample—and apply this knowledge to pick a B to match a B. However, because it is responding to O-sameness on an implicit basis, the animal is limited to making an immediate response to a pair.

Here is where the special training of Sarah and the other language-trained chimpanzees triggered another mental leap, up to the level of relational mapping. To do relational mapping, the animal must first be able to translate correspondences obtained by attribute mapping, such as

$A_1 \leftrightarrow A_2$

and (in a separate attribute mapping)

$B_1 \leftrightarrow B_2$,

into representations that at least approximate propositions expressing *the fact that* the mapped objects are related. Furthermore, the relation used in these propositions must be general enough to apply to different kinds of objects that can be grouped on the basis of attribute similarity. The necessary relation is provided by an explicit concept of O-sameness. Armed with this concept, the animal can recode the above attribute mappings into the propositions

O-same (A_1, A_2)

O-same (B_1, B_2)

for any pair of examples of A's or of B's. An animal that can form such propositions has taken a giant stride toward abstract thought. For now it is prepared to map not simply objects one by one, but objects taken in pairs—mappings driven not by direct physical similarity of objects but by similarity of the relation between the objects. By mapping two propositions of the above sort, an animal like Sarah can arrive at the correspondences

O-same ↔ O-same

$A_1 \leftrightarrow B_1$

$A_2 \leftrightarrow B_2$

based on mapping objects in pairs. These mappings are not justified by direct similarity between A's and B's. Instead, the mapping is justified by the fact that A's and B's can play parallel roles with respect to O-sameness: the A's are O-same as each other just as the B's are O-same as each other. An animal with this much mental equipment is able not only to solve the basic match-to-sample task and perform well on generalized transfer tests with completely different objects, but also to solve the pairwise version of the task. Apples may be freely replaced by hammers, and the problem will still be seen as the same.

The remarkable progression in abstract thought that we have just outlined required several million years of biological evolution. With Sarah's final step, the progression moved from the timescale of evolution to that of learning. Her feat in solving the pairwise version of the match-to-sample task was the mental equivalent of the first flight of the Wright brothers—only a short wobbly hop, but one that we can look back on as the dawn of space travel.

If we look carefully, we can see that the progression from reacting to physical similarity to performing relational mapping is based on a general strategy for deepening the abstraction of thinking, moving beyond direct sensory experience into the realm of concepts. Of course, by calling it a "strategy" we do not mean that it is deliberate or purposeful—it is simply a general description of how abstract thought appears to have evolved. The strategy might be sketched this way:

Step 1: Based on whatever means of mapping elements the animal already has at its disposal, it finds correspondences and identifies sets of elements that map consistently. The earliest basis for mapping was global physical similarity of objects.

Step 2: Explicit concepts are formed to capture the basis of the mappings.

Step 3: The strategy cycles back to step 1. It will now be possible for the altered animal to map elements on the basis of the new concepts that have been formed. Armed with this new way to justify mappings, steps 1 through 3 are repeated.

The full power of the strategy lies in its call to "repeat." What exactly does that mean? The intellectual developments roughly bounded by a pigeon and Sarah appear to have required two cycles through the strategy. In the first cycle, step 2 provided the capacity to represent attributes of objects explicitly. For example, instead of just reacting in the same way to different apples because they look alike, the animal could now think about the basis for the similarity of different apples, expressed as a concept that applies to any apple, as in

apple (apple-1),

in which "apple" serves as a shorthand for those attributes generally shared by apples. Roughly speaking, this evolutionary move takes us from the pigeon to the monkey, which returns from step 3 armed with the ability to perform attribute mapping. In the second cycle, this more sophisticated animal is able to form mappings by reacting to the relation of O-sameness between attributes of objects, which allows it to solve the generalized match-to-sample task with single objects. At step 2, the basis for these mappings is coded as the explicit concept of O-sameness. This move required both evolution (roughly, from the level of the monkey to that of the chimpanzee) and the special training that Sarah received.

Sarah represents an animal that could return from step 3 to step 1 with the novel capacity to perform relational mapping. She can now map elements in pairs based on similarity between different occurrences of her new relational concepts, such as O-sameness, which allows her to solve the pairwise version of the match-to-sample task.

The case of Sarah is especially striking, because for the first time our strategy operated in part by learning within an individual animal, rather than solely by evolution of new species. What would a third cycle of the strategy bring? We have already hinted at how an animal might go beyond Sarah's level of relational mapping. The basis for a relational mapping (step 1 repeated) is the sameness of a relation in the source to a relation in the target—in other words, R-sameness. An animal that could form an explicit concept of R-sameness (step 2 repeated) would have a new and more abstract basis for mapping. Before we consider animals capable of this deeper level of abstraction, let us look more carefully at the thinking of our nearest evolutionary cousins.

The Analogies of Apes

Sarah's training allowed her to become the first nonhuman animal to solve analogy problems in the proportional format used on human intelligence tests. As we mentioned in chapter 2, a proportional analogy has the form A:B::C:D ("A is to B as C is to D," as in "A can opener is related to a can as a key is related to a lock"). The problem solver might have to judge whether A:B is the same as or different than C:D, or to choose the best completion for an analogy from a set of alternatives, as in A:B::C:? in which the answer is the best "D" term. We can treat the A:B pair as the source analog and C:D as the intended target analog. Taking the "can opener" problem as an example, the obvious representation of the source would be

open (can opener, can),

and the representation of the target would be

open (key, lock).

The source and target form a relational mapping based on the R-sameness of the two relations.

The resemblance to the form of the pairwise match-to-sample task should be apparent. In the pairwise matching task, although the problem solver is not directly asked to map the sample to the alternatives, this is in fact how the task can be performed. The analogy format simply makes the requirement explicit. In general, an analogy problem can use any relation to link the objects in the source and in the target ("open" in the above example). In the pairwise matching task, the relation happens to be O-sameness. The higher-order relation of R-sameness, however, is special. In a basic proportional analogy, the justification for mapping the source and target is always that the relation in the source analog is R-same as that in the target, regardless of what specific relation is used in each analog. However, a relational mapping only requires an implicit reaction to R-sameness rather than an explicit representation of the concept.

Analogy problems in which a missing term has to be generated, as in A:B::C:? pose yet another cognitive requirement—the ability to consider a question. The idea of a question is one we take for granted, but like the idea of sameness of relations, it in fact represents a considerable cognitive achievement. Notice that understanding or formulating an explicit question requires representing a missing slot filler in a proposition. For example, the question "What is the color of an apple?" has the logical form

color-of (apple, X?),

in which "X?" represents a missing slot filler. The answer to the question is some definite value, such as red, that can fill the empty slot denoted by "X?" to generate a complete proposition that is true. Similarly, "How old is Neil?" has the form

age-of (Neil, X?),

where the answer is some particular age, say four years, that could replace "X?"

The mental representation of a missing slot filler such as "X?" that holds a slot open for something that might fill it can be termed a *query marker.* Our English words for types of questions—"who," "what," "where," "when," "why," "how"—pick out different types of fillers for the empty slot: "what" typically calls for an object of some sort, "who" for a person, "where" for a place, and so on. Questions can be thought of as incomplete propositions in which a query marker holds open a slot that needs a filler.

The language training Sarah and the other chimpanzees received gave them a great deal of experience in answering questions. Early on, Sarah was taught that a special token acted as a query marker, just as

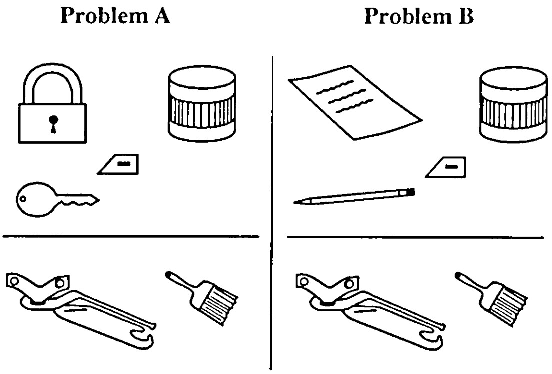


Figure 3.7
Examples of functional analogy problems that Sarah solved. (*A*) When shown a closed lock and a key and a closed painted can, she selected the can opener (rather than the paint brush) as the appropriate completion to make the two relations the "same." (*B*) When shown a marked paper and a pencil, and the same closed painted can, Sarah selected the paint brush as the better completion. From Gillan et al. (1981). Copyright (1981) by the American Psychological Association. Reprinted by permission.

"X?" does in the above examples. If Sarah had already learned the tokens for "color of" and for "apple," and was being taught the token for "red," she would be shown a string of tokens representing "X? color of apple" and receive a reward if she correctly replaced the "X?" with the token for "red." Sarah and the other language-trained chimpanzees caught on quickly as to how the query marker was to be interpreted, suggesting that chimpanzees may have some natural concept of this sort. With a query marker, an animal can start to think about what it does not yet know, not just about what it currently does know. Along with her experience working with a token for "same," Sarah's experience with the question token provided an important way for her to answer analogy questions.

Premack and his colleagues were able to test Sarah on analogies involving other relations besides O-sameness. She proved proficient in solving analogies based on functional relations, in which each analog involves an instrument that can operate in some specific way to change the state of another object. Figure 3.7A shows the key–can opener analogy on which Sarah was tested, along with the pair of alternatives she had to choose between to fill in the missing term in order to make the source and target the "same." (The shape with an equal sign in figure

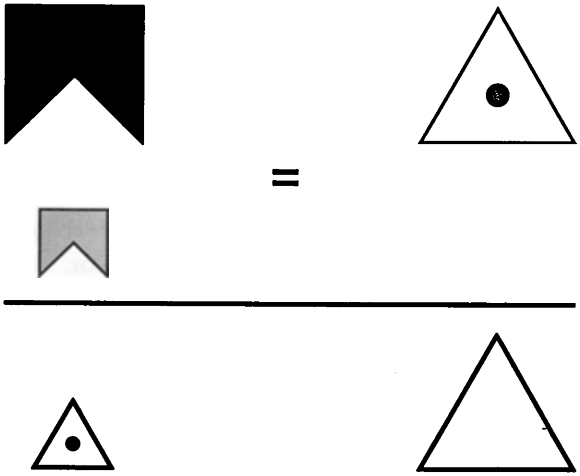


Figure 3.8
Example of a geometric analogy problem that Sarah solved. When shown a large, dark sawtooth and a small, dark sawtooth, and a large, white triangle with a dot, she selected the small, white triangle with a dot (rather than the large, white triangle with no dot). From Gillan et al. (1981). Copyright (1981) by the American Psychological Association. Reprinted by permission.

3.7 represents Sarah's token for "same," which was actually a yellow, plastic rectangle.) The can in the picture was both closed and painted, and the incorrect alternative was an object closely associated with a painted can, namely a paintbrush. Nonetheless, Sarah selected the can opener as the object that best completed the analogy.

Besides the relation of opening, as in the key–can opener analogy, Sarah could map analogs on the basis of "marking." In fact, as the display in figure 3.7B indicates, it was possible to design a marking analogy that reversed Sarah's choice between the same two alternatives. Given pictures of a marked paper and a pencil as the source pair ("pencil marks paper") and a picture of the painted can, she now selected the paintbrush ("brush marks can") as the better completion, rather than the can opener.

Notice that these analogies cannot be answered in any obvious way on the basis of direct physical resemblances. Not only does a can opener not look like a key, nor a can like a lock, but also the motion of opening a can with a can opener actually has very little physical overlap with the motion of opening a lock with a key. The fact that we call both these actions opening—just as Sarah learned a single token that applied in both cases—calls attention to their commonalities. But these commonalities have less to do with physical appearance than with their similar functions: both openings are actions that achieve access to a space that was initially blocked. As we will see shortly, Sarah's problem solving, like her analogy solutions, showed that she was able to represent the purposes of actions.

Using various matching tasks, Premack and his colleagues were able to show that Sarah could solve other analogy problems. These included geometric analogies, such as that depicted in figure 3.8. These problems were based on arbitrary forms that could vary in shape, color, and marking (either unmarked or with a black dot). In this example, when shown a large dark sawtooth form and a small dark sawtooth form and a large white triangle with a dot inside, Sarah selected the small white triangle with a dot rather than the large white triangle with no dot. That is, she was able to pick the form that created the relation "larger than" between the two target forms, the relation that was illustrated for the source pair.

One of Sarah's tasks required her to respond on the basis of a mapping between proportional relations. For example, Sarah was shown a "sample" of a half-filled glass cylinder. Then she had to choose between half of an apple or three-quarters of an apple as the better match to the sample. Sarah consistently solved problems of this sort correctly, revealing that she was sensitive to the fact that a half-filled cylinder is relationally similar to half an apple, even though the objects themselves are very dissimilar. Sarah's success on the proportions test indicates that chimpanzees can perceive something akin to part-whole relations.

Another interesting example of Sarah's ability to match on the basis of abstract relational similarities involved same/different judgments between pairs of three-item sequences. The sequences were composed of physically dissimilar nonsense shapes. Each sequence of three was presented twice, and Sarah was required to judge whether the order was the same in the two presentations. In a challenging form of the test, the first presentation consisted of the three objects' being presented one at a time, in serial order, whereas the second presentation consisted of the same three objects' being presented simultaneously in a certain spatial order. To compare the two sequences, Sarah had to map an ordering based on time to an ordering based on space. (On other trials, the first sequence was spatial and the second was temporal.) Although her performance was imperfect on this matching task, it was well above chance. It therefore appears that Sarah was able to recognize that order of objects in time is relationally similar to order of objects in space.

In several types of analogy tests, then, Sarah's language training gave her a qualitative advantage over chimpanzees that lacked language training. Her apparent mastery of explicit concepts for relations appears to be the most important factor that helps explain what Sarah's successes have in common with each other. In later work, Premack and his colleagues trained other chimpanzees in a similar fashion, systematically testing their performance on analogy problems at different steps in the training process. The crucial event appeared to be the introduction of plastic words for "same" and "different." Moreover, it was actually sufficient simply to teach the use of "same" and "different" as they apply to objects (e.g., two apples were called "same," but an apple and a banana were "different"). Without further direct training, the chimpanzees were then able to apply their tokens for "same" and "different" to relations as well as objects. In contrast, the researchers found that neither teaching tokens for words nor teaching sentence-like strings of tokens was sufficient to produce success on an analogy task.

In fact, studies that Premack has performed reveal that chimpanzees will eventually learn to respond on the basis of relational similarity after one to three years of "dogged training," even without being taught tokens for "same" and "different." For example, the animals will eventually learn to match samples on the basis of like proportions (e.g., matching a quarter of an apple to a quarter of a glass of water). Furthermore, unlike arbitrary learned relations, such as "if the light is red, turn right," chimpanzees do not have to be rewarded for making correct match-to-sample choices. Eventually they simply seem to notice, for example, that a certain proportion of one object is similar to that proportion of another object and begin to respond on that basis. But without an explicit name for "same," the animals do not transfer what they learn to new relations. Thus, having learned to match samples of like proportions does not lead to success in forming analogies on the basis of other relations. In contrast, animals who have acquired tokens for "same" and "different" are able to handle any of a wide range of tasks that depend on reacting to sameness of relations. Teaching tokens for "same" and "different" helps the animal to break loose from the immediate training context and solve a much wider variety of relational problems.

We are still left wondering how training with symbolic tokens for "same" and "different" allows explicit relational concepts to provide new mental tools for guiding the behavior of chimpanzees. One possibility that Premack has suggested is that various parts of Sarah's training, which were directed specifically at the generation of propositions based on explicit concepts for relations, caused her to shift greater attention to this more abstract level of description. Sarah and her fellow students spent a great deal of time in situations in which they were encouraged to treat relations expressed in propositions as more important than the particular items filling slots at any particular time. This is a very different regimen from that which confronts an ordinary chimpanzee living in the wild, whose experience is likely to encourage a decidedly more practical bent. In the wild, it is the salient perceptual properties of objects and relations between objects that are usually most important to survival. Consequently, although they also may perceive the deeper relations, untrained chimpanzees presumably generally ignore them. In contrast, Sarah was encouraged to pay less attention to surface properties of objects, and instead to focus on deeper relations. Indeed, Premack found that she sometimes experienced a surprising degree of difficulty solving simpler problems that required attention to more primitive object similarities. There may be an inherent trade-off between attention to abstract and to concrete aspects of mental representations. The focus on the abstract at the expense of the concrete, apparent in the thinking of the first chimpanzee intellectual, may also characterize the thinking of the human variety.

The Invention of Problems

Let us step back a moment. Sarah's accomplishments may seem like so many puzzles invented by clever psychologists—amusing to us, perhaps even to Sarah, but having nothing to do with the real world of a chimpanzee. But this would be a gross underestimation. There is reason to think that Sarah's abilities and those of other apes are telling us about the dawn of the kind of intelligence that makes us human. Let us look again at the capacities for analogy that have been revealed in the minds of chimpanzees. Without specialized training, we find evidence of

- explicit attributes representing object categories based on physical similarity
- the ability to perceive O-sameness as well as more concrete physical relations.

At least when chimpanzees are given training in the manipulation of a symbol system, we find in addition

- explicit representation of O-sameness as a concept
- relational mappings based on similarity of relations without similarity of objects
- explicit representation of query markers
- perception of relational similarity, R-sameness.

What advantages might these capacities convey for an animal in its natural habitat?

While spending the period of the First World War on the island of Tenerife, the psychologist Wolfgang Köhler conducted a series of studies demonstrating that chimpanzees can systematically solve problems by using simple tools. If the animal was presented with food that was in some way inaccessible—out of reach overhead, or outside the cage—it would take advantage of objects in its cage to obtain the food. For example, if the food was outside of the cage, the chimpanzee might take a long stick and use it to draw the food within reach. Simple problem solving of this sort has also been observed in the natural behavior of chimpanzees in the wild.

It therefore appears that the higher primates have, in at least some crude sense, invented tools. But actually this statement skips a crucial step: it would be more accurate to say primates have invented problems. This may seem like a surprising claim, since problems seem to be part and parcel of what it means for any animal to be alive on this planet. There is no doubt that a laboratory rat, for example, miserable with hunger as it scours the experimenter's maze searching for food, is confronting what we would all recognize as a problem. But we have no good reason to suppose the rat is thinking about, rather than simply reacting to, its state of deprivation. A problem is the recognition of a gap between the present state of affairs and some desired goal state. To represent *the fact that* it has a problem, an animal needs some explicit representation of what is absent—a solution that would fill the gap. Only then will the animal be able to reason about how the problem might be solved or recognize that some object provides a means to close the gap—that is, recognize that an object might be used as a tool.

This description of what it means to have a concept of a problem should sound familiar. We have already talked about what it takes to represent a gap in knowledge: a query marker. A problem is really just a kind of question to oneself. It has the basic form, "How can the current state of affairs be transformed into a state in which my goal is achieved?" or

transform (<solution?>, <initial state>, <goal state>).

The symbol "solution?" represents the missing knowledge of how to accomplish the desired transformation. The solution is generally an action or sequence of actions, and a tool is an object that is used to help perform the required actions. A tool, then, is more than just an object used to do something: it has to fill a special slot that is defined by its role in the overall schema for a problem.

With the invention of some simple version of the concept of a problem, it becomes possible to solve a problem by recognizing that it involves objects and relations similar to those involved in previous experiences. In other words, the capacities for analogy can become mental tools for problem solving. Let us imagine a fanciful, but perhaps not entirely implausible, scenario for the invention of a tool by a chimpanzee living in the wild. Suppose that one day it playfully bangs a rock against a nut, and the nut happens to break open. With a basic capacity for recognizing causal sequences and for storing explicit propositions in memory, the chimpanzee might code the initial state of the nut as

closed (nut-1) *name*: closed-1,

its own action of striking the nut with the rock as

strike (self, nut-1, rock-1) *name*: strike-1,

the final state of the nut as

open (nut-1) *name*: open-1,

and the causal connection between the action and the change in the state of the nut, from closed to open, as

break-open (strike-1, closed-1, open-1) *name*: break-open-1.

Now let us suppose that on some future occasion this chimpanzee is hungry. It finds a nut (nut-2) but is at first unable to open it. A rock (rock-2, different than the earlier one) is lying within its view. The animal may be able to formulate its problem as wanting to take this closed nut,

closed (nut-2) *name*: closed-2

and have it open,

open (nut-2) *name*: **open-2**,

by opening it somehow,

make-open (action?, closed-2, **open-2**) *name*: make-open-2

The boldface on **open-2** signifies that the animal must be able to understand this to be a desired goal state—it can imagine the nut to be open, even though at this moment it remains closed. In addition, the symbol "action?" serves as a query marker for the (so far unknown) action that would accomplish the desired physical transformation.

If the animal can succeed in formulating its problem along the above lines, its basic analogical tools can be brought into play. Physical similarity of elements may be sufficient for the present problem to cue the animal's prior experience in opening nuts. It will then be straightforward to establish the mappings:

nut-2 ↔ nut-1

closed-2 ↔ closed-1

open-2 ↔ open-1

self ↔ self.

Furthermore, the animal may be able to detect that its present goal of opening nut-2 is relationally similar to breaking open nut-1, so that

make-open-2 ↔ break-open-1.

Now the animal can make some inferences by analogical substitution. By trying to complete the mapping of make-open-2 to break-open-1, matching corresponding slot fillers, the chimpanzee finds that

action? ↔ strike-1.

Its final step is to construct a specific description of "action?" using the established correspondences. By performing copying with substitution, the basic device for generating analogical inferences, the animal should now be able to formulate a much more specific description of its problem—how to break the nut by striking it with something:

strike (self, nut-2, object?) *name*: strike-2.

At this point the animal has a new question: what can be used to strike the nut? Given that all other slots have been mapped from strike-2 to strike-1, consistency requires that

object? ↔ rock-1.

If the chimpanzee had not noticed already, this should be enough to allow simple attribute mapping to find a filler for the query marker "object?" namely

rock-2 ↔ rock-1.

The last gap has now been filled. The original query marker, "action?" has been replaced by a plan for action:

strike (self, nut-2, rock-2) *name*: strike-2.

This may seem like a lot of mental work just to figure out that a rock can be used to crack a nut now, just as on another day some other rock was used to crack some other nut. However, the "just" reflects our own proficiency in such routine tasks. If this kind of problem keeps recurring, and the animal has some basic learning ability, the plan for a solution may eventually be recast as a direct rule, something like "If I want to break open a nut, strike it with a handy rock." But the situation is quite different for an animal at the dawn of problem solving or a child first learning about its world or for any of us when faced with a novel problem for which we have no prestored recipe for reaching the goal. Analogy provides a way to fill the gaps in a novel problem by mapping it to a past experience involving similar objects and relations.

Given our earlier discussion of the actual evidence about their mapping ability, have we credited our hypothetical wild chimpanzee with more mental skill than is justified? After all, our wild chimpanzee did not have Sarah's schooling, but we have assumed it could explicitly represent relations and query markers, imagine possible states, and detect R-sameness. However, in many ways our scenario is not that demanding. The only relations required are ones that express frequently used physical actions, such as those required to open nuts, that are closely tied to achieving the animal's basic goals. There is good reason to suspect that coding of causal connections based on one's own physical actions is at the evolutionary leading edge of cognitive sophistication. Coding causal relations is essential for explicit problem solving, a mental skill that surely had survival value for the primate ancestors of both humans and chimpanzees. Furthermore, the mappings that we assumed were heavily guided by similarity of objects, such as nut-1 and nut-2, as well as similarity of relations, such as break-open-1 and make-open-2. Much of the work required for mapping the source and target situations could be done by attribute rather than relational mapping.

Analogy and problem solving may well have evolved together. Sarah is certainly proficient at both analogy and problem solving. We already mentioned that analogies based on functional relations, such as the "can opener to the can," are among the problems on which she excels. It is also clear that she can perceive causal relations of considerable complexity. When shown a videotape of a human actor struggling to escape from a locked cage, Sarah would select a photograph of a key, the instrument for a solution. (At least she would if the actor was someone she liked. If she disliked the person in the videotape, she would often pick a photograph showing some "bad" outcome!) In other tests she demonstrated that she understood the roles that actions play in transforming objects. For example, she was taught actions that had opposite causal effects, such as marking versus erasing a piece of paper. She was also taught to read sequences from left to right. Then she was given sequences such as

blank paper ——— marked paper

and the reverse,

marked paper ——— blank paper,

along with three alternative instruments, including pencil and eraser. After preliminary training, she could reliably choose the correct instrument that would effect each change (i.e., pencil in the first example, eraser in the second). This performance shows that Sarah did not simply know that blank paper, marked paper, pencils, and erasers go together in some loose, associative way. Rather, she knew what slot each state or instrument fills in the relational structure that represents a causal event sequence, in which an instrument is used to change an initial state into a final state. Sarah's ability to distinguish between causal sequences with reversed initial and end states resembles the human ability (discussed in chapter 2) to distinguish such pairs of propositions as "Hercules chases Fifi" and "Fifi chases Hercules."

Animals without language training failed even simpler versions of such tests. But notice that what Sarah was asked to do was considerably more complex than figuring out that if one rock can crack a nut, another rock might open another nut. It seems that Sarah's understanding of causal sequences, like her ability to recognize relations between relations, builds on mental skills that in simpler forms are exhibited by untutored members of her species. Present-day apes may well have the kind of mind that was once possessed by the long-vanished evolutionary ancestors of humans.

Hitting the Wall

The performances of Sarah and other similarly trained apes reveals that the basic constraints postulated by our multiconstraint theory of analogy—similarity, structure, and purpose—are already in place in the chimpanzee. Sarah is able to detect and represent similarities between objects and relations. By mapping similar relations, she can place pairs of objects in correspondence on the basis of their structural roles. For example, by mapping the relation of opening a can to the similar relation of opening a lock, Sarah can determine that a can opener is to a can what a key is to a lock, even though a can opener does not look like a key nor a can like a lock. Sarah is thus capable of relational as well as attribute mapping. Moreover, Sarah's skill in the use of analogies is matched by her proficiency in understanding causal relations and the purposes of actions. She is able to understand questions that depend on such functional relationships as that linking a desired state with a means of attaining it. The cognitive tools that the chimpanzee has available to represent queries and relations between objects appear to be sufficient to allow simple problem solving by analogy.

And yet something is still missing. Sarah was a bright chimpanzee with the benefit of the best education. She broke new intellectual ground for her species, but she never became a rocket scientist. Not only did she never discover the wave theory of sound (a small complaint—not many people have either), but she would never have been able to make the slightest sense of it. It is always a bit dangerous to draw negative conclusions from what an animal fails to do. Perhaps her training was somehow less than ideal after all, and the sheer time and effort required obviously made it impossible to attempt teaching her everything. But it is clear that at some point the chimpanzee "hits the wall" and can follow the path of human intelligence no further.

One striking failure that Sarah exhibited in her understanding of causal relations is instructive. She was shown a videotape in which a human actor discovers a small fire of paper burning on the floor; he looks concerned as if he wishes to put the fire out. Sarah was then shown three alternatives (matches, knife, clay) and trained to associate a token with the likely cause of the fire (matches). In addition, she was shown three other alternatives (water, tape, eraser) and trained to associate another token with the instrument that could provide a solution to the actor's problem (water). After being trained to make the appropriate choices, Sarah was tested on her ability to use the tokens appropriately on tests with other problems. For example, suppose she saw videotape of a scene in which a person is concerned because a piece of paper has been cut. Would Sarah associate the cause token with knife, and the solution token with tape? Premack reports that the answer is "no." Thus even though Sarah can clearly understand what is the specific cause of a specific outcome (burning paper or cut paper), she seems unable to acquire the more abstract concept of cause itself, abstracted from any specific causal scenario.

It is certainly not surprising that chimpanzees are incapable of full human intelligence. To begin with, the gross neurological differences between chimpanzees and humans are massive. The human brain is over three times larger in volume than that of the chimpanzee. Most of the increased size of the human brain involves the cerebral hemispheres, the site of advanced cognitive functioning; and the greatest changes of all involve the frontal lobes that lie just behind our brows. Clearly, these changes point toward some major advances in cognitive capacity. Furthermore, chimpanzees are considerably weaker than human children at a number of tasks involving representations. For example, although chimpanzees can be taught to match pictures of objects to the objects themselves, they find this task very difficult. In the course of learning, they are prone to match pictures with pictures and objects to objects, regardless of their content. Thus when shown a banana as the sample and forced to choose between a picture of a banana and an actual shoe, the animal is likely to pick the shoe. Interestingly, similar errors are made by severely retarded children.

Earlier we sketched a general strategy for deepening the abstraction of thinking by generating new concepts that explicitly represent the common basis for certain analogical correspondences. O-sameness, we argued, was a concept that explicitly represents the basis for attribute mappings that depend on physical similarity. Armed with this relation, an animal like Sarah can now map pairs of objects in which the two items are O-same as each other. We suggested that this process of progressive abstraction could be extended. What should the next intellectual move be? To follow our strategy, the animal (either altered by evolution or guided by its own learning capability) should now use its new mappings to aid it in forming new concepts that explicitly represent the basis for the mappings.

What would this mean? For Sarah, it would mean forming explicit concepts of R-sameness and R-difference to express the basis for her relational mappings. These concepts are higher-order relations between relations and hence would allow Sarah to move from relational to system mappings. We can illustrate an analogy requiring system mapping by further generalizing the match-to-sample task. Sarah can solve problems in which the sample is a single object or a pair of objects; now let us consider a version based on a sample of four objects, such as the problem in figure 3.9. Which of the two alternatives would you choose as the best match? There is no real right answer, but there is a deeper answer. Notice that both alternatives involve entirely different objects than the sample. The first alternative (the left one in figure 3.9) also involves different relations than the sample: in the sample, both pairs are O-same

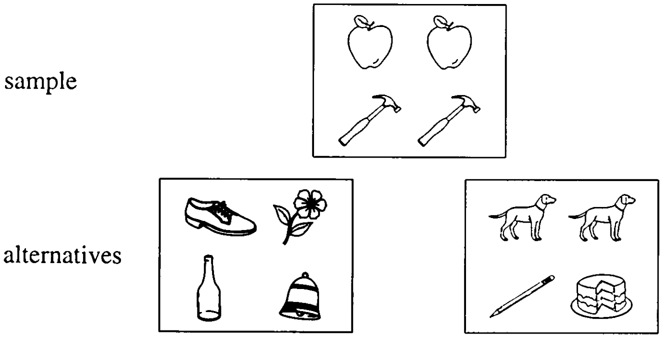


Figure 3.9
A match-to-sample problem based on quadruples of objects, which can only be solved by system mapping.

as each other, whereas in the first alternative both pairs are O-different. In the second alternative, the top pair is O-same, but the bottom pair is O-different. The second alternative thus has more overlap with the sample at the level of first-order relations than does the first.

However, as is illustrated in figure 3.10, only the first alternative is truly isomorphic to the sample. If we represent the explicit concept of R-sameness, we can see that the two O-different relations in this alternative are R-same as each other, just as the two O-same relations in the sample are R-same as each other. That is, an implicit reaction to *R^2-sameness,* the higher-order sameness of relations between relations, supports consistently mapping the relation of R-same in the sample to R-different in the alternative. This system mapping provides a basis for choosing the first alternative as the better match to the sample.

We conjecture that Sarah would be unable to appreciate the deeper match in the quadruple version of the match-to-sample task, because chimpanzees are unable to form mappings at the system level. If this conjecture is correct, chimpanzees will have other related cognitive limitations. For example, an animal capable of system mapping might be able to represent the fact that the relations of marking with a pencil and erasing with an eraser are R-different in a special way: the slot fillers for the initial and final states are reversed. That is, given

mark (pencil, blank-paper, marked-paper) *name*: mark-1

erase (eraser, marked-paper, blank-paper) *name*: erase-1,

the mappings are

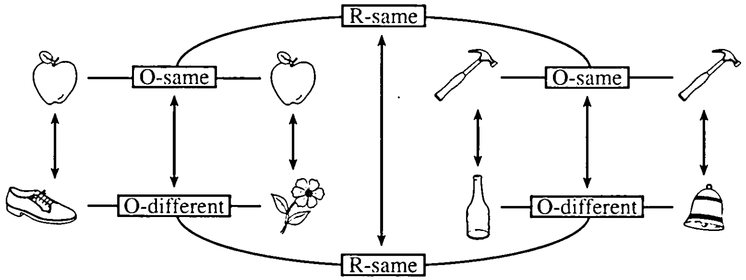


Figure 3.10
The mapping between the sample and the first alternative in figure 3.9, which share only the higher-order relation of R-sameness.

mark ↔ erase

pencil ↔ eraser

blank-paper ↔ marked-paper

marked-paper ↔ blank-paper.

Let us call this special kind of R-difference between mark-1 and erase-1 "R-reverse," since it basically is the mapping between actions that reverse each other's effect. So we have

R-reverse (mark-1, erase-1) *name*: R-reverse-1.

Notice that the relation of R-reversal depends on finding a mapping between two triples of slots, where the slot fillers cannot be mapped on the basis of element similarity (note that blank paper maps to marked paper rather than to blank paper, opposite to the obvious attribute mapping) nor on the basis of perceptual similarity between actions (since marking and erasing do not look that much alike). Forming an explicit concept for the higher-order relation of R-reversal thus depends on the capacity to perform a system mapping, exploiting the structural constraint of isomorphism.

Now, it happens that the relation of R-reversal is the basis for other mappings. For example, consider examples of cutting and joining:

cut (knife, whole-paper, cut-paper) *name*: cut-1

join (tape, cut-paper, whole-paper) *name*: join-1.

These two actions can also be mapped, resulting in correspondences that satisfy R-reversal, which can be summarized as

R-reverse (cut-1, join-1) *name*: R-reverse-2.

Now we have two examples of the higher-order relation R-reverse, which we might say are R^2-*same* as each other—that is, we have sameness of relations between relations. As in the system mapping required to solve the quadruple version of the match-to-sample task, we have moved another notch up the abstraction hierarchy.

Now suppose the animal not only could respond implicitly to R^2-sameness but was able to formulate an explicit concept to represent the way in which the relation between marking and erasing, reversibility of a transformation, is R^2-same as the relation between cutting and joining. Such an animal would understand the explicit concept of "reversible transformation" and could potentially use this concept to think about such issues as what properties of transformations are required for them to be reversible. Our strategy for deepening the abstraction of thought would have been repeated yet again.

As far as we can tell, neither Sarah nor any other chimpanzee could ever be this insightful. Evolution did not innately provide the chimpanzee with the concept of sameness of relations between relations; nor did it provide the learning capability that would enable an individual chimpanzee to form such concepts. A creature that was born with, or that could learn, such concepts would be able to move beyond the simple relational mappings at which Sarah excels, to perform more complex system mappings. Rather than being satisfied by having solved a problem by analogy, it might go on to ponder the basis for the analogy, internally reorganizing its knowledge in a quest for hidden regularities. In addition to seeing correspondences between known objects and relations, it might elaborate partial mappings by positing hidden causes—God in Heaven, invisible ripples of sound in the air.

Evolution did not endow the chimpanzee with these capacities. But it so favored another species that walks this earth today. Taking another leap, we arrive at ourselves.

Summary

Analogical thinking is the product of evolutionary changes in the way animals represent knowledge of relations. All vertebrates are able to respond implicitly to similarity between relations in a way that goes beyond direct similarity of objects. However, only in primates do we find clear evidence of explicit knowledge of relations. Evidence from match-to-sample tasks reveals that monkeys can perform simple attribute mapping based on explicit representations of sameness of objects. But monkeys are unable to solve similar problems that would require explicit representations of sameness of relations, rather than just sameness of objects. In contrast, chimpanzees have the capability to explicitly think about sameness of relations; however, this capability is only fully revealed after special training in the use of symbols for "same" and "different." The development of explicit knowledge of relations appears to be related not only to the use of analogy but also to the kind of deliberative thinking needed to reason about how to solve problems. Chimpanzee intelligence is nonetheless bounded in ways that suggest this species is incapable of system mapping, which is required for more abstract forms of analogical thinking. Although chimpanzees can think explicitly about first-order relations, they do not seem to be able to think explicitly about higher-order relations, such as cause. This limitation restricts their ability to use structure to guide analogical thinking and also their ability to think about the purpose of analogies.