Chapter 4

High-level Perception, Representation, and Analogy:

A Critique of Artificial-intelligence Methodology

DAVID CHALMERS, ROBERT FRENCH,
and DOUGLAS HOFSTADTER

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 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 the understanding of 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. In support of this claim, we will examine some existing models of scientific discovery and analogical thought, 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 points the way to alternative architectures (which are discussed in the next several chapters of this book).

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 *relations.* 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 bowtie, 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.

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 pre-packaged as slots and fillers, and organizing them into a coherent structure is likely to be a highly nontrivial 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 apprehended 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 (James, 1890, pages 222–224):

> 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 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, page 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 (Marr, 1977, page 44) 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.

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 multidimensional 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 and his colleagues (Holland *et al.*, 1986), 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 be able to 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 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 of electrical resistance.

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, page 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 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 university students starting with the data that BACON was given could make essentially the same "discoveries" within an hour-long period. 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.

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 *essences* 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 helps to flesh 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 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 SME, the Structure Mapping Engine (Falkenhainer, Forbus, & 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 (1983), Holyoak & Thagard (1989), Kedar-Cabelli (1988a), 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 IV–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 IV–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 the analogy-making task has already been accomplished.

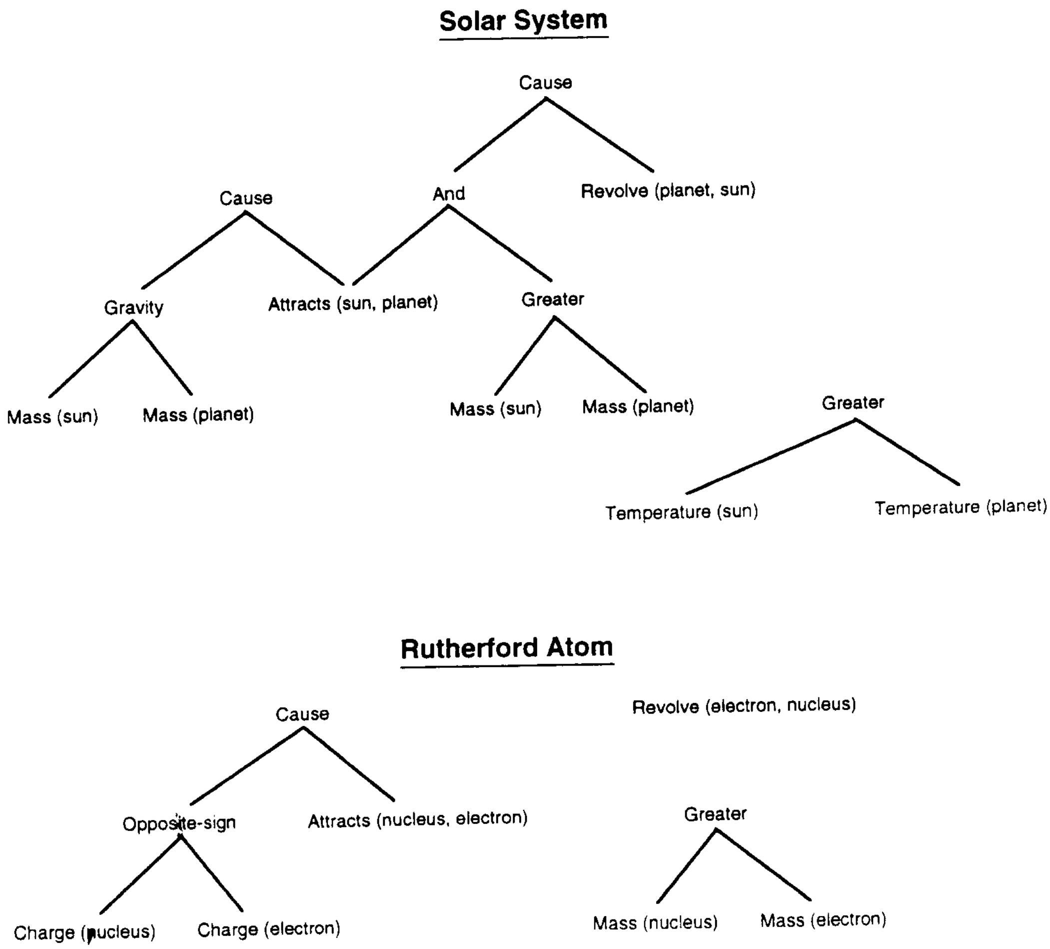

Figure IV–1. Predicate-calculus representations of two situations between which the SME program constructs a mapping.

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 is 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, never to a 4-place predicate, no matter how semantically close the predicates might be. So it is vitally important for SME that every relation be represented by precisely the right kind of predicate structure in every representation. It seems unlikely, however, 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 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.[1]

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 (1988a) 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 (1983), and Winston (1982), 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 fusing 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.

1. Disclaimer: Since this article was written, Ken Forbus, one of the authors of SME, has worked on modules that build representations in "qualitative physics". Some work has also been done on using these representations as input to SME. Despite this conceptual advance in the architecture, the two processes are still not intertwined and interdependent, as we claim they are in human thought.

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, leaving 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, though, some senior figures in Christian churches in America had a quite 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" (as Reagan once characterized them) 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 superfluous 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 now and then to deal with especially tough 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 the time being, however, suppose we 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 an isolated 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 allow the full system 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, 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 very 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 are gradually built up as the various pressures evoked by a given context manifest themselves. In such a system, not only would the mapping be determined by perceptual processes, but the perceptual processes would in turn be influenced by the mapping process. Representations would be built up gradually by means of this continual interaction between perception and mapping. If a particular representation seemed appropriate for a given mapping, then that representation would continue to be developed while the mapping continued to be fleshed out. If the representation seemed less promising, then alternative directions would be explored by the perceptual process. It would be of the essence that the processes of perception and mapping be *interleaved* at all stages. Gradually, an appropriate analogy would emerge, based on structured representations that dovetail with the final mapping.

In fact, two such systems have been built in our research group (they are described in the next few chapters of this book). Such systems are 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.

The Utility of Small Domains

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.

In our research group, we have built two systems along the lines sketched above. One of them, the Copycat program (see Chapter 5 as well as Mitchell, 1993), 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.

Copycat 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.

Copycat's highly parallel and nondeterministic 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 of Copycat could be said to fall somewhere along the spectrum stretching 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.

Copycat illustrates 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 successful implementation of Copycat amounts to a feasibility proof that these ideas actually can be made to work.

The architectural core of Copycat, incidentally, is applicable much more widely than to just the particular domain in which Copycat functions. For instance, substantially the same architecture has been used in the Tabletop program to deal with the problem of perceiving structure and making analogies involving configurations of silverware and crockery on a tabletop — a micro-domain with a more "real world" feel (see Chapters 8 and 9, as well as French, 1995). Additionally, Letter Spirit, a related architecture involving perception of the shapes and styles of visual letterforms, as well as generation of new letterforms sharing an abstract style suggested by other letterforms, has been proposed and partially implemented (see Chapter 10).

There is certainly nothing in Copycat or these other projects that corresponds to the messy low-level perception that goes on in the visual and auditory systems. It might well be argued by a skeptic, however, that just as we have insisted herein that 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. This point is well taken. We acknowledge that in the end, a complete model of high-level perception will have to take all levels of low-level perception into account as well, but we believe that for now, the complexity of this task is such that key features of the high-level perceptual processes must be studied in isolation from their low-level base.

The Tabletop program takes a few steps further than Copycat does 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. Letter Spirit goes even further toward low-level vision, but still stops far short of rock bottom.

We are familiar with a couple of AI projects that attempt 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.

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 models of concept-based perception that we have developed go 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 Copycat/Tabletop architecture and its generalizations. 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.