High-Level Perception, Representation, and Analogy: A Critique of Artificial Intelligence Methodology

David J. Chalmers, Robert M. French, Douglas R. Hofstadter

Center for Research on Concepts and Cognition
Indiana University
Bloomington, Indiana 47408

CRCC Technical Report 49 — March 1991

E-mail addresses:
dave@cogsci.indiana.edu
french@cogsci.indiana.edu
dughof@cogsci.indiana.edu

To appear in *Journal of Experimental and Theoretical Artificial Intelligence*.

High-Level Perception, Representation, and Analogy: A Critique of Artificial Intelligence Methodology

Abstract

High-level perception—the process of making sense of complex data at an abstract, conceptual level—is fundamental to human cognition. Through high-level perception, chaotic environmental stimuli are organized into the mental representations that are used throughout cognitive processing. Much work in traditional artificial intelligence has ignored the process of high-level perception, by starting with hand-coded representations. In this paper, we argue that this dismissal of perceptual processes leads to distorted models of human cognition. We examine some existing artificial-intelligence models—notably BACON, a model of scientific discovery, and the Structure-Mapping Engine, a model of analogical thought—and argue that these are flawed precisely because they downplay the role of high-level perception. Further, we argue that perceptual processes cannot be separated from other cognitive processes even in principle, and therefore that traditional artificial-intelligence models cannot be defended by supposing the existence of a "representation module" that supplies representations ready-made. Finally, we describe a model of high-level perception and analogical thought in which perceptual processing is integrated with analogical mapping, leading to the flexible build-up of representations appropriate to a given context.

1 The Problem of Perception

One of the deepest problems in cognitive science is that of understanding how people make sense of the vast amount of raw data constantly bombarding them from their environment. The essence of human perception lies in the ability of the mind to hew order from this chaos, whether this means simply detecting movement in the visual field, recognizing sadness in a tone of voice, perceiving a threat on a chessboard, or coming to understand the Iran–Contra affair in terms of Watergate.

It has long been recognized that perception goes on at many levels. Immanuel Kant divided the perceptual work of the mind into two parts: the faculty of Sensibility, whose job it is to pick up raw sensory information, and the faculty of Understanding, which is devoted to organizing these data into a coherent, meaningful experience of the world. Kant found the faculty of Sensibility rather uninteresting, but he devoted much effort to the faculty of Understanding. He went so far as to propose a detailed model of the higher-level perceptual processes involved, dividing the faculty into twelve Categories of Understanding.

Today Kant's model seems somewhat baroque, but his fundamental insight remains valid. Perceptual processes form a spectrum, which for convenience we can divide into two components. Corresponding roughly to Kant's faculty of Sensibility, we have low-level perception, which involves the early processing of information from the various sensory modalities. High-level perception, on the other hand, involves taking a more global view of this information, extracting *meaning* from the raw material by accessing concepts, and making sense of situations at a conceptual level. This ranges from the recognition of objects to the grasping of abstract relations, and on to understanding entire situations as coherent wholes.

Low-level perception is far from uninteresting, but it is high-level perception that is most relevant to the central problems of cognition. The study of high-level perception leads us directly to the problem of mental *representation*. Representations are the fruits of perception. In order for raw data to be shaped into a coherent whole, they must go through a process of filtering and organization, yielding a structured representation that can be used by the mind for any number of purposes. A primary question about representations, currently the subject of much debate, concerns their precise structure. Of equal importance is the question of how these representations might be *formed* in the first place, via a process of perception, starting from raw data. The process of representation-formation raises many important questions: How are representations influenced by context? How can our perceptions of a situation radically reshape themselves when necessary? Where in the process of perception are concepts accessed? Where does meaning enter, and where and how does understanding emerge?

The main thesis of this paper is that high-level perception is deeply interwoven with other cognitive processes, and that researchers in artificial intelligence must therefore integrate perceptual processing into their modeling of cognition. Much work in artificial intelligence has attempted to model conceptual processes independently of perceptual processes, but we will argue that this approach cannot lead to a satisfactory understanding of the human mind. We will examine some existing models of scientific discovery and analogical thought in support of this claim, and will argue that the exclusion of perceptual processes from these models leads to serious limitations. The intimate link between analogical thought and high-level perception will be investigated in detail, and we will describe a computational model in which the two processes are integrated.

Low-level and high-level perception

The lowest level of perception occurs with the reception of raw sensory information by various sense organs. Light impinges on the retina, sound waves cause the eardrum to vibrate, and so on. Other processes further along the information-processing chain may also be usefully designated as low-level. In the case of vision, for instance, after information has passed up the optic nerve, much basic processing occurs in the lateral geniculate nuclei and the primary visual cortex, as well as the superior colliculus. Included here is the processing of brightness contrasts, of light boundaries, and of edges and corners in the visual field, and perhaps also location processing.

Low-level perception is given short shrift in this paper, as it is quite removed from the more cognitive questions of representation and meaning. Nonetheless, it is an important subject of study, and a complete theory of perception will necessarily include low-level perception as a fundamental component.

The transition from low-level to high-level perception is of course quite blurry, but we may delineate it roughly as follows. High-level perception begins at that level of processing where *concepts* begin to play an important role. Processes of high-level perception may be subdivided again into a spectrum from the concrete to the abstract. At the most concrete end of the spectrum, we have *object recognition*, exemplified by the ability to recognize an apple on a table, or to pick out a farmer in a wheat field. Then there is the ability to grasp *relation*s. This allows us to determine the relationship between a blimp and the ground (“above”), or a swimmer and a swimming pool (“in”). As one moves further up the spectrum towards more abstract relations (“George Bush is *in* the Republican Party”), the issues become distant from particular sensory modalities. The most abstract kind of perception is the processing of entire complex *situations*, such as a love affair or a war.

One of the most important properties of high-level perception is that it is extremely flexible.

A given set of input data may be perceived in a number of different ways, depending on the context and the state of the perceiver. Due to this flexibility, it is a mistake to regard perception as a process that associates a fixed representation with a particular situation. Both contextual factors and top-down cognitive influences make the process far less rigid than this. Some of the sources of this flexibility in perception are as follows.

Perception may be influenced by belief. Numerous experiments by the "New Look" theorists in psychology in the 1950's (e.g., Bruner 1957) showed that our expectations play an important role in determining what we perceive even at quite a low level. At a higher level, that of complete situations, such influence is ubiquitous. Take for instance the situation in which a husband walks in to find his wife sitting on the couch with a male stranger. If he has a prior belief that his wife has been unfaithful, he is likely to perceive the situation one way; if he believes that an insurance salesman was due to visit that day, he will probably perceive the situation quite differently.

Perception may be influenced by goals. If we are trying to hike on a trail, we are likely to perceive a fallen log as an obstacle to be avoided. If we are trying to build a fire, we may perceive the same log as useful fuel for the fire. Another example: Reading a given text may yield very different perceptions, depending on whether we are reading it for content or proofreading it.

Perception may be influenced by external context. Even in relatively low-level perception, it is well known that the surrounding context can significantly affect our perception of visual images. For example, an ambiguous figure halfway between an "A" and an "H" is perceived one way in the context of "C—T", and another in the context of "T—E". At a higher level, if we encounter somebody dressed in tuxedo and bow-tie, our perception of them may differ depending on whether we encounter them at a formal ball or at the beach.

Perceptions of a situation can be radically reshaped where necessary. In Maier's well-known two-string experiment (Maier 1931), subjects are provided with a chair and a pair of pliers, and are told to tie together two strings hanging from the ceiling. The two strings are too far apart to be grasped simultaneously. Subjects have great difficulty initially, but after a number of minutes some of them hit upon the solution of tying the pliers to one of the strings, and swinging the string like a pendulum. Initially, the subjects perceive the pliers first and foremost as a special tool; if the weight of the pliers is perceived at all, it is very much in the background. To solve this problem, subjects have to radically alter the emphasis of their perception of the pair of pliers. Its function as a tool is set aside, and its weightiness is brought into the foreground as the key feature in this situation.

The distinguishing mark of high-level perception is that it is semantic: it involves drawing *meaning* out of situations. The more semantic the processing involved, the greater the role played by *concepts* in this processing, and thus the greater the scope for top-down influences. The most abstract of all types of perception, the understanding of complete situations, is also the most flexible.

Recently both Pylyshyn (1980) and Fodor (1983) have argued against the existence of top-down influences in perception, claiming that perceptual processes are "cognitively impenetrable" or "informationally encapsulated". These arguments are highly controversial, but in any case they apply mostly to relatively low-level sensory perception. Few would dispute that at the higher, conceptual level of perception, top-down and contextual influences play a large role.

2 Artificial Intelligence and the Problem of Representation

The end product of the process of perception, when a set of raw data has been organized into a coherent and structured whole, is a *representation*. Representations have been the object of much study and debate within the field of artificial intelligence, and much is made of the "representation problem". This problem has traditionally been phrased as "What is the correct structure for mental representations?", and many possibilities have been suggested, ranging from predicate calculus through frames and scripts to semantic networks and more. We may divide representations into two kinds: long-term knowledge representations that are stored passively somewhere in the system, and short-term representations that are active at a given moment in a particular mental or computational process. (This distinction corresponds to the distinction between long-term memory and working memory.) In this discussion, we will mostly be concerned with short-term, active representations, as it is these that are the direct product of perception.

The question of the structure of representations is certainly an important one, but there is another, related problem that has not received nearly as much attention. This is that of understanding how such a representation could be arrived at, starting from environmental data. Even if it were possible to discover an optimal type of representational structure, this would leave unresolved two important problems, namely:

> *The problem of relevance*: How is it decided which subsets of the vast amounts of data from the environment get used in various parts of the representational structure? Naturally, much of the information content at the lowest level will be quite irrelevant at the highest representational level. To determine which parts of the data are relevant to a given representation, a complex filtering process is required.
>
> *The problem of organization*: How are these data put into the correct form for the

> representation? Even if we have determined precisely which data are relevant, and we have determined the desired framework for the representation—a frame-based representation, for instance—we still face the problem of organizing the data into the representational form in a useful way. The data do not come prepackaged as slots and fillers, and organizing them into a coherent structure is likely to be a highly non-trivial task.

These questions, taken together, amount in essence to the problem of high-level perception, translated into the framework of artificial intelligence.

The traditional approach in artificial intelligence has been to *start* by selecting not only a preferred type of high-level representational structure, but also the data assumed to be relevant to the problem at hand. These data are organized by a human programmer who appropriately fits them into the chosen representational structure. Usually, researchers use their prior knowledge of the nature of the problem to hand-code a representation of the data into a near-optimal form. Only after all this hand-coding is completed is the representation allowed to be manipulated by the machine. The problem of representation-formation, and thus the problem of high-level perception, is ignored. (These comments do not, of course, apply to work in machine vision, speech processing, and other perceptual endeavors. However, work in these fields usually stops short of modeling processes at the conceptual level and is thus not directly relevant to our critique of high-level cognitive modeling.)

The formation of appropriate representations lies at the heart of human high-level cognitive abilities. It might even be said that the problem of high-level perception forms the central task facing the artificial-intelligence community: the task of understanding how to draw *meaning* out of the world. It might not be stretching the point to say that there is a "meaning barrier", which has rarely been crossed by work in AI. On one side of the barrier, some models in low-level perception have been capable of building primitive representations of the environment, but these are not yet sufficiently complex to be called "meaningful". On the other side of the barrier, much research in high-level cognitive modeling has *started* with representations at the conceptual level, such as propositions in predicate logic or nodes in a semantic network, where any meaning that is present is already built in. There has been very little work that bridges the gap between the two.

Objectivism and traditional AI

Once AI takes the problem of representation-formation seriously, the next stage will be to deal with the evident flexibility of human high-level perceptual processes. As we have seen, objects and situations can be comprehended in many different ways, depending on context and top-down influences. We must find a way of ensuring that AI representations have a corresponding degree of flexibility. William James, in the late nineteenth century, recognized this aspect of cognitive representations:

> "There is no property ABSOLUTELY essential to one thing. The same property which figures as the essence of a thing on one occasion becomes a very inessential feature upon another. Now that I am writing, it is essential that I conceive my paper as a surface for inscription. . . . But if I wished to light a fire, and no other materials were by, the essential way of conceiving the paper would be as a combustible material. . . . The essence of a thing is that one of its properties which is so *important for my interests* that in comparison with it I may neglect the rest. . . . The properties which are important vary from man to man and from hour to hour. . . . many objects of daily use—as paper, ink, butter, overcoat—have properties of such constant unwavering importance, and have such stereotyped names, that we end by believing that to conceive them in those ways is to conceive them in the only true way. Those are no truer ways of conceiving them than any others; there are only more frequently serviceable ways to us." (James 1890, pp. 222–224)

James is saying, effectively, that we have different representations of an object or situation at different times. The representational process adapts to fit the pressures of a given context.

Despite the work of philosopher-psychologists such as James, the early days of artificial intelligence were characterized by an objectivist view of perception, and of the representation of objects, situations, and categories. As the linguist George Lakoff has characterized it, "On the objectivist view, reality comes complete with a unique correct, complete structure in terms of entities, properties and relations. This structure exists, independent of any human understanding." (Lakoff 1987, p. 159) While this objectivist position has been unfashionable for decades in philosophical circles (especially after Wittgenstein's work demonstrating the inappropriateness of a rigid correspondence between language and reality), most early work in AI implicitly accepted this set of assumptions.

The Physical Symbol System Hypothesis (Newell & Simon 1976), upon which most of the traditional AI enterprise has been built, posits that thinking occurs through the manipulation of symbolic representations, which are composed of atomic symbolic primitives. Such symbolic representations are by their nature somewhat rigid, black-and-white entities, and it is difficult for their representational content to shift subtly in response to changes in context. The result, in practice—irrespective of whether this was intended by the original proponents of this framework—is a structuring of reality that tends to be as fixed and absolute as that of the objectivist position outlined above.

By the mid-seventies, a small number of AI researchers began to argue that in order to progress, the field would have to part ways with its commitment to such a rigid representational framework. One of the strongest early proponents of this view was David Marr, who noted that

> "the perception of an event or object must include the simultaneous computation of several different descriptions of it, that capture diverse aspects of the use, purpose or circumstances of the event or object." (Marr 1977, p. 44)

Recently, significant steps have been taken toward representational flexibility with the advent of sophisticated connectionist models whose distributed representations are highly context-dependent (Rumelhart & McClelland 1986). In these models, there are no representational primitives in internal processing. Instead, each representation is a vector in a multi-dimensional space, whose position is not anchored but can adjust flexibly to changes in environmental stimuli. Consequently, members of a category are not all represented by identical symbolic structures; rather, individual objects will be represented in subtly different ways depending upon the context in which they are presented. In networks with recurrent connections (Elman 1990), representations are even sensitive to the current internal state of the model. Other recent work taking a flexible approach to representation includes the classifier-system models of Holland (1986) and his colleagues, where genetically-inspired methods are used to create a set of "classifiers" that can respond to diverse aspects of various situations.

In these models, a flexible perceptual process has been integrated with an equally flexible dependence of action upon representational content, yielding models that respond to diverse situations with a robustness that is difficult to match with traditional methods. Nonetheless, the models are still somewhat primitive, and the representations they develop are not nearly as complex as the hand-coded, hierarchically-structured representations found in traditional models; still, it seems to be a step in the right direction. It remains to be seen whether work in more traditional AI paradigms will respond to this challenge by moving toward more flexible and robust representational forms.

On the possibility of a representation module

It might be granted that given the difficulty of the problem of high-level perception, AI researchers could be forgiven for starting with their representations in a made-to-order form. They might plausibly claim that the difficult problem of representation-formation is better left until later. But it must be realized that behind this approach lies a tacit assumption: that it is possible to model high-level cognitive processes independently of perceptual processes. Under this assumption, the representations that are currently, for the most part, tailored by human hands, would eventually be built up by a separate lower-level facility—a "representation module" whose job it would be to funnel data into representations. Such a module would act as a "front end" to the models of the cognitive processes currently being studied, supplying them with the appropriately-tailored representations.

We are deeply skeptical, however, about the feasibility of such a separation of perception from the rest of cognition. A representation module that, given any situation, produced the single "correct" representation for it, would have great difficulty emulating the flexibility that characterizes human perception. For such flexibility to arise, the representational processes would have to be sensitive to the needs of all the various cognitive processes in which they might be used. It seems most unlikely that a single representation would suffice for all purposes. As we have seen, for the accurate modeling of cognition it is necessary that the representation of a given situation can vary with various contextual and top-down influences. This, however, is directly contrary to the "representation module" philosophy, wherein representations are produced quite separately from later cognitive processes, and then supplied to a "task-processing" module.

To separate representation-building from higher-level cognitive tasks is, we believe, impossible. In order to provide the kind of flexibility that is apparent in cognition, any fully cognitive model will probably require a continual interaction between the process of representation-building and the manipulation of those representations. If this proves to be the case, then the current approach of using hand-coded representations not only is postponing an important issue but will, in the long run, lead up a dead-end street.

We will consider this issue in greater depth later, when we discuss current research in the modeling of analogical thought. For now, we will discuss in some detail one well-known AI program for which great claims have been made. We argue that these claims represent a lack of appreciation of the importance of high-level perception.

BACON: A case study

A particularly clear case of a program in which the problem of representation is bypassed is BACON, a well-known program that has been advertised as an accurate model of scientific discovery (Langley *et al* 1987). The authors of BACON claim that their system is "capable of representing information at multiple levels of description, which enables it to discover complex laws involving many terms". BACON was able to "discover", among other things, Boyle's law of ideal gases, Kepler's third law of planetary motion, Galileo's law of uniform acceleration, and Ohm's law.

Such claims clearly demand close scrutiny. We will look in particular at the program's "discovery" of Kepler's third law of planetary motion. Upon examination, it seems that the success of the program relies almost entirely on its being given data that have already been represented in near-optimal form, using after-the-fact knowledge available to the programmers.

When BACON performed its derivation of Kepler's third law, the program was given only data about the planets' average distances from the sun and their periods. These are *precisely the data required to derive the law*. The program is certainly not "starting with essentially the same initial conditions as the human discoverers", as one of the authors of BACON has claimed (Simon 1989, p. 375). The authors' claim that BACON used "original data" certainly does not mean that it used *all* of the data available to Kepler at the time of his discovery, the vast majority of which were irrelevant, misleading, distracting, or even wrong.

This pre-selection of data may at first seem quite reasonable: after all, what could be more important to an astronomer-mathematician than planetary distances and periods? But here our after-the-fact knowledge is misleading us. Consider for a moment the times in which Kepler lived. It was the turn of the seventeenth century, and Copernicus' *De Revolutionibus Orbium Cœlestium* was still new and far from universally accepted. Further, at that time there was no notion of the forces that produced planetary motion; the sun, in particular, was known to produce light but was not thought to influence the motion of the planets. In that prescientific world, even the notion of using mathematical equations to express regularities in nature was rare. And Kepler believed—in fact, his early fame rested on the discovery of this surprising coincidence—that the planets' distances from the sun were dictated by the fact that the five regular polyhedra could be fit between the five "spheres" of planetary motion around the sun, a fact that constituted seductive but ultimately misleading data.

Within this context, it is hardly surprising that it took Kepler thirteen years to realize that conic sections and not Platonic solids, that algebra and not geometry, that ellipses and not Aristotelian "perfect" circles, that the planets' distances from the sun and not the polyhedra in which they fit, were the *relevant* factors in unlocking the regularities of planetary motion. In making his discoveries, Kepler had to reject a host of conceptual frameworks that might, for all he knew, have applied to planetary motion, such as religious symbolism, superstition, Christian cosmology, and teleology. In order to discover his laws, he had to make all of these creative leaps. BACON, of course, had to do nothing of the sort. The program was given precisely the set of variables it needed from the outset (even if the values of some of these variables were sometimes less than ideal), and was moreover supplied with precisely the right biases to induce the algebraic form of the laws, it being taken completely for granted that mathematical laws of a type now recognized by physicists as standard were the desired outcome.

It is difficult to believe that Kepler would have taken thirteen years to make his discovery if his working data had consisted entirely of a list where each entry said "Planet X: Mean Distance from Sun Y, Period Z". If he had further been told "Find a polynomial equation relating these entities", then it might have taken him a few hours. Addressing the question of why Kepler took thirteen years to do what BACON managed within minutes, Langley *et al* (1987) point to "sleeping time, and time for ordinary daily chores", and other factors such as the time taken in setting up experiments, and the slow hardware of the human nervous system (!). In an interesting juxtaposition to this, researchers in a recent study (Qin & Simon 1990) found that starting with the data that BACON was given, university students could make essentially the same "discoveries" within an hour-long experiment. Somewhat strangely, the authors (including one of the authors of BACON) take this finding to support the plausibility of BACON as an accurate model of scientific discovery. It seems more reasonable to regard it as a demonstration of the vast difference in difficulty between the task faced by BACON and that faced by Kepler, and thus as a *reductio ad absurdum* of the BACON methodology.

So many varieties of data were available to Kepler, and the available data had so many different ways of being interpreted, that it is difficult not to conclude that in presenting their program with data in such a neat form, the authors of BACON are inadvertently guilty of 20–20 hindsight. BACON, in short, works only in a world of hand-picked, prestructured data, a world completely devoid of the problems faced by Kepler or Galileo or Ohm when they made their original discoveries. Similar comments could be made about STAHL, GLAUBER, and other models of scientific discovery by the authors of BACON. In all of these models, the crucial role played by high-level perception in scientific discovery, through the filtering and organization of environmental stimuli, is ignored.

It is interesting to note that the notion of a "paradigm shift", which is central to much scientific discovery (Kuhn 1970), is often regarded as the process of *viewing the world* in a radically different way. That is, scientists' frameworks for representing available world knowledge are broken down, and their high-level perceptual abilities are used to organize the available data quite differently, building a novel representation of the data. Such a new representation can be used to draw different and important conclusions in a way that was difficult or impossible with the old representation. In this model of scientific discovery, unlike the model presented in BACON, the process of high-level perception is central.

The case of BACON is by no means isolated—it is typical of much work in AI, which often fails to appreciate the importance of the representation-building stage. We will see this in more depth in the next section, in which we take a look at the modeling of analogy.

3 Models of Analogical Thought

Analogical thought is dependent on high-level perception in a very direct way. When people make analogies, they are perceiving some aspects of the structures of two situations—the *essenc-*

es of those situations, in some sense—as identical. These structures, of course, are a product of the process of high-level perception.

The quality of an analogy between two situations depends almost entirely on one's perception of the situations. If Ronald Reagan were to evaluate the validity of an analogy between the U.S. role in Nicaragua and the Soviet Union's role in Afghanistan, he would undoubtedly see it as a poor one. Others might consider the analogy excellent. The difference would come from different perceptions, and thus representations, of the situations themselves. Reagan's internal representation of the Nicaraguan situation is certainly quite different from Daniel Ortega's.

Analogical thought further provides one of the clearest illustrations of the flexible nature of our perceptual abilities. Making an analogy requires highlighting various different aspects of a situation, and the aspects that are highlighted are often not the most obvious features. The perception of a situation can change radically, depending on the analogy we are making.

Let us consider two analogies involving DNA. The first is an analogy between DNA and a zipper. When we are presented with this analogy, the image of DNA that comes to mind is that of two strands of paired nucleotides (which can come apart like a zipper for the purposes of replication). The second analogy involves comparing DNA to the source code (i.e., non-executable high-level code) of a computer program. What comes to mind now is the fact that information in the DNA gets "compiled" (via processes of transcription and translation) into enzymes, which correspond to machine code (i.e., executable code). In the latter analogy, the perception of DNA is radically different—it is represented essentially as an information-bearing entity, whose physical aspects, so important to the first analogy, are of virtually no consequence.

In cases such as these, it seems that no single, rigid representation can capture what is going on in our heads. It is true that we probably have a single rich representation of DNA sitting passively in long-term memory. However, in the contexts of different analogical mappings, very different facets of this large representational structure are selected out as being relevant, by the pressures of the particular context. Irrespective of the *passive* content of the long-term representation of DNA, the *active* content that is processed at a given time is determined by a flexible representational process.

Furthermore, not only is analogy-making dependent on high-level perception, but the reverse holds true as well: perception is often dependent on analogy-making itself. The high-level perception of one situation in terms of another is ubiquitous in human thought. If we perceive Nicaragua as "another Vietnam", for example, the making of the analogy is fleshing out our representation of Nicaragua. Analogical thought provides a powerful mechanism for the enrichment of a representation of a given situation. This is well understood by good educators and writers, who know that there is nothing like an analogy to provide a better mental picture of a given situation. Analogies affect our perception all the time: in a love affair, for instance, it is difficult to stop parallels with past romances from modulating one's perception of the current situation. In the large or the small, such analogical perception—the grasping of one situation in terms of another—is so common that we tend to forget that what is going on is, in fact, analogy. Analogy and perception are tightly bound together.

It is useful to divide analogical thought into two basic components. First, there is the process of *situation-perception*, which involves taking the data involved with a given situation, and filtering and organizing them in various ways to provide an appropriate representation for a given context. Second, there is the process of *mapping*. This involves taking the representations of two situations and finding appropriate correspondences between components of one representation with components of the other to produce the match-up that we call an analogy. It is by no means apparent that these processes are cleanly separable; they seem to interact in a deep way. Given the fact that perception underlies analogy, one might be tempted to divide the process of analogy-making sequentially: first situation perception, then mapping. But we have seen that analogy also plays a large role in perception; thus mapping may be deeply involved in the situation-perception stage, and such a clean division of the processes involved could be misleading. Later, we will consider just how deeply intertwined these two processes are.

Both the situation-perception and mapping processes are essential to analogy-making, but of the two the former is more fundamental, for the simple reason that the mapping process requires representations to work on, and representations are the product of high-level perception. The perceptual processes that produce these representations may in turn deeply involve analogical mapping; but each mapping process requires a perceptual process to precede it, whereas it is not the case that each perceptual process necessarily depends upon mapping. Therefore the perceptual process is conceptually prior, although perception and mapping processes are often temporally interwoven. If the appropriate representations are already formed, the mapping process can often be quite straightforward. In our view, the most central and challenging part of analogy-making is the perceptual process: the shaping of situations into representations appropriate to a given context.

The mapping process, in contrast, is an important object of study especially because of the immediate and natural use it provides for the products of perception. Perception produces a particular structure for the representation of a situation, and the mapping process emphasizes certain aspects of this structure. Through the study of analogy-making, we obtain a direct window onto high-level perceptual processes. The study of which situations people view as analogous can tell us much about how people represent those situations. Along the same lines, the computational modeling of analogy provides an ideal testing-ground for theories of high-level perception. Considering all this, one can see that the investigation of analogical thought has a huge role to play in the understanding of high-level perception.

Current models of analogical thought

In light of these considerations, it is somewhat disheartening to note that almost all current work in the computational modeling of analogy bypasses the process of perception altogether. The dominant approach involves starting with fixed, preordained representations, and launching a mapping process to find appropriate correspondences between representations. The mapping process not only takes center stage; it is the only actor. Perceptual processes are simply ignored; the problem of representation-building is not even an issue. The tacit assumption of such research is that correct representations have (somehow) already been built.

Perhaps the best-known computational model of analogy-making is the Structure-Mapping Engine (SME) (Falkenhainer, Forbus, and Gentner 1990), based upon the structure-mapping theory of Dedre Gentner (1983). We will examine this model within the context of our earlier remarks. Other models of analogy-making, such as those of Burstein (1986), Carbonell (1986), Holyoak & Thagard (1989), Kedar-Cabelli (1988), and Winston (1982), while differing in many respects from the above work, all share the property that the problem of representation-building is bypassed.

Let us consider one of the standard examples from this research, in which the SME program is said to discover an analogy between an atom and the solar system. Here, the program is given representations of the two situations, as shown in Figure 1. Starting with these representations, SME examines many possible correspondences between elements of the first representation and elements of the second. These correspondences are evaluated according to how well they preserve the high-level structure apparent in the representations. The correspondence with the highest score is selected as the best analogical mapping between the two situations.

A brief examination of Figure 1 shows that the discovery of the similar structure in these representations is not a difficult task. The representations have been set up in such a way that the common structure is immediately apparent. Even for a computer program, the extraction of such common structure is relatively straightforward.

We are in broad sympathy with Gentner's notion that the mappings in an analogy should preserve high-level structure (although there is room to debate over the details of the mapping process). But when the program's discovery of the correspondences between the two situations is a direct result of its being explicitly given the appropriate structures to work with, its victory in finding the analogy becomes somewhat hollow. Since the representations are tailored (perhaps unconsciously) to the problem at hand, it is hardly surprising that the correct structural correspondences are not difficult to find. A few pieces of irrelevant information are sometimes thrown in as decoys, but this makes the task of the mapping process only slightly more complicated. The point is that if appropriate representations come presupplied, the hard part of

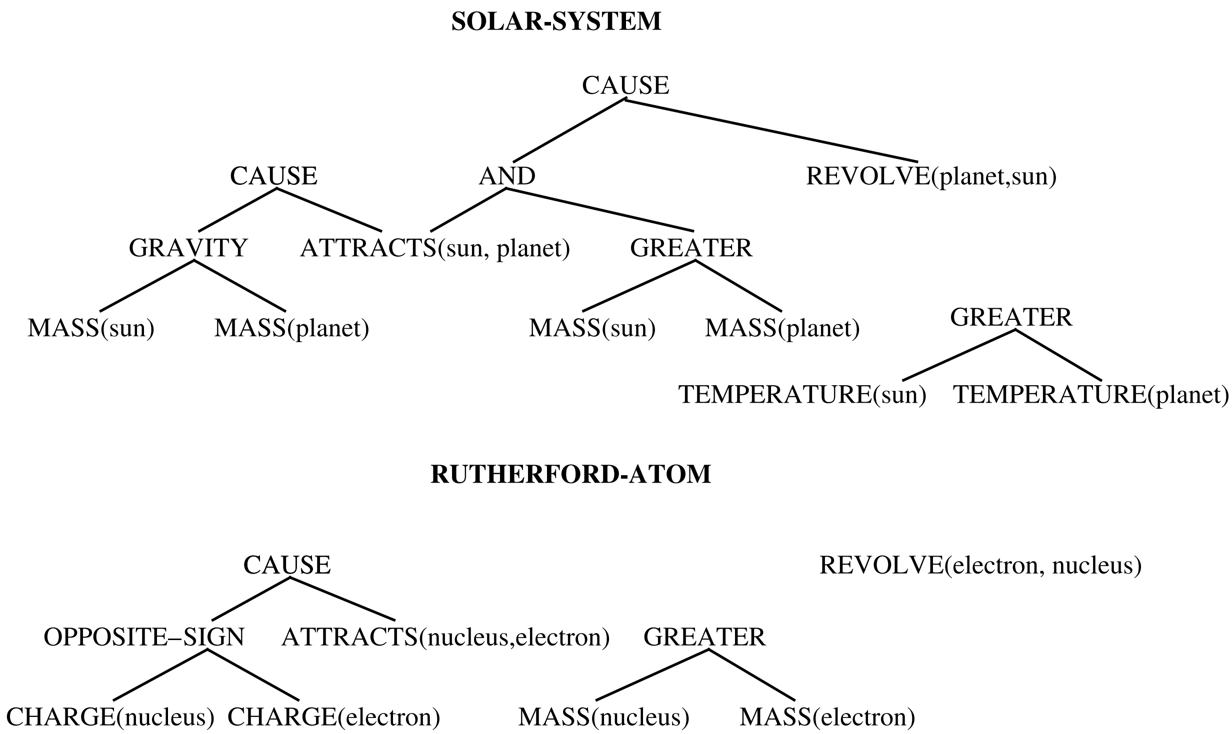

Figure 1. The representations used by SME in finding an analogy between the solar system and the atom. (From Falkenhainer *et al*, 1990.)

the analogy-making task has already been accomplished.

Imagine what it would take to devise a representation of the solar system or an atom independent of any context provided by a particular problem. There are so many data available: one might, for instance, include information about the moons revolving around the planets, about the opposite electric charges on the proton and the electron, about relative velocities, about proximities to other bodies, about the number of moons, about the composition of the sun or the composition of the nucleus, about the fact that the planets lie in one plane and that each planet rotates on its axis, and so on. It comes as no surprise, in view of the analogy sought, that the only relations present in the representations that SME uses for these situations are the following: "attracts", "revolves around", "gravity", "opposite-sign" and "greater" (as well as the fundamental relation "cause"). These, for the most part, are precisely the relations that are relevant factors in this analogy. The criticisms of BACON discussed earlier apply here also: the representations used by both programs seem to have been designed with 20–20 hindsight.

A related problem arises when we consider the distinction that Gentner makes between *objects*, *attributes*, and *relations*. This distinction is fundamental to the operation of SME, which works by mapping objects exclusively to objects and relations to relations, while paying little attention to attributes. In the atom/solar-system analogy such things as the nucleus, the sun, and the electrons are labeled as "objects", while mass and charge, for instance, are considered to be "attributes". However, it seems most unclear that this representational division is so clean in human thought. Many concepts, psychologically, seem to float back and forth between being objects and attributes, for example. Consider a model of economics: should we regard "wealth" as an object that flows from one agent, or as an attribute of the agents that changes with each transaction? There does not appear to be any obvious *a priori* way to make the decision. A similar problem arises with the SME treatment of relations, which are treated as n-place predicates. A 3-place predicate can be mapped only to a 3-place predicate, and never to a 4-place predicate, no matter how semantically close the predicates might be. So it is vitally important that every relation be represented by precisely the right kind of predicate structure in every representation. It seems unlikely that the human mind makes a rigid demarcation between 3-place and 4-place predicates—rather, this kind of thing is probably very blurry.

Thus, when one is designing a representation for SME, a large number of somewhat arbitrary choices have to be made. The performance of the program is highly sensitive to each of these choices. In each of the published examples of analogies made by SME, these representations were designed in just the right way for the analogy to be made. It is difficult to avoid the conclusion that at least to a certain extent, the representations given to SME were constructed with those specific analogies in mind. This is again reminiscent of BACON.

In defense of SME, it must be said that there is much of interest about the mapping process itself; and unlike the creators of BACON, the creators of SME have made no great claims for their program's "insight". It seems a shame, however, that they have paid so little attention to the question of just how the SME's representations could have been formed. Much of what is interesting in analogy-making involves extracting structural commonalities from two *situations*, finding some "essence" that both share. In SME, this problem of high-level perception is swept under the rug, by starting with preformed representations of the situations. The essence of the situations has been drawn out in advance in the formation of these representations, leaving only the relatively easy task of discovering the correct mapping. It is not that the work done by SME is necessarily *wrong*: it is simply not tackling what are, in our opinion, the really difficult issues in analogy-making.

Such criticisms apply equally to most other work in the modeling of analogy. It is interesting to note that one of the earliest computational models of analogy, Evans' ANALOGY (Evans 1968), attempted to build its own representations, even if it did so in a fairly rigid manner. Curiously, however, almost all major analogy-making programs since then have ignored the problem of representation-building. The work of Kedar-Cabelli (1988) takes a limited step in this direction by employing a notion of "purpose" to direct the selection of relevant information, but still starts with all representations pre-built. Other researchers, such as Burstein (1986), Carbonell (1986), and Winston (1982), all have models that differ in significant respects from the work outlined above, but none of these addresses the question of perception.

The ACME program of Holyoak and Thagard (1989) uses a kind of connectionist network to satisfy a set of "soft constraints" in the mapping process, thus determining the best analogical correspondences. Nevertheless, their approach seems to have remained immune to the connectionist notion of context-dependent, flexible representations. The representations used by ACME are preordained, frozen structures of predicate logic; the problem of high-level perception is bypassed. Despite the flexibility provided by a connectionist network, the program has no ability to change its representations under pressure. This constitutes a serious impediment to the attempts of Holyoak and Thagard to capture the flexibility of human analogical thought.

The necessity of integrating high-level perception with more abstract cognitive processing

The fact that most current work on analogical thought has ignored the problem of representation-formation is not necessarily a damning charge: researchers in the field might well defend themselves by saying that this process is far too difficult to study at the moment. In the meantime, they might argue, it is reasonable to assume that the work of high-level perception could be done by a separate "representation module", which takes raw situations and converts them into structured representations. Just how this module might work, they could say, is not their concern. Their research is restricted to the mapping process, which takes these representations as input. The problem of representation, they might claim, is a completely separate issue. (In fact, Forbus, one of the authors of SME, has also worked on modules that build representations in "qualitative physics". Some preliminary work has been done on using these representations as input to SME.)

This approach would be less ambitious than trying to model the entire perception-mapping cycle, but lack of ambition is certainly no reason to condemn a project *a priori*. In cognitive science and elsewhere, scientists usually study what seems within their grasp, and leave problems that seem too difficult for later. If this were all there was to the story, our previous remarks might be read as pointing out the limited scope of the present approaches to analogy, but at the same time applauding their success in making progress on a small part of the problem. There is, however, more to the story than this.

By ignoring the problem of perception in this fashion, artificial-intelligence researchers are making a deep implicit assumption—namely, that the processes of perception and of mapping are temporally separable. As we have already said, we believe that this assumption will not hold up.

We see two compelling arguments against such a separation of perception from mapping. The first argument is simpler, but the second has a broader scope.

The first argument stems from the observation, made earlier, that much perception is dependent on processes of analogy. People are constantly interpreting new situations in terms of old ones. Whenever they do this, they are using the analogical process to build up richer representations of various situations. When the controversial book *The Satanic Verses* was attacked by Iranian Moslems and its author threatened with death, most Americans were quick to condemn the actions of the Iranians. Interestingly, some senior figures in American Christian churches had a somewhat different reaction. Seeing an analogy between this book and the controversial film *The Last Temptation of Christ*, which had been attacked in Christian circles as blasphemous, these figures were hesitant about condemning the Iranian action. Their perception of the situation was significantly altered by such a salient analogy.

Similarly, seeing Nicaragua as analogous to Vietnam might throw a particular perspective on the situation there, while seeing the Nicaraguan rebels as "the moral equivalent of the founding fathers" is likely to give quite a different picture of the situation. Or consider rival analogies that might be used to explain the role of Saddam Hussein, the Iraqi leader who invaded Kuwait, to someone who knows little about the situation. If one were unsympathetic, one might describe him as analogous to Hitler, producing in the listener a perception of an evil, aggressive figure. On the other hand, if one were sympathetic, one might describe him as being like Robin Hood. This could produce in the listener a perception of a relatively generous figure, redistributing the wealth of the Kuwaitis to the rest of the Arab population.

Not only, then, is perception an integral part of analogy-making, but analogy-making is also an integral part of perception. From this, we conclude that it is impossible to split analogy-making into "first perception, then mapping". The mapping process will often be needed as an important part of the process of perception. The only solution is to give up on any clean temporal division between the two processes, and instead to recognize that they interact deeply.

The modular approach to the modeling of analogy stems, we believe, from a perception of analogical thought as something quite separate from the rest of cognition. One gets the impression from the work of most researchers that analogy-making is conceived of as a special tool in reasoning or problem-solving, a heavy weapon wheeled out occasionally to deal with difficult problems. Our view, by contrast, is that analogy-making is going on constantly in the background of the mind, helping to shape our perceptions of everyday situations. In our view, analogy is not separate from perception: analogy-making itself is a perceptual process.

For now, however, let us accept this view of mapping as a "task" in which representations, the products of the perceptual process, are used. Even in this view, the temporal separation of perception from mapping is, we believe, a misguided effort, as the following argument will demonstrate. This second argument, unlike the previous one, has a scope much broader than just the field of analogy-making. Such an argument could be brought to bear on almost any area within artificial intelligence, demonstrating the necessity for "task-oriented" processes to be tightly integrated with high-level perception.

Consider the implications of the separation of perception from the mapping process, by the use of a separate representation module. Such a module would have to supply a single "correct" representation for any given situation, independent of the context or the task for which it is being used. Our earlier discussion of the flexibility of human representations should already suggest that this notion should be treated with great suspicion. The great adaptability of high-level perception suggests that no module that produced a single context-independent representation could ever model the complexity of the process.

To justify this claim, let us return to the DNA example. To understand the analogy between DNA and a zipper, the representation module would have to produce a representation of DNA that highlights its physical, base-paired structure. On the other hand, to understand the analogy between DNA and source code, a representation highlighting DNA's information-carrying properties would have to be constructed. Such representations would clearly be quite different from each other.

The only solution would be for the representation module to always provide a representation all-encompassing enough to take in *every possible aspect* of a situation. For DNA, for example, we might postulate a single representation incorporating information about its physical, double-helical structure, about the way in which its information is used to build up cells, about its properties of replication and mutation, and much more. Such a representation, were it possible to build it, would no doubt be very large. But its very size would make it far too large for immediate use in processing by the higher-level task-oriented processes for which it was intended—in this case, the mapping module. The mapping processes used in most current computer models of analogy-making, such as SME, all use very small representations that have the relevant information selected and ready for immediate use. For these programs to take as input large representations that include all available information would require a radical change in their design.

The problem is simply that a vast oversupply of information would be available in such a representation. To determine precisely which pieces of that information were relevant would require a complex process of filtering and organizing the available data from the representation. *This process would in fact be tantamount to high-level perception all over again.* This, it would seem, would defeat the purpose of separating the perceptual processes into a specialized module.

Let us consider what might be going on in a human mind when it makes an analogy. Presumably people have somewhere in long-term memory a representation of all their knowledge about, say, DNA. But when a person makes a particular analogy involving DNA, only certain information about DNA is used. This information is brought from long-term memory and probably used to form a temporary active representation in working memory. This second representation will be much less complex, and consequently much easier for the mapping process to manipulate. It seems likely that this smaller representation is what corresponds to the specialized representations we saw used by SME above. It is in a sense a projection of the larger representation from long-term memory—with only the relevant aspects being projected. It seems psychologically implausible that when a person makes an analogy, their working memory is holding all the information from an all-encompassing representation of a situation. Instead, it seems that people hold in working memory only a certain amount of relevant information with the rest remaining latent in long-term storage.

But the process of forming the appropriate representation in working memory is undoubtedly not simple. Organizing a representation in working memory would be another specific example of the action of the high-level perceptual processes—filtering and organization—responsible for the formation of representations in general. And most importantly, this process would necessarily interact with the details of the task at hand. For an all-encompassing representation (in long-term memory) to be transformed into a usable representation in working memory, the nature of the task at hand—in the case of analogy, a particular attempted mapping—must play a pivotal causal role.

The lesson to be learned from all this is that separating perception from the "higher" tasks for which it is to be used is almost certainly a misguided approach. The fact that representations have to be adapted to particular contexts and particular tasks means that an interplay between the task and the perceptual process is unavoidable, and therefore that any "modular" approach to analogy-making will ultimately fail. It is therefore essential to investigate how the perceptual and mapping processes can be integrated.

One might thus envisage a system in which representations can gradually be built up as the various pressures evoked by a given context manifest themselves. We will describe such a system in the next section. In this system, not only is the mapping determined by perceptual processes: the perceptual processes are in turn influenced by the mapping process. Representations are built up gradually by means of this continual interaction between perception and mapping. If a particular representation seems appropriate for a given mapping, then that representation continues to be developed, while the mapping continues to be fleshed out. If the representation seems less promising, then alternative directions are explored by the perceptual process. It is of the essence that the processes of perception and mapping are *interleaved* at all stages. Gradually, an appropriate analogy emerges, based on structured representations that dovetail with the final mapping. We will examine this system in greater detail shortly.

Such a system is very different from the traditional approach, which assumes the representation-building process to have been completed, and which concentrates on the mapping process in isolation. But in order to be able to deal with the great flexibility of human perception and representation, analogy researchers must integrate high-level perceptual processes into their work. We believe that the use of hand-coded, rigid representations will in the long run prove to be a dead end, and that flexible, context-dependent, easily adaptable representations will be recognized as an essential part of any accurate model of cognition.

Finally, we should note that the problems we have outlined here are by no means unique to the modeling of analogical thought. The hand-coding of representations is endemic in traditional AI. Any program that uses pre-built representations for a particular task could be subject to such a "representation module" argument similar to that given above. For most purposes in cognitive science, an integration of task-oriented processes with those of perception and representation will be necessary.

4 A Model that Integrates High-Level Perception with Analogy-Making

A model of high-level perception is clearly desirable, but a major obstacle lies in the way. For any model of high-level perception to get off the ground, it must be firmly founded on a base of low-level perception. But the sheer amount of information available in the real world makes the problem of low-level perception an exceedingly complex one, and success in this area has understandably been quite limited. Low-level perception poses so many problems that for now, the modeling of full-fledged high-level perception of the real world is a distant goal. The gap between the lowest level of perception (cells on the retina, pixels on the screen, waveforms of sound) and the highest level (conceptual processes operating on complex structured representations) is at present too wide to bridge.

This does not mean, however, that one must admit defeat. There is another route to the goal. The real world may be too complex, but if one *restricts the domain*, some understanding may be within our grasp. If, instead of using the real world, one carefully creates a simpler, artificial world in which to study high-level perception, the problems become more tractable. In the absence of large amounts of pixel-by-pixel information, one is led much more quickly to the problems of high-level perception, which can then be studied in their own right.

Such restricted domains, or microdomains, can be the source of much insight. Scientists in all fields throughout history have chosen or crafted idealized domains to study particular phenomena. When researchers attempt to take on the full complexity of the real world without first having some grounding in simpler domains, it often proves to be a misguided enterprise. Unfortunately, microdomains have fallen out of favor in artificial intelligence. The "real world" modeling that has replaced them, while ambitious, has often led to misleading claims (as in the case of

BACON), or to limited models (as we saw with models of analogy). Furthermore, while "real world" representations have impressive labels—such as "atom" or "solar system"—attached to them, these labels conceal the fact that the representations are nothing but simple structures in predicate logic or a similar framework. Programs like BACON and SME are really working in stripped-down domains of certain highly idealized logical forms—their domains merely *appear* to have the complexity of the real world, thanks to the English words attached to these forms.

While microdomains may superficially seem less impressive than "real world" domains, the fact that they are explicitly idealized worlds allows the issues under study to be thrown into clear relief—something that generally speaking is not possible in a full-scale real-world problem. Once we have some understanding of the way cognitive processes work in a restricted domain, we will have made genuine progress towards understanding the same phenomena in the unrestricted real world.

The model that we will examine here works in a domain of alphabetical letter-strings. This domain is simple enough that the problems of low-level perception are avoided, but complex enough that the main issues in high-level perception arise and can be studied. The model, the "Copycat" program (Hofstadter 1984; Mitchell 1990; Hofstadter and Mitchell 1992), is capable of building up its own representations of situations in this domain, and does so in a flexible, context-dependent manner. Along the way, many of the central problems of high-level perception are dealt with, using mechanisms that have a much broader range of application than just this particular domain. Such a model may well serve as the basis for a later, more general model of high-level perception.

This highly parallel, non-deterministic architecture builds its own representations and finds appropriate analogies by means of the continual interaction of perceptual structuring-agents with an associative concept network. It is this interaction between perceptual structures and the concept network that helps the model capture part of the flexibility of human thought. The Copycat program is a model of both high-level perception and analogical thought, and it uses the integrated approach to situation perception and mapping that we have been advocating.

The architecture could be said to fall somewhere on the spectrum between the connectionist and symbolic approaches to artificial intelligence, sharing some of the advantages of each. On the one hand, like connectionist models, Copycat consists of many local, bottom-up, parallel processes from whose collective action higher-level understanding emerges. On the other hand, it shares with symbolic models the ability to deal with complex hierarchically-structured representations.

We shall use Copycat to illustrate possible mechanisms for dealing with five important problems in perception and analogy. These are:

- the gradual building-up of representations;
- the role of top-down and contextual influences;
- the integration of perception and mapping;
- the exploration of many possible paths toward a representation;
- the radical restructuring of perceptions, when necessary.

The description of Copycat given here will necessarily be brief and oversimplified, but further details are available elsewhere (Hofstadter 1984; Mitchell and Hofstadter 1990; Mitchell 1990; Hofstadter and Mitchell 1992).

The Copycat domain

The task of the Copycat program is to make analogies between strings of letters. For instance, it is clear to most people that **abc** and **iijjkkll** share common structure at some level. The goal of the program is to capture this by building, for each string, a representation that highlights this common structure, and by finding correspondences between the two representations.

The program uses the result of this correspondence-making to solve analogy problems of the following form: "If **abc** changes to **abd**, what does **iijjkkll** change to?" Once the program has discovered common structure in the two strings **abc** and **iijjkkll**, deciding that the letter **a** in the first corresponds to the group **ii** in the second and that **c** corresponds to **ll**, it is relatively straightforward for it to deduce that the best answer must be **iijjkkmm**. The difficult task for the program—the part requiring high-level perception—is to build the representations in the first place. We will shortly examine in more detail just how these representations are built.

Before we begin a discussion of the details of Copycat, we should note that the program knows nothing about the shapes of letters, their sounds, or their roles in the English language. It does know the order of the letters in the alphabet, both forwards and backwards (to the program, the alphabet is in no sense "circular"). The alphabet consists of 26 "platonic" letter entities, each with no explicit relation to anything except its immediate neighbors. When instances of these simple concepts, the letters, are combined into strings of various lengths, quite complex "situations" can result. The task of the program is to perceive structure in these situations, and to use this structure to make good analogies.

The architecture used by the program, incidentally, is applicable much more widely than to just the particular domain used here. For instance, the architecture has also been implemented to deal with the problem of perceiving structure and making analogies involving the dinner implements on a tabletop (a microdomain with a more "real world" feel) (French 1988). An application involving perception of the shapes and styles of visual letterforms, and generation of new letterforms sharing the given style, has also been proposed (Hofstadter *et al*, 1987).

Building up representations

The philosophy behind the model under discussion is that high-level perception emerges as a product of many independent but cooperating processes running in parallel. The system is at first confronted with a raw situation, about which it knows almost nothing. Then a number of perceptual agents swarm over and examine the situation, each discovering small amounts of local structure adding incrementally to the system's perception, until finally a global understanding of the situation emerges.

These perceptual agents, called *codelets*, are the basic elements of Copycat's perceptual processing. Each codelet is a small piece of code, designed to perform a particular type of task. Some codelets seek to establish relations between objects; some chunk objects that have been perceived as related into groups; some are responsible for describing objects in particular ways; some build the correspondences that determine the analogy; and there are various others. Each codelet works locally on a small part of the situation. There are many codelets waiting to run at any given time, in a pool from which one is chosen nondeterministically at every cycle. The codelets often compete with each other, and some may even break structures that others have built up, but eventually a coherent representation emerges.

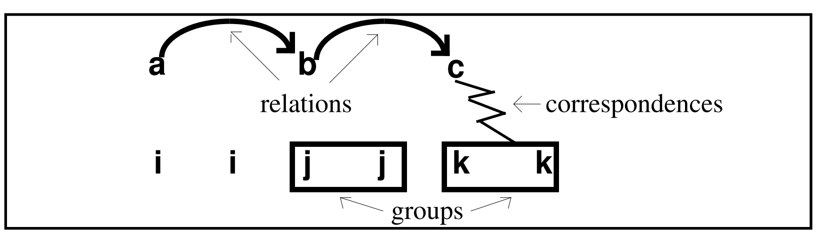

Figure 2. Examples of perceptual structures built by Copycat.

When it starts to process a problem in the letter-string domain, Copycat knows very little about the particular problem at hand. It is faced with three strings, of which it knows only the platonic type of each letter, which letters are spatially adjacent to each other, and which letters are leftmost, rightmost, and middle in each string. The building-up of representations of these strings and of their interrelationships is the task of codelets. Given a string such as **ppqqrrss**, one codelet might notice that the first and second letters are both instances of the same platonic letter-type ("**P**"), and build a "sameness" bond between them. Another might notice that the physically adjacent letters **r** and **s** are in fact alphabetical neighbors, and build a "successor" bond between them. Another "grouping" codelet might chunk the two bonded letters **p** into a group, which can be regarded at least temporarily as a unit. After many such codelets have run, a highly structured representation of the situation emerges, which might, for instance, see the string as a sequence of four chunks of two letters each, with the "alphabetic successor" relation connecting each chunk with its right neighbor. Figure 2 gives an stripped-down example of Copycat's perceptual structuring.

Different types of codelets may come into play at different stages of a run. Certain types of codelets, for example, can run only after certain types of structures have been discovered. In this way, the codelets cause structure to be built up *gradually*, and in a context-sensitive manner. Due to the highly nondeterministic selection of codelets, several directions can be simultaneously explored by the perceptual process. Given the string **abbccd**, for instance, some codelets might try to organize it as a sequence of "sameness" groups, **a-bb-cc-d**, while others might simultaneously try to organize it quite differently as a sequence of "successor" groups, **ab-bc-cd**. Eventually, the program is likely to focus on one or the other of these possibilities, but because of the nondeterminism, no specific behavior can be absolutely guaranteed in advance. However, Copycat usually comes up in the end with highly structured and cognitively plausible representations of situations it is given.

The role of context and top-down influences

As we have seen, one of the most important features of high-level perception is its sensitivity to context. A model of the perceptual process that proceeds in a manner that disregards context will necessarily be inflexible.

The Copycat model captures the dependence of perception on contextual features by means of an associative concept-network (Figure 3), the *Slipnet*, which interacts continually with the perceptual process. Each node in this network corresponds to a concept that might be relevant in the letter-string domain, and each node can be activated to a varying degree depending on the perceived relevance of the corresponding concept to the given situation. As a particular feature of the situation is noted, the node representing the concept that corresponds to the feature is activated in the concept network. In turn, the activation of this concept has a biasing effect on the perceptual processing that follows. Specifically, it causes the creation of some number of associated

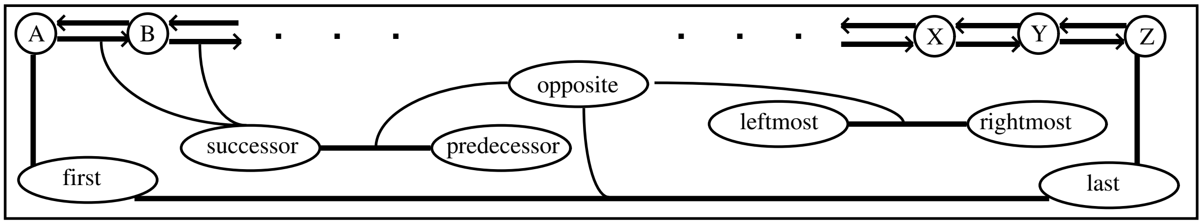

Figure 3. A small portion of Copycat's concept network, the Slipnet.

codelets, which are placed in the pool of codelets waiting to run. For instance, if the node corresponding to the concept of "alphabetic successor" is activated in the Slipnet, then several codelets will be spawned whose task is to look for successorship relations elsewhere in the situation.

Further, the activation of certain nodes means that it is more likely that associated perceptual processes will succeed. If the "successor" node is highly active, for example, not only is it more likely that codelets that try to build successorship relations will be *spawned*, but it is also more likely that once they run, they—rather than some competing type of codelet—will *succeed* in building a lasting relation as part of the representation. In both of these ways, perceptual processing that has already been completed can have a contextual, top-down influence on subsequent processing through activation of concepts in the Slipnet.

For instance, in the string **kkrrtt** it is likely that the two **r**'s will be perceived as a "sameness group" (a group all of whose members are the same); such a perception will be reinforced by the presence of two similar groups on either side, which will activate the node representing the concept of "sameness group". On the other hand, in the string **abcijkpqrrst**, the presence of the groups **abc** and **ijk** will cause the node representing "successor group" (a group consisting of alphabetically successive letters) to be active, making it more likely that **pqr** and **rst** will be perceived in the same way. Here, then, it is more likely that the two adjacent **r**'s will be perceived separately, as parts of two different "successor groups" of three letters each. The way in which two neighboring **r**'s are perceived (i.e., as grouped or not) is highly dependent on the context that surrounds them, and this contextual dependence is mediated by the Slipnet.

This two-way interaction between the perceptual process and the concept network is a combination of top-down and bottom-up processing. The perceptual work performed by the codelets is an inherently bottom-up process, achieved by competing and cooperating agents each of which acts locally. The Slipnet, however, by modulating the action of the codelets, acts as a top-down influence on this bottom-up process. The Slipnet can thus be regarded as a dynamic controller, allowing global properties such as the activation of concepts to influence the local action of perceptual agents. This top-down influence is vitally important, as it it ensures that perceptual processes do not go on independently of the system's understanding of the global context.

Integration of perception and mapping in analogy-making

We have already discussed the necessity of a fully-integrated system of perceptual processing and mapping in analogy-making. The Copycat model recognizes this imperative. The situation-perception and mapping processes take place simultaneously. Certain codelets are responsible for building up representations of the given situations, while others are responsible for building up a mapping between the two. Codelets of both types are in the pool together.

In the early stages of a run, perceptual codelets start to build up representations of the individual situations. After some structure has been built up, other types of codelets begin to make tentative mappings between the structures. From then on, the situation-perception and mapping processes proceed hand in hand. As more structure is built within the situations, the mapping becomes more sophisticated, and aspects of the evolving mapping in turn exert pressure on the developing perceptions of the situations.

Consider, for example, two analogies involving the string **ppqrss**. If we are trying to find an analogy between this and, say, the string **aamnxx**, then the most successful mapping is likely to map the group of **p**'s to the group of **a**'s, the group of **s**'s to the group of **x**'s, and **qr** to the successor group **mn**. The most natural way to perceive the second string is in the form **aa-mn-xx**, and this in turn affects the way that the first string is perceived, as three two-letter groups in the form **pp-qr-ss**. On the other hand, if we are trying to find an analogy between **ppqrss** and the string **aijklx**, then recognition of the successor group **ijkl** inside the latter string is likely to arouse perceptual biases toward seeking successor relations and groups, so that the system will be likely to spot the successor group **pqrs** within **ppqrss**, and to map one successor group to the other. This leads to the original string being perceived as **p-pqrs-s**, which maps in a natural way to **a-ijkl-x**.

Thus we can see that different mappings act as different contexts to evoke quite different perceptions of the same string of letters. This is essentially what was going on in the two analogies described earlier involving DNA. In both cases, the representation of a given situation is made not in isolation, but under the influence of a particular mapping.

We should note that the Copycat model makes no important distinction between structures built for the purpose of situation-perception (such as bonds between adjacent letters, or groups of letters), and those built for the purpose of mapping (such as correspondences between letters or groups in the two strings). Both types of structure are built up gradually over time, and both contribute to the program's current understanding of the overall situation. The mapping structures can themselves be regarded as perceptual structures: the mapping is simply an understanding of the analogy as a whole.

Exploring different paths and converging on a solution

A model of perception should, in principle, be able to explore all of the different plausible ways in which a situation might be organized into a representation. Many representations may be possible, but some will be more appropriate than others. Copycat's architecture of competing codelets allows for the exploration of many different pathways toward a final structure. Different codelets will often begin to build up structures that are incompatible with each other. This is good—it is desirable that many possibilities be explored. In the end, however, the program must converge on one particular representation of a given situation.

In Copycat, the goal of homing in on a particular solution is aided by the mechanism of *computational temperature*. This is a number that measures the amount and quality of structure present in the current representation of the situation. Relevant structures here include bonds, groups, and correspondences, as well as some others. The term "quality of structure" refers to how well different parts of the structure cohere with each other. Computational temperature is used to control the amount of randomness in the local action of codelets. If a large amount of good structure has been built up, the temperature will be low and the amount of randomness allowed will be small. Under these circumstances, the system will proceed in a fairly deterministic way, meaning that it sticks closely to a single pathway with few rival side-explorations being considered. On the other hand, if there is little good structure, the temperature will be high, which will lead to diverse random explorations being carried out by codelets.

At the start of a run, before any structure has been built, the temperature is maximally high, so the system will behave in a very random way. This means that many different pathways will be explored in parallel by the perceptual processes. If no promising structural organization emerges, then the temperature will remain high and many different possibilities will continue to be explored. Gradually, in most situations, certain structures will prove more promising, and these are likely to form the basis of the final representation. At any given moment, a single structural view is dominant, representing the system's current most coherent worldview, but many other tentative structures may be present in the background, competing with it.

As good structures build up, the temperature gradually falls and so the system's exploratory behavior becomes less random. This mean that structures that have already been built have a lowered chance of being replaced by new ones and are thus favored. The more coherent a global structure, the less likely parts of it are to be broken. As structure builds up and temperature falls, the system concentrates more and more on developing the structure that exists. Eventually, the program will converge on a good representation of a given situation. In practice, Copycat frequently comes up in different runs with different representations for the same situation, but these representations usually seem to be cognitively plausible. Its final "solutions" to various analogy problems are distributed in a fashion qualitatively similar to the distributions found with human subjects (Mitchell 1990; Hofstadter and Mitchell 1992).

The process of exploring many possibilities and gradually focusing on the most promising ones has been called a "parallel terraced scan" (Hofstadter 1984; Hofstadter and Mitchell 1992). The process is akin to the solution to the "two-armed bandit" problem (Holland 1975) where a gambler has access to two slot machines with fixed but distinct probabilities of payoff. These payoff probabilities are initially unknown to the gambler, who wishes to maximize payoffs over a series of trials. The best strategy is to start by sampling both machines equally, but to gradually focus one's resources probabilistically on the machine that appears to be giving the better payoff. The Copycat program has to perform an analogous task. To function flexibly, it has to sample many representational possibilities and choose those that promise to lead to the most coherent worldview, gradually settling down to a fixed representation of the situation. In both the two-armed bandit and in Copycat, it takes time for certain possibilities to emerge as the most fruitful, and a biased stochastic sampling technique is optimal for this purpose.

Radical restructuring

Sometimes representations that have been built up for a given situation turn out to be inappropriate, in that they do not lead to a solution to the problem at hand. When people find themselves in this situation, they need to be able to completely restructure their representations, so that new ones can evolve that are more adequate for the current task. Maier's two-string experiment provides an example of radical restructuring; people have to forget about their initial representation of a pair of pliers as a tool for bending things, and instead see it as a heavy weight.

In Copycat, when a representation has been built up, the temperature has necessarily gone down, which makes it difficult to change to another representation. But it is obviously not advantageous for the program to keep a representation that does not led to a solution. For this reason, the program has a special set of mechanisms to deal with such situations.

For instance, when the program is given the analogy "If **abc** changes to **abd**, what does **xyz** change to?", it usually builds a detailed representation of **abc** and **xyz** as successor groups, and quite reasonably maps one string to the other accordingly (**a** maps to **x**, **c** maps to **z**, etc). But now, when it tries to carry out the transformation for **xyz** that it feels is analogous to **abc** becoming **abd**, it finds itself blocked, since it is impossible to take the successor of the letter **z** (the Copycat alphabet is non-circular). The program has hit a "snag"; the only way to deal with it is to find an alternative representation of the situation.

The program deals with the problem firstly by *raising the temperature*. The temperature shoots up to its maximal value. This produces a great deal of randomness at the codelet level. Secondly, "breaker codelets" are brought in for the express purpose of destroying representations. The result is that many representations that have been carefully built up are broken down. At the same time, much activation is poured into the concept representing the source of the snag—the concept **Z**—and much perceptual attention is focused on the *specific* **z** inside **xyz** (that is, it becomes very salient and attracts many codelets). This causes a significant change in the representation-building process the second time around. To make a long story short, the program is thereby able to come up with a completely new representation of the situation, where **abc** is still perceived as a successor group, but **xyz** is re-perceived as a *predecessor* group, starting from **z** and going backwards. Under this new representation, the **a** in the first string is mapped to the **z** in the second.

Now if the program attempts to complete its task, it discovers that the appropriate transformation on **xyz** is to take the *predecessor* of the *leftmost* letter, and it comes up with the insightful answer **wyz**. (We should stress that the program, being nondeterministic, does not always or even consistently come up with this answer. The answer **xyd** is actually given more often than **wyz**.) Further details are given by Mitchell and Hofstadter (1990).

This process of reperception can be regarded as a stripped-down model of a "scientific revolution" (Kuhn 1970) in a microdomain. According to this view, when a field of science comes up against a problem it cannot solve, clamor and confusion result in the field, culminating in a "paradigm shift" where the problem is viewed in a completely different way. With the new worldview, the problems may be straightforward. The radical restructuring involved in the above letter-string problem seems quite analogous to this scientific process.

What Copycat doesn't do

Some have argued that in employing hand-coded mechanisms such as codelets and the Slipnet, Copycat is guilty of 20-20 hindsight in much the same fashion as BACON and SME. But there is a large difference: BACON and SME use fixed *representations*, whereas Copycat develops flexible representations using fixed *perceptual mechanisms*. Whereas we have seen that the use of fixed representations is cognitively implausible, it is clear that human beings at any given time have a fixed repertoire of mechanisms available to the perceptual process. One might justifiably ask where these mechanisms, and the corresponding mechanisms in Copycat, come from, but this would be a question about *learning*. Copycat is not intended as a model of learning: its performance, for instance, does not improve from one run to the next. It would be a very interesting further step to incorporate learning processes into Copycat, but at present the program should be taken as a model of the perceptual processes in an individual agent at a particular time.

There are other aspects of human cognition that are not incorporated into Copycat. For instance, there is nothing in Copycat that corresponds to the messy low-level perception that goes on in the visual and auditory systems. It might well be argued that just as high-level perception exerts a strong influence on and is intertwined with later cognitive processing, so low-level perception is equally intertwined with high-level perception. In the end, a complete model of high-level perception will have to take low-level perception into account, but for now the complexity of this task means that key features of the high-level perceptual processes must be studied in isolation from their low-level base.

The Tabletop program (French and Hofstadter 1991; French 1992) takes a few steps towards lower-level perception, in that it must make analogies between visual structures in a two-dimensional world, although this world is still highly idealized. There is also a small amount of related work in AI that attempts to combine perceptual and cognitive processes. It is interesting to note that in this work, microdomains are almost always used. Chapman's "Sonja" program (Chapman 1991), for instance, functions in the world of a video game. Starting from simple graphical information, it develops representations of the situation around it and takes appropriate action. As in Tabletop, the input to Sonja's perceptual processes is a little more complex than in Copycat, so that these processes can justifiably be claimed to be a model of "intermediate vision" (more closely tied to the visual modality than Copycat's high-level mechanisms, but still abstracting away from the messy low-level details), although the representations developed are less sophisticated than Copycat's. Along similar lines, Shrager (1990) has investigated the central role of perceptual processes in scientific thought, and has developed a program that builds up representations in the domain of understanding the operation of a laser, starting from idealized two-dimensional inputs.

5 Conclusion

It may sometimes be tempting to regard perception as not truly "cognitive", something that can be walled off from higher processes, allowing researchers to study such processes without getting their hands dirtied by the complexity of perceptual processes. But this is almost certainly a mistake. Cognition is infused with perception. This has been recognized in psychology for decades, and in philosophy for longer, but artificial-intelligence research has been slow to pay attention.

Two hundred years ago, Kant provocatively suggested an intimate connection between concepts and perception. “Concepts without percepts”, he wrote, “are empty; percepts without concepts are blind.” In this paper we have tried to demonstrate just how true this statement is, and just how dependent on each other conceptual and perceptual processes are in helping people make sense of their world.

“Concepts without percepts are empty.” Research in artificial intelligence has often tried to model concepts while ignoring perception. But as we have seen, high-level perceptual processes lie at the heart of human cognitive abilities. Cognition cannot succeed without processes that build up appropriate representations. Whether one is studying analogy-making, scientific discovery, or some other area of cognition, it is a mistake to try to skim off conceptual processes from the perceptual substrate on which they rest, and with which they are tightly intermeshed.

“Percepts without concepts are blind.” Our perception of any given situation is guided by constant top-down influence from the conceptual level. Without this conceptual influence, the representations that result from such perception will be rigid, inflexible, and unable to adapt to the problems provided by many different contexts. The flexibility of human perception derives from constant interaction with the conceptual level. We hope that the model of concept-based perception that we have described goes some way towards drawing these levels together.

Recognizing the centrality of perceptual processes makes artificial intelligence more difficult, but it also makes it more interesting. Integrating perceptual processes into a cognitive model leads to flexible representations, and flexible representations lead to flexible actions. This is a fact that has only recently begun to permeate artificial intelligence, through such models as connectionist networks, classifier systems, and the architecture presented here. Future advances in the understanding of cognition and of perception are likely to go hand in hand, for the two types of process are inextricably intertwined.

References

Bruner, J. (1957). On perceptual readiness. *Psychological Review*, 64: 123-152.

Burstein, M. (1986). Concept formation by incremental analogical reasoning and debugging. In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell (eds.), *Machine learning: An artificial intelligence approach,* Vol. 2 (Los Altos, CA: Morgan Kaufmann).

Carbonell, J. G. (1986). Learning by analogy: Formulating and generalizing plans from past experience. In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell (eds.), *Machine learning: An artificial intelligence approach,* Vol. 2 (Los Altos, CA: Morgan Kaufmann).

Chapman, D. (1991). *Vision, instruction, and action.* Cambridge, MA: MIT Press.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14: 179-212.

Evans, T. G. (1968). A program for the solution of a class of geometric-analogy intelligence-test questions. In M. Minsky (ed.), *Semantic information processing* (Cambridge, MA: MIT Press).

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1990). The structure-mapping engine. *Artificial Intelligence*, 41: 1-63.

Fodor, J. A. (1983). *The modularity of mind* (Cambridge, MA: MIT Press).

French, R. M., and Hofstadter, D. R. (1991). Tabletop: A stochastic, emergent model of analogy-making. *Proceedings of the 13th annual conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum.

French, R. M. (1992). Tabletop: A stochastic, emergent model of analogy-making. Doctoral dissertation, University of Michigan.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2).

Hofstadter, D. R. (1984). The Copycat project: An experiment in nondeterminism and creative analogies. M.I.T. AI Laboratory memo 755.

Hofstadter, D. R., and Mitchell, M. (1992). An overview of the Copycat project. In K. J. Holyoak and J. Barnden (eds.), *Connectionist Approaches to Analogy, Metaphor, and Case-Based Reasoning* (Norwood, NJ: Ablex).

Hofstadter, D. R., Mitchell, M., and French, R. M. (1987). Fluid concepts and creative analogies: A theory and its computer implementation. CRCC Technical Report 18, Indiana University.

Holland, J. H. (1975). *Adaptation in natural and artificial systems* (Ann Arbor, MI: University of Michigan Press).

Holland, J. H. (1986). Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell (eds.), *Machine learning: An artificial intelligence approach,* Vol. 2 (Los Altos, CA: Morgan Kaufmann).

Holyoak, K. J. and Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13: 295-355.

James, W. (1890). *The principles of psychology* (Henry Holt & Co).

Kedar-Cabelli, S. (1988). Towards a computational model of purpose-directed analogy. In A. Prieditis (ed.), *Analogica* (Los Altos, CA: Morgan Kaufmann).

Kuhn, T. (1970). *The structure of scientific revolutions* (2nd edition) (Chicago: University of Chicago Press).

Lakoff, G. (1987). *Women, fire and dangerous things* (Chicago: University of Chicago Press).

Langley, P., Simon, H. A., Bradshaw, G. L., and Zytkow, J. M. (1987). *Scientific discovery: Computational explorations of the creative process* (Cambridge, MA: MIT Press).

Maier, N. R. F (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Cognitive Psychology*, 12: 181-194.

Marr, D. (1977). Artificial intelligence—a personal view. *Artificial Intelligence*, 9: 37-48.

Mitchell, M. (1990). Copycat: A computer model of high-level perception and conceptual slippage in analogy-making. Doctoral dissertation, University of Michigan.

Mitchell, M., and Hofstadter, D. R. (1990). The emergence of understanding in a computer model of concepts and analogy-making. *Physica D*, 42: 322-334.

Newell, A. and H. A. Simon (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 19: 113-126.

Pylyshyn, Z. (1980). Cognition and computation. *Behavioral and Brain Sciences*, 3: 111-132.

Qin, Y., and Simon, H. A. (1990). Laboratory replication of scientific discovery processes. *Cognitive Science* 14: 281-310.

Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986). *Parallel distributed processing* (Cambridge, MA: MIT Press).

Shrager, J. (1990). Commonsense perception and the psychology of theory formation. In J. Shrager and P. Langley (eds.) *Computational models of scientific discovery and theory formation* (San Mateo, CA: Morgan Kaufmann).

Simon, H. A. (1989). The scientist as problem solver. In D. Klahr and K. Kotovsky (eds.) *Complex information processing* (Hillsdale, NJ: Lawrence Erlbaum).

Winston, P. H. (1982). Learning new principles from precedents and exercises. *Artificial Intelligence* 19: 321-350.