Metaphor
and Common-Sense Reasoning

Jaime G Carbonell and Steven Minton
Carnegie-Mellon University
Pittsburgh, PA 15213

5 March 1983

Abstract

Inferences based on metaphors appear to play a major role in human common sense reasoning. This paper identifies and analyzes general inference patterns based upon underlying metaphors, in particular the pervasive *balance principle*. Strategies for metaphor comprehension are explored, and *analogical mapping structures* are proposed as a means of representing metaphorical relationships between domains. In addition, a framework for a computational model embodying principles of metaphorical common sense reasoning is discussed.[1]

Metaphor and Common Sense Reasoning

Table of Contents

Metaphor and Common Sense Reasoning

Jaime G. Carbonell and Steven Minton

1. Introduction

The theory that metaphor dominates large aspects of human thinking, as well playing a significant role in linguistic communication, has been argued with considerable force [26, 24, 8, 5]. However, the validity of such a theory is a matter of continuing debate that appears neither to dissuade its proponents nor convince its detractors. Being among the proponents, we propose to develop a computational reasoning system for performing metaphorical inferences. If such a system exhibits cognitively plausible common sense reasoning capabilities, it will demonstrate, at the very least, the utility of metaphorical inference in modeling significant aspects of naive human reasoning. This paper reviews our initial steps towards the development of a computational model of metaphor-based reasoning.

2. Experiential Reasoning vs Formal Systems

Humans reason and learn from experience to a degree that no formal system, AI model, or philosophical theory has yet been able to explain. The statement that the human mind is (or contains) the sum total of its experiences is in itself rather vacuous. A more precise formulation of experience-based reasoning must be structured in terms of coordinated answers to the following questions: *How* are experiences brought to bear in understanding new situations? *How* is long term memory modified and indexed? *How* are inference patterns acquired in a particular domain and adapted to apply in novel situations? *How* does a person "see the light" when a previously incomprehensible problem is viewed from a new perspective? *How* are the vast majority of irrelevant or inappropriate experiences and inference patterns filtered out in the understanding process? Answering all these "how" questions requires a *process model* capable of organizing large amounts of knowledge and mapping relevant aspects of past experience to new situations. Some meaningful starts have been made towards large-scale episodic-based memory organization [32, 33, 34, 28, 25]and towards episodic-based analogical reasoning [9, 12, 7]. Bearing these questions in mind, we examine the issue of common sense reasoning in knowledge-rich mundane domains.

Our central hypothesis is:

> **Experiential reasoning hypothesis:** *Reasoning in mundane, experience-rich recurrent situations is qualitatively different from formal, deductive reasoning evident in more abstract, experimentally contrived, or otherwise non-recurrent situations (such as some mathematical or puzzle-solving domains).*

In the statement of our hypothesis we do not mean to exclude experience-rich metaphorical inference from scientific or mathematical thought. Rather, we claim that formal deductive inference is definitely not the dominant process in mundane reasoning. In essence, the experiential reasoning hypothesis states that structuring new information according to relevant past experience is an important aspect of human comprehension -- perhaps more important than other aspects studied thus far in much greater depth.

Common-sense experience-rich reasoning consists of recalling appropriate past experiences and inference patterns, whereas solving abstract problems divorced from real-world experience requires knowledge-poor search processes more typical of past and present AI problem solving systems. Since computer programs perform much better in simple, elegant, abstract domains than in "scruffy" experience-rich human domains, it is evident that a fundamental reasoning mechanism is lacking from the AI repertoire. The issue is not merely that AI systems lack experience in mundane human scenarios -- they would be unable to benefit from such experience if it were encoded in their knowledge base. We postulate that the missing reasoning method is based on the transfer of proven inference patterns and experiential knowledge across domains. This is not to say that humans are incapable of more formal reasoning, but rather that such reasoning is seldom necessary, and when applied it requires a more concerted cognitive effort than mundane metaphorical inference.

There is evidence that human expertise, far beyond what we would label common sense reasoning, draws upon past experience and underlying analogies. For instance, the master chess player is not a better deductive engine than his novice counterpart. Rather, as Chase and Simon [14] have shown, he commands a vast repertoire of chess-board patterns and associated strategies that comprise his past experience. And, when encountering a new chessboard situation he uses the relevant patterns (which may only partially match the current position) to index the appropriate knowledge. Mechanics problems in physics are often solved by creating a simple mental model -- an analog of the real situation -- that preserves the significant properties. The model, created on the basis of past experience solving similar problems, is then used to instantiate one or more well known principles of physics in a familiar manner and thereby obtain a solution [27, 15, 10].

People's well-developed ability to perform analogical reasoning is at least partly responsible for what we call "common-sense" reasoning. Roughly speaking, *analogical reasoning* is the process by which one recognizes that a new situation is similar to some previously encountered situation, and uses the relevant prior knowledge to structure and enrich one's understanding of the new situation. We refer to *metaphorical reasoning* as that subset of analogical reasoning in which the analogy is explicitly stated or otherwise made evident to the understander. For instance, comprehending "John

is an encyclopedia" entails metaphorical reasoning since the analogy between John and an encyclopedia is explicitly suggested. However, constructing a novel analogy in order to explain some new situation is a different task which requires searching memory for a previously encountered similar situation. Both of these forms of inference may be labeled common-sense reasoning in so far as they require access to large amounts of past knowledge and reaching conclusions without benefit of formal deduction.

3. Patterns of Metaphorical Inference

A metaphor, simile or analogy can be said to consist of 3 parts: a *target*, a *source* and an *analogical mapping*. For example:

> John was embarrassed. His face looked like a beet.

Here the target is "John's face" and the source is "a beet". The analogical mapping transmits information from the source to the target domain. In this case, the mapping relates the color of John's face to the color of a beet. Our use of the same terminology to describe metaphors, similes and analogies reflects our opinion that they are all merely different linguistic manifestations of the same underlying cognitive process: analogical reasoning. That is, they differ primarily in their form of presentation rather than in their internal structure. Consequently, although our choice of terminology may indicate that we are centrally concerned with the phenomenon of metaphor, we mean to include simile and analogy as well.

3.1. The Balance Principle

Consider a prevalent metaphor: reasoning about imponderable or abstract entities as though they were objects with a measurable weight. One of several reasoning patterns based on this simple metaphor is the *balance principle*. The physical analog of this reasoning pattern is a prototypical scale with two balanced plates. Large numbers of metaphors appeal to this simple device coupled with the processes of bringing the system into (and out of) equilibrium. First, consider some examples of the basic metaphor, in which *the relevant aspect of an abstract concept maps onto the weight[2] of an unspecified physical object.*

> Arms control is a *weighty* issue.

> The worries of a nation *weigh heavily* upon his shoulders.

> The Argentine air force launched a *massive* attack on the British fleet. One frigate was *heavily* damaged, but only *light* casualties were suffered by British sailors. The Argentines payed a *heavy* toll in downed aircraft.

---

[2] Mass is virtually synonymous with weight in naive reasoning.

Not being in the mood for *heavy* drama, John went to a *light* comedy, which turned out to be a piece of meaningless *fluff*.

Pendergast was a real *heavyweight* in the 1920s Saint Louis political scene.

The crime *weighed heavily* upon his conscience.

The *weight* of the evidence was overwhelming.

3.2. The Physical Metaphor Hypothesis

Weight clearly represents different things in the various metaphors: the severity of a nation's problems, the number of attacking aircraft, the extent of physical damage, the emotional affect on audiences of theatrical productions, the amount of political muscle (to use another metaphor), the reaction to violated moral principles, and the degree to which evidence is found to be convincing. In general, more is heavier; less is lighter. One may argue that since language is heavily endowed with words that describe weight, mass and other physical attributes (such as height and orientation [26, 8]), one borrows such words when discussing more abstract entities [1] for lack of alternate vocabulary. Whereas this argument is widely accepted, it falls far short of the conjecture we wish to make.

> **Physical metaphor hypothesis:** *Physical metaphors directly mirror the underlying inference processes. Inferences patterns valid for physical attributes are used via the analogical mapping to generate corresponding inferences in the target domain.*

In order to illustrate the validity of this hypothesis, consider a common inference pattern based on the weight of physical objects: The inference pattern is the *balance principle* mentioned earlier as applied to a scale with two plates. The scale can be in balance or tipped towards either side, as a function of the relative weights of objects placed in the respective plates. Inference consists of placing objects in the scale and predicting the resultant situation -- no claim is made as to whether this process occurs in a propositional framework or as visual imagery, although we favor the former. How could such a simple inference pattern be useful? How could it apply to complex, non-physical domains? Consider the following examples of metaphorical communication based on this inference pattern:

> The jury found the *weight* of the evidence favoring the defendant. His impeccable record *weighed heavily* in his favor, whereas the prosecution witness, being a confessed con-man, carried *little weight* with the jury. *On balance* the state failed to *amass* sufficient evidence for a *solid* case.

> The SS-20 missile *tips the balance* of power in favor of the Soviets.

> Both conservative and liberal arguments appeared to *carry equal weight* with the president, and his decision *hung on the balance*. However, his long-standing opposition to abortion *tipped the scale* in favor of the conservatives.

> The Steelers were the *heavy* pre-game favorites, but the Browns started *piling up* points and accumulated a *massive* half-time lead. In spite of a late rally, the Steelers did not score *heavily* enough to pull the game out.

> The job applicant's shyness *weighed* against her, but her excellent recommendations *tipped the scales* in her favor.

In each example above the same basic underlying inference pattern recurs, whether representing the outcome of a trial, statements of relative military power, decision-making processes, or the outcome of a sporting event. The inference pattern itself is quite simple: it takes as input signed quantities -- whose magnitudes are analogous to their stated "weight" and whose signs depend on which side of a binary issue those weights correspond -- and selects the side with the maximal weight, computing some qualitative estimate of how far out of balance the system is. Moreover, the inference pattern also serves to infer the rough weight of one side if the weight of the other side and the resultant balance state are known. (E.g., If Georgia won the football game scoring only 14 points, Alabama's scoring must have been *really light*)

Our point is that this very simple inference pattern accounts for large numbers of inferences in mundane human situations. Given the existence of such a simple and widely applicable pattern, why should one suppose that more complicated inference methods explain human reasoning more accurately? We believe that there exist a moderate number of general inference patterns such as the present one which together span a large fraction of mundane human reasoning situations. Moreover, the few other patterns we have found thus far are also rooted on simple physical principles or other directly experienced phenomena.

4. The Role of Metaphor in Common-Sense Reasoning

Recently one of us developed a model of analogical problem solving [12, 11] based on the principle that past experience in solving particular classes of problems should play a central role in solving new problems of a similar nature. At the risk of oversimplification, analogical problem solving can be summarized as a four-stage process:

1. Recalling one or more past problems that bear strong similarity to the new problem.

2. Constructing a mapping from the old problem solution process into a solution process for the new problem, exploiting known similarities in the two problem situations.

3. Instantiating, refining and testing the potential solution to the new problem.

4. Generalizing recurring solution patterns into reusable plans for common types of problems.

The analogical problem solving model finessed most of the issues in building a computationally effective mechanism to recall similar problem solving episodes from memory, but suggested a model built along the lines of Schank's MOPS [34, 28, 25], utilizing relative invariance measures [8] as a memory organization principle.

Most of the "action" in the analogical problem solving model dealt with the issue of constructing a mapping that would transfer, modify, and augment a solution from a similar past problem situation to satisfy the requirements of the new problem situation. Here, we propose that the role of metaphors is to capture and communicate mappings from well known experiential domains to new, less structured domains. These mappings often fail to provide deep insight into the target phenomenon they seek to explain, but do provide easily and quickly a shallow level of understanding sufficient for most common, everyday purposes. For instance, stating that the prosecutor's evidence was counterbalanced by the defendant's alibi gives us all a "feel" for the present state of the trial. But, consider a situation where you, the reader, must step in for the temporarily ill prosecuting attorney. Suddenly, your understanding of the trial is woefully inadequate. Questions arise such as: "Just how did the defense witness counter the prosecutions evidence? Did it undermine the credibility of our witnesses? Did it beef up the defendant's story? Did new evidence surface? Does the case now hinge upon a possible breakdown in our chain of evidence?" (to use other metaphors). Deeper reasoning about a particular topic may well be metaphorically based, or more deductive in nature. Whether or not a particular metaphor provides more than a casual level of understanding depends on the validity of the source as a model, as well as the structure of the mapping [20]. In the next section we will examine further the close relationship between mappings and metaphorical understanding.

5. Metaphorical Inference and The Mapping Problem

Metaphors, similes and analogies are more than clever ways of re-stating the obvious. They are extraordinarily concise devices by which a writer can convey new information, simply by signaling his audience that information in the source domain is applicable to the target domain. Presumably the reader has a coherent body of knowledge about the source and can transfer and adapt this information to the target. This saves both the reader and the writer a good deal of time and cognitive effort. However, before the reader can initiate the transfer he must identify the correspondences between the two domains and establish exactly what information in the source is applicable to the target, since much of it is clearly inappropriate. This is the central issue in metaphor comprehension;

we refer to it as the *analogical mapping problem*.

One might suppose that the reader finds the analogical mapping simply by comparing the source and target domains for similarities. However, several researchers [8, 30] have demonstrated that this *domain comparison* model is in fact too simplistic. Although discovering similarities is an important component of the comprehension process, other strategies must be used as well. Indeed, when learning new material, the reader knows little about the target, hence domain comparison is, by itself, inadequate. Clearly, if the purpose of the metaphor is to transfer information to the target domain from the source domain, then this information does not already exist in the target domain, and hence it can play no part in the comparison process. Furthermore, when one considers requirements for computational tractability, it becomes evident that there must be strategies to help focus the comparison and constrain the matching process. Focusing limits the complexity of the comparison process and reduces spurious matches between the two domains.[3] Our purpose in this section is to explore some cognitively plausible focusing strategies which might be used to facilitate the construction of analogical mappings.

5.1. Knowledge Acquisition via Analogical Mappings

The following example, found in a children's book, illustrates an explanation in which the reader (presumably a child) is expected to create an analogical mapping and transfer information across domains.

> A motorcycle is a vehicle. Like a car it has a motor. But it looks
> more like a bicycle.

The author attempts to explain the concept of a motorcycle by referring to other, presumably more familiar, objects. But his statement implies much more than is explicitly stated. For instance, it suggests not only that a motorcycle has a motor, but that it has a motor in the same way that a *car* has a motor: that the motor is an internal combustion engine, it uses gasoline, it causes the machine to move, etc. The reference to a *car* is essential; consider the effect of substituting "electric shaver" for "car" in the example. (After all, electric shavers have motors too, but their motors are not a means of propulsion). Certainly, drawing an analogy to electric shavers would not be nearly as helpful in communicating what a motorcycle is.

---

[3] Implementing domain comparisons in a computer is typically accomplished by attempting to find matches between the representations of the target and source domains. As we shall see in the next section, these representations are typically graphs or equivalent structures. Although the details of the matching process vary considerably depending on the representation system used, the computation can be quite expensive if performed upon arbitrary domains. Indeed, the related "Subgraph Isomorphism" problem is NP-complete [19]. Given a precise formulation of the matching problem, it is easy to demonstrate that it too is intractable unless it is bounded in some principled way.

˙Although analogies, such as the one above, can obviously be used to great advantage in transmitting new information, the reader is often left in the position of not knowing how far carry the analogy. To a child who has never seen a motorcycle, the previous description of a motorcycle, though informative, is still quite ambiguous. "Does a motorcycle have pedals?" he may ask. In order to gauge the extent of the analogy, to verify which of his inference patterns relevant to cars and bicycles are valid for motorcycles, the child must either read further or find a picture of a motorcycle. A priori, there is no way for him to be sure which inferences to make. But at least the set of sensible questions he may ask will be focused by the analogy. Thus, it is reasonable to ask about handlebars or pedals, but not about whiskers or wings.

In most mundane situations, knowing which inferences are correct seldom poses a significant problem for people, largely due to the fact that there are characteristic ways of expressing metaphors so that the mapping problem is easier to solve. For instance, consider that the truly novel metaphor is rarely encountered. Through frequent use, many metaphors acquire idiomatic meanings, to a greater or lesser degree. We refer to these metaphors as *frozen*. "John is a hog" and "Sheila is a dog" both exemplify frozen metaphors. The latter would probably be interpreted as a rude comment concerning Sheila's looks, rather than a compliment on her loyalty, which seems to be an equally reasonable interpretation given only one's knowledge about dogs. Frozen metaphors are easy to understand because the analogical mapping has been (to some degree) precomputed, and so does not have to be reconstructed, only remembered and reapplied. Hence, neither a complex matching process nor prior knowledge about the target are necessary in order to find the mapping. There is little question of which are the right inferences and which are the wrong ones.

5.2. Salience and Novel Metaphors

If a metaphor is novel, other strategies are available for coping with the complexity of the mapping problem. One way is to focus on *salient* features of the source [30, 35]. Consider the example "Billboards are like warts" in which both the target and source are familiar objects. Most people interpret this as meaning that billboards stick out, and are ugly. Their mapping relates attributes that are common to both source and target, but particularly emphasizes those such as "ugliness" that are "high-salient" attributes of warts, the source. It is our contention that by focusing on prominent features and ignoring unimportant ones, the computational complexity of the mapping problem is reduced.

˙Concentrating the initial mapping to salient features of the source is an effective strategy even when one's knowledge of the target domain is limited. In fact, it is likely that the salient features are the very ones that should be mapped into the target domain, as is the case the following metaphor:

`RTO Inc. is the Freddie Laker of consumer electronics.`
Although the target is an unknown company, and the metaphor is novel, it is understandable simply because the source, Freddie Laker of Laker Airlines, has certain outstandingly salient features. Of course, the creator of the metaphor expects that his audience will all have the same opinion as to which of Laker's features are salient. Why certain features are considered universally salient whereas others are not is a difficult problem in its own right, one which we will not pause to consider here.

We have examined two types of metaphors which can be understood in spite of incomplete knowledge of the target: frozen metaphors and metaphors based on the source's salient features. These illustrate just two of the many ways *pragmatic* considerations enable one to bypass much of the complexity of the mapping problem. Occasionally, however, one cannot avoid more complex versions of the mapping problem. For us, this is the most interesting case. It occurs frequently during explanations involving extended analogies, such as when a grade-school mathematics teacher begins his algebra class by proclaiming:

> `An equation is like a balance. You must keep the same amount of`
> `weight on each side of the equals sign....`

Certainly there will be students in the class for whom this is a novel idea, and who spend the next 10 minutes desperately trying to find the intended analogical mapping. Or consider a secondary school biology text which begins a chapter on the human nervous system by comparing it to a telephone network. Or a treatise on "Hamlet" whose thesis is that the protagonist's life is a metaphor for adolescence. When confronted with one of these analogies in context, one may need to search for appropriate hypotheses; one's analogical mapping will be elaborated and changed as one's understanding of the target domain grows.

A good example of an extended analogy is provided by Milton Friedman in his Newsweek Column (of 12/27/82), as he attempts to explain recent fluctuations in the nation's money supply.

> ...Consider the record: M_1 grew at annual reates of 15.3 percent from October 1981 to January 1982; 1.2 percent from January 1982 to July 1982; 16.3 percent from July 1982 to November 1982. Is it really conceivable that the Fed produced these gyrations on purpose, given its repeated protestations that it was committed to a steady and moderate rate of monetary growth?

> Why the gyrations? A better explanation is that the Fed is, as it were, driving a car with a highly defective steering gear. It is driving down a road with walls on both sides. It can go down the middle of the road on the average only by first bouncing off one wall and then off the opposite wall. Not very good for the car or its passengers or bystanders, but one way to get down the road.

> This interpretation raises two key questions: first, why doesn't the Fed replace the defective steering gear? Second, what course will this defective steering gear lead to over

coming months...

This metaphor provides the reader with a clear, albeit simplistic, understanding of the situation *without requiring much prerequisite knowledge about how the Fed works.* A possible remedy (replacing the steering wheel) is suggested based on inferences valid in the source domain. Most importantly, the passage implies that the Fed's control over the economy and where it is headed (one can hardly help but use metaphors here!) is not very accurate. Simply by invoking this metaphor Friedman communicates his belief that this situation is bad -- but not totally disastrous -- without ever having to explain the underlying monetary and fiscal reasons. In fact, when we informally questioned people about what exactly Friedman is referring to when he speaks of "walls", most admitted that they weren't really sure. A typical response was that the "walls" represented some sort of "limits". And yet, these people felt that they had understood, or gotten the gist of, the metaphor. Apparently one's analogical mapping does not have to be particularly detailed, as long as *key* inferences can be made. It seems that once certain connections or *beachheads* have been established between the target and source domains, people are content to incrementally elaborate the mapping as they find it necessary during further reading or problem solving.

5.3. A Classification Based on Processing Requirements

In our discussion thus far, we have identified various analogical mapping strategies whose applicability depends upon the properties of the metaphor under consideration. We therefore offer the following pragmatic classification, based on what we believe are meaningful distinctions in the type of processing employed during comprehension. We caution that these categories should not be viewed as distinct; it seems more reasonable to view metaphors as occurring along a continuum with respect the criteria presented below.

- **Frozen Metaphors** -- .Example: "John is a hog." These have idiosyncratic, well-established meanings and therefore require little, if any, analysis or domain comparison during comprehension. Spurious inferences, such as John having a curly tail, snoutish nose or a tendency to crawl on all fours, do not typically enter the readers's mind.

- **Partially Frozen Metaphors** -- Example: "The new evidence weighed heavily in the defendant's favor". The balance metaphor is a partially frozen metaphor; the details of the metaphor may vary from instance to instance, but the mappings remain fairly standard. Previous experience guides the mapping process, and thereby reduces the amount of domain comparison necessary to establish a new mapping.

- **Novel One-shot Metaphors** -- Example: "RTO electronics is the Freddie Laker of consumer electronics". These metaphors may require a considerable amount of computation in order to construct a satisfactory mapping. Various strategies, such as focusing on salient attributes, are employed both by the writer in creating the metaphor, and the reader in comprehending the metaphor. Typically these metaphors are used

once, and then forgotten.

- **Extended Metaphors** -- The quote from Friedman (above) is an example of an extended metaphor. These are characterized by a relatively extensive mapping that is incrementally elaborated over time. The metaphor provides a model for reasoning about the target domain. *Scientific* analogies [10, 20] such as the nervous system/telephone network analogy fall into this category.

6. Representing Metaphors: The LIKE Relation

In the previous paragraphs we have discussed the problems involved in finding an analogical mapping and making metaphorical inferences. We now turn our attention from the process of comprehension to issues of representation. How do we represent an analogical mapping in a computational model? We know that our representation must satisfy two requirements:

1. The representation must facilitate the transfer of information from the source domain to the target domain, and

2. it must be dynamic, enabling the analogy to be elaborated over time.

In this section we discuss how analogies (and metaphors) can be represented in semantic networks so as to satisfy these requirements. Although our work was motivated by representation languages such as KL-ONE [4], SRL [39], NETL [17] and KRL [2], we intend the ideas presented below to be applicable on a broad basis, and therefore make no commitment to any particular representational scheme. The notation presented in our diagrams is meant to be purely illustrative.
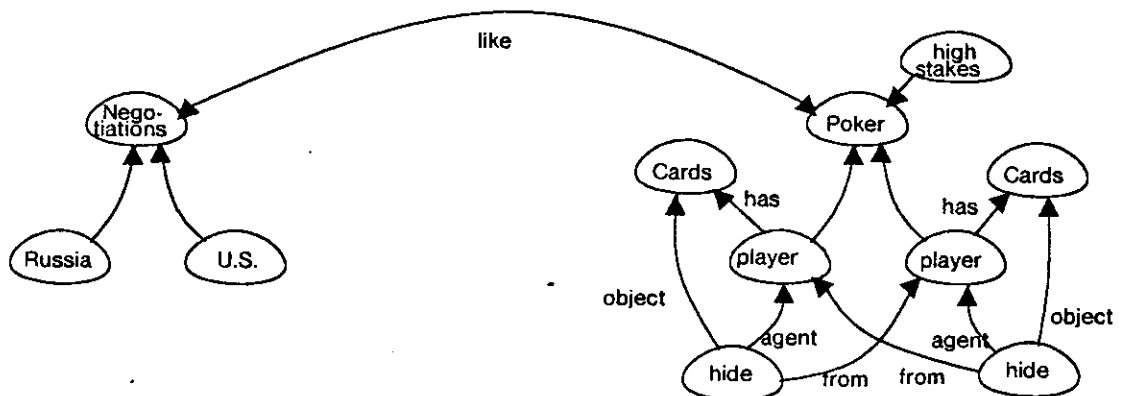

Figure 6-1: Using a single LIKE link to relate two domains

In a semantic network, knowledge about a domain is encoded as a graph. Typically the nodes

represent *concepts* and the links *relations*.[4]   In order to represent an analogy we need a way of indicating that one domain is "like" another domain. For instance, we might want to represent the metaphor "The U.S./Russian arms negotiations is a high stakes poker game". Unfortunately, simply "linking" the target concept to the source concept with a "LIKE" relation 6-1 does not provide us with the necessary functionality.   We require a means of representing the various correspondences between the two graphs so that it is apparent *how* arms negotiations are like a game of poker. Russia and the U.S. are both being compared to poker players, each of which is withholding information from the other, etc. The representation must identify precisely which subconcepts and corresponding relations map into each other. A single link between two domains is just not expressive enough to represent mappings conveying our understanding of the metaphor.

We can attempt to remedy the situation by connecting all the corresponding subconcepts in the two domains with LIKE links. However, this multiple-link solution also fails since the lower-level links have meaning only in the context of the entire analogy. To see why, consider what might occur after representing and storing "The Democrats and the Republicans are like Coke and Pepsi", meaning that they are virtually indistinguishable except by self-chosen labels. One of the assertions contained in the semantic network would be "The Republicans are like Pepsi". In isolation, this fact is meaningless and potentially misleading. We must associate it with its analogical context. But, if only subconcept correspondences are stored, it is impossible to reconstruct the context because there is nothing in the knowledge base that represents the analogy per se. The problem grows worse for more complex analogies because it is increasingly difficult to keep track of the various inter-relationships between concepts given the growing forest of LIKE links. Essentially, the multiple-link solution is inadequate because it relies completely on a reductionistic representation for the analogy. (See Bobrow and Winograd [2] and Schank [34] for discussions of more wholistic, reconstructive representations.)

In order to represent an arbitrary mapping between two domains, we propose to use a distinct entity, which we term a **mapping structure**. A mapping structure functions as a filter, allowing explicitly specified types of information to be transferred from one domain to another. (In this respect it plays a role similar to that of Winston's transfer frames [38, 37].) A mapping structure identifies the various correspondences in the source and target by providing a skeleton that specifies the analogical mapping in terms of common elements found in both domains. Figure 6-2 is a schematic

---

[4]Concepts may be decomposable, in which case a single node can be replaced by a network of lower-level nodes if desired. Exactly how this is managed is of importance to the domain matching process, but it need not concern us greatly for the purposes of this discussion.
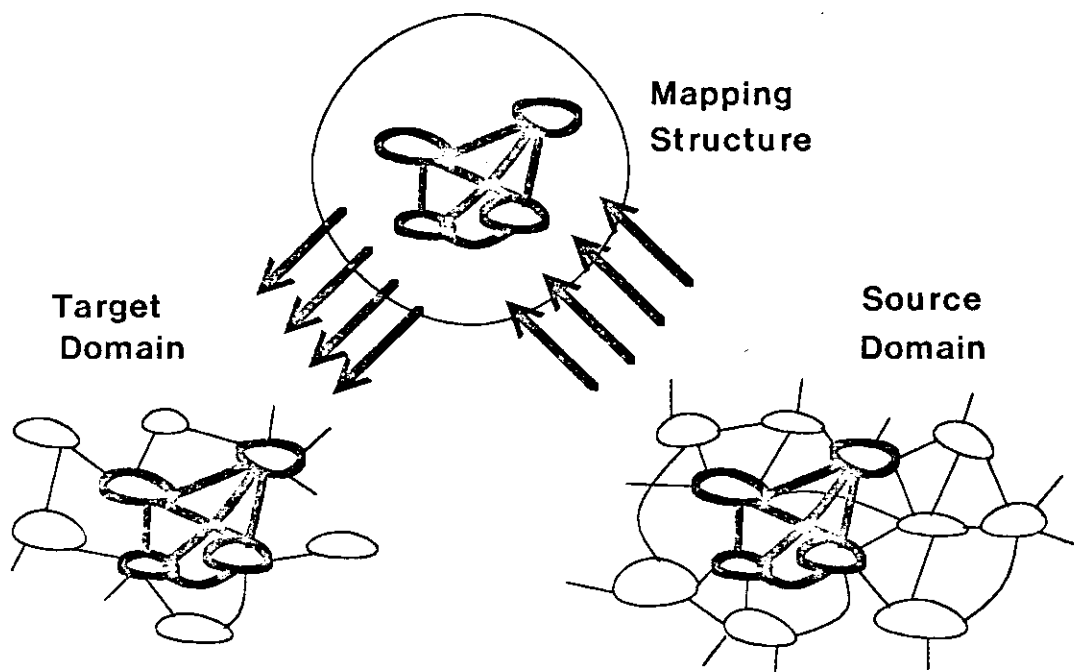
Figure 6-2: Schematic illustrating the role of mapping structures

which illustrates how the structure serves to coordinate the mapping by associating the two domains. The small graphs representing the target and source in the figure are meant to reside within much larger semantic networks.

In addition to the role they play in associating the source and target domains, mapping structures serve to organize meta-information about the mapping. Inferences made as a result of the analogy are represented by collections of new nodes in the target domain. These new nodes are associated with *data-dependencies* [13, 16] referring back to the mapping structure. (A data-dependency is essentially an indication of how the information was derived.) This enables the mapping structure to be extended incrementally while permitting subsequent verification or retraction of inferences. Because the mapping structure can be modified dynamically, at any particular time it represents the current conception of what the metaphor means. Of course, this implies that the mapping structure must be retained for some unspecified duration. We assume that the mapping structure will be "forgotten" (i.e. discarded by some autonomous process supporting the representation system) if it dos not continue to be accessed as a source of inferences when target domain information is retrieved from or added to memory.

Figure 6-3 shows a simplified representation for "Bulgaria is a Russian Puppet". The dotted lines around the CONTROL node in the target domain are meant to suggest that this is an inference made
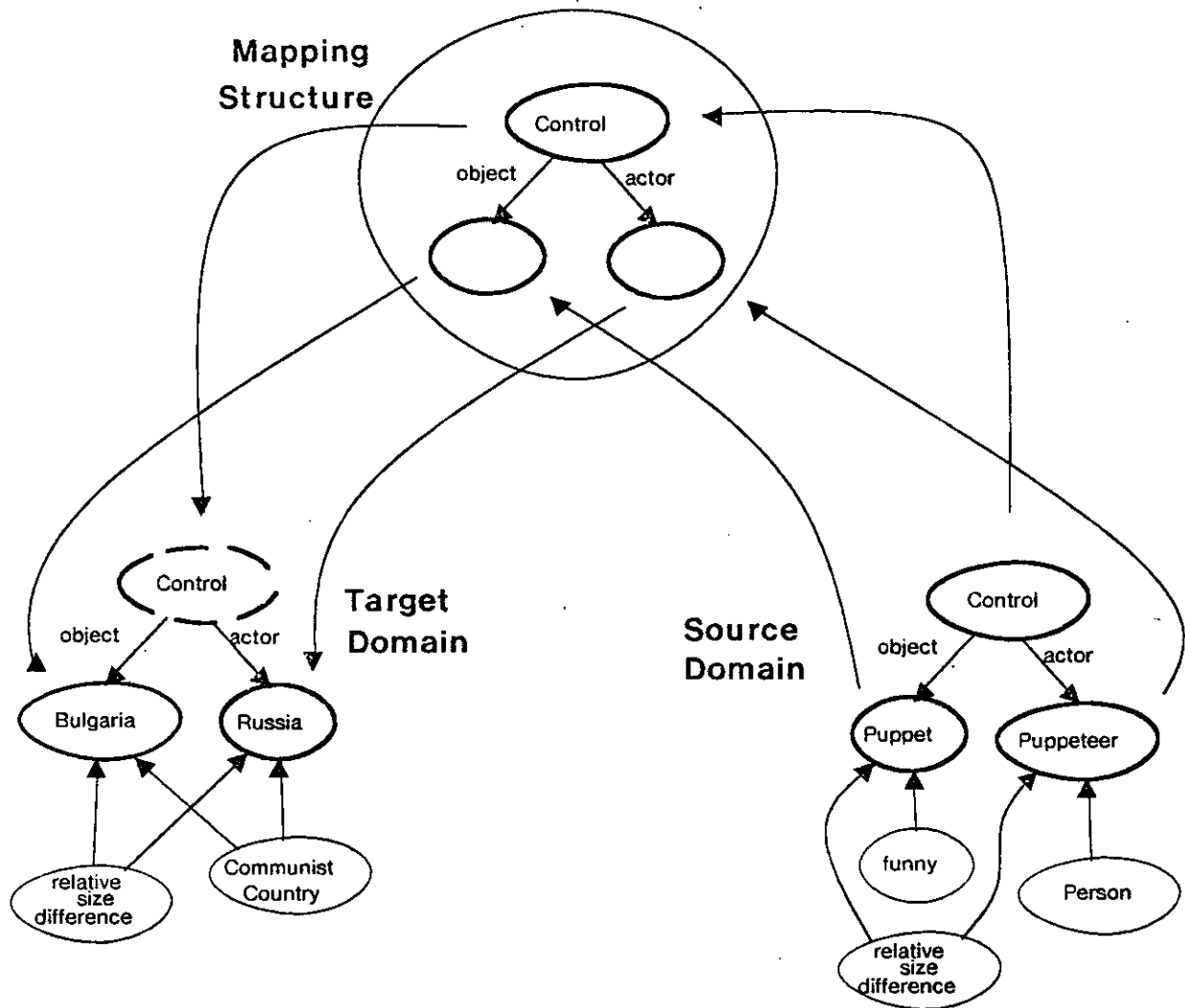
Figure 6-3: Simplified representation of "Bulgaria is a Russian puppet".

as a result of the metaphor. For simplicity, only the directly relevant sections of the target and source domains are actually shown. In addition, several nodes which do not take part in the mapping, such as the FUNNY attribute describing puppets, are included for illustrative purposes. Note that in both the target and the source, there is a node signifying the "relative size differences" of the objects (admittedly a gross representational simplification). Although this node is not part of the mapping, it might very well be included later if the mapping is extended.

As we pointed out in the previous section, a metaphor may become frozen through frequent use. An advantage gained by the use of mapping structures is that we can model computationally this "freezing" process quite naturally. When presented with a novel metaphor, there is no recourse but

to construct a brand new mapping structure. However, suppose a metaphor is encountered whose source is similar to the source of a previously understood metaphor. In this case we can use the mapping structure built earlier as a template to help build the new structure. This constrains the domain comparison process because the relevant features of the source concept are identified by the pre-existing mapping structure. While one must still locate corresponding features in the new target domain, this too may be done efficiently if the two targets are similar.

The statement "John eats like a pig" is a typical example of a frozen metaphor. Notice that it is understandable even though we are using "John" as a generic person. In our model, the mapping structure corresponding to "...eats like a pig" is associated with the section of the knowledge network where information about pigs' eating habits is stored. Parsing "John eats like a pig" requires retrieving this mapping structure, noticing the exact correspondence between the source in the structure and the source in the new metaphor, and then instantiating the structure with "John" as the target domain. Instantiation is relatively easy to do, because the mapping structure specifies which nodes map from the source [8]. Obviously we have glossed over many important problems in this description, such as how mapping structures can be retrieved given a source description, and whether a new physical copy of the mapping structure must be generated for each instantiation of a frozen metaphor. These questions are being studied at the present time.

7. Generalizing Mapping Structures

In the previous section, mapping structures were proposed as a means for representing arbitrary inter-domain correspondences. It is our intention that mapping structures be viewed as data-structures which *implement* LIKE relations. That is, a LIKE relation still exists between the source and target domain of an analogy, but it is too complex to be implemented with a simple link. Instead, a more elaborate mechanism is required to represent the internal structure of the analogical relationship. The indirect implementation of an analogical relationship as a data structure declaratively specifying the mapping process provides a necessary extra level of abstraction along its functional dimension. Thus, one can refer to the entire analogy as a unit, or one can access and elaborate the constituent parts of the mapping structure.

At the present time, we are considering other relations that may be better implemented by mapping structures rather than by simple links. Perhaps the most obvious candidate the IS-A relation, which provides a way to structure a knowledge network into a **type hierarchy** so that properties of a class representative can be mapped automatically to members of that class. We refer to this as *vertical*

inheritance, because each concept inherits from those above it in the type hierarchy.[5]   Historically, vertical inheritance has been used in knowledge representation systems to implement certain types of default reasoning. For example, knowing that Clyde is an elephant, and elephants have trunks, a system might use inheritance to infer that Clyde has a trunk.[6]
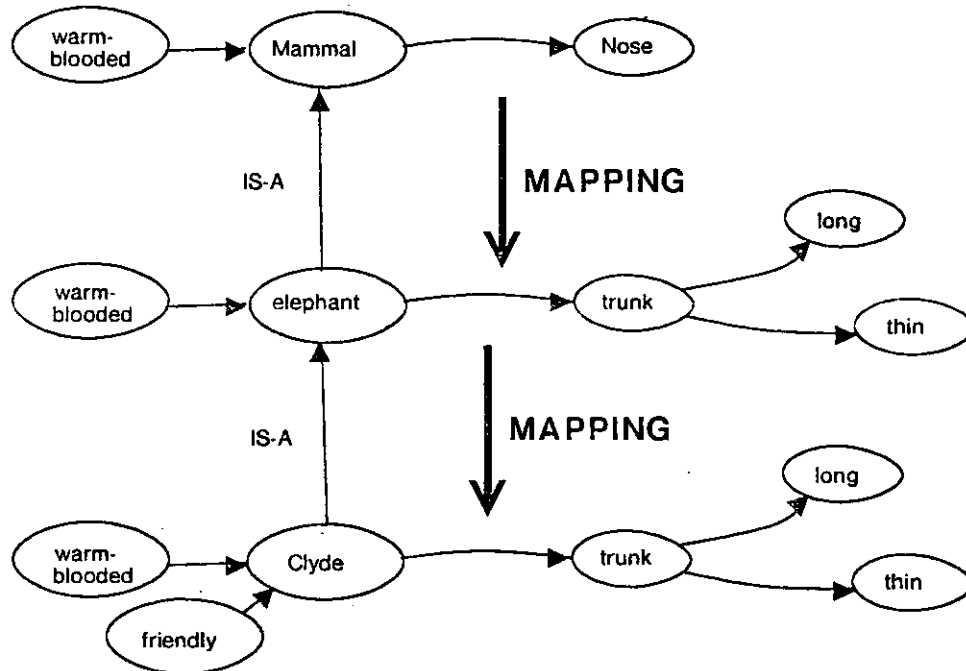

Figure 7-1:   IS-A links and other inheritance relations entail mappings between concepts.

The IS-A relation requires that a mapping be established between the superordinate and subordinate concepts, so that properties of the superordinate can be inherited by the subordinate (Figure 7-1). Accordingly, it is natural to use a mapping structure to accomplish this as was done for the LIKE relation. The mapping structure allows us to record what information may be transmitted from a particular superordinate, as well as to indicate explicitly what information *should not* be inherited. Although one assumes by default that all the information in the superordinate is shared by the subordinate, in many cases this can be contradicted. For example, if it is known that "Penguins do not fly", then the PENGUIN node should not inherit the FLIES attribute from BIRD. The mapping

---

[5]In AI applications, the type hierarchy is almost never a strict hierarchy, but rather is a *directed acyclic graph* or a *tangled hierarchy*, in which a concept can have more than one superordinate. See Brachman [3] for an overview of network-based knowledge representation methods.

[6]Other reasons for employing inheritance include space savings, although this is countered by a corresponding time cost during retrieval operations. In addition, updating the knowledge base can be simpler; to add the fact that mammals are warm blooded, only single node must be changed, that of "TYPICAL-MAMMAL", rather than having to find and change each and every node in the network denoting John, Mary, Felix the Cat, saber-tooth tigers, and all other mammals.

structure provides us with a *view* of a penguin as a bird[7] The key point we mean to emphasize is that both the LIKE relation and the IS-A relation require mappings between concepts. This similarity has prompted some to describe the analogical mapping process as *lateral inheritance*, since information is mapped laterally between concepts in the type hierarchy. If one accepts this viewpoint, then mapping structures begin to play a broad role in representing complex situations.

8. Towards a Computational Model of Metaphorical Inference

The pervasiveness of metaphor in every aspect of human communication has been convincingly demonstrated by Lakoff and Johnson [26], Ortony [31], Hobbs [23], and many others. However, with a few exceptions [23, 8] the creation of a process model of metaphor comprehension and inference has not been of central concern. From a computational standpoint, metaphor has been viewed as an obstacle, to be tolerated at best and ignored at worst. For instance, Wilks [36] presents a few rules on how to relax semantic constraints in order for a parser to process a sentence in spite of the metaphorical usage of a particular word. From our point of view, this attitude is not surprising, since few AI systems to date have used analogical reasoning as a primary inference method. Analogical reasoning has been viewed as a difficult problem in its own right, which must be solved before it can be incorporated in applications systems (such as parsers and medical diagnosis systems). However, a robust system must be able to operate analogically, especially if intended for naive users, otherwise they would find its lack of "common sense" intolerable. For example, a parser which could not understand metaphors, analogies, or similes would be useful only in the most limited of situations.[8] With these thoughts in mind we have begun initial work towards a parser which can reason metaphorically, and below present the following conceptual steps in the metaphor-recognition parsing process:

1. Identification of the source and target concepts. This is done during the parser's normal, non-metaphorical operation.

2. Recognition that the input currently being parsed cannot be handled literally, and is in fact an instance of a metaphor. This is actually a non-trivial task requiring considerable

---

[7]Views have played a central role in many knowledge representation schemes, including those of and Moore and Newell [29], Bobrow and Winograd [2], and Wright and Fox [39]. To indicate the need for more flexible mappings that simple all-or-nothing inheritance, consider the fact that the average mammal may be 3 feet tall, or may range from a 1/2 inch to 21 feet tall. Whereas we want our concept of "Giraffe" to inherit most of our knowledge of mammals, we clearly do not want to say that the average giraffe is 3 feet tall, nor that giraffes range in height from 1/2 inch to 21 feet tall. Hence, the IS-A relation inherits only certain classes of attributes and excludes others; typically intrinsic properties of individual members are inherited whereas aggregate set properties are not. A mapping structure can be used to make explicit statements, such as the one above, regarding the information that may be transmitted from one concept to another via any particular inheritance link [18, 6].

[8]Skeptics who dispute this claim are invited to examine any source of common everyday text, such as a copy of Time magazine or even the New York Times Financial section, and count the number of metaphors occurring on a single page.

sophistication.   For example, the parser must realize that the input is not simply erroneous. This judgement depends to a large degree on pragmatic considerations.

3. Creation of an analogical mapping from the source domain unto the target domain so that corresponding subconcepts in the two domains map to each other. This phase may be broken down further as follows:

   a. Search for a pre-existing mapping structure associated with the source domain.

   b. If any such structure is found, check whether it is appropriate with respect to the new target domain. This is done by building incrementally a new mapping structure containing the same nodes as the old structure. As each new node is created a corresponding node in the target domain must be identified.

   c. If no pre-existing mapping structure is found for the source, or those that are found prove to be inappropriate for the new target, then a new mapping must be constructed from scratch. A matching algorithm must search the two domains in order to find similarities.   In this case, one should use as many heuristics as possible for reducing the amount of domain comparison that must be done. Possible heuristics include focusing on salient concepts in the source, and focusing on certain categories of knowledge which tend to be mapped invariant in meaningful metaphors [8].

4. Once corresponding nodes in the two domains have been identified (by constructing a mapping structure), knowledge from the source can be added to the mapping, thereby generating corresponding inferences within the target domain.[9] In an abstract sense, this mechanism accomplishes an implicit transfer of information from the source to the target. Verification that the metaphorical inferences are compatible with the target domain is an integral part of this process.

Whether or not it is possible to develop a *robust* metaphor comprehension system with today's technology is a matter of debate. Metaphorical understanding requires a potentially vast amount of world knowledge, as well as an efficient way of comparing large domains for similarities. However, we feel that even a fragile, partial model built along these lines is a worthwhile endeavor, since eventually these problems must be solved in order to create a truly intelligent parser and inference system.

In cooperation with work towards a model of metaphor understanding, we are also studying the role that metaphorical inference plays in scientific reasoning. As discussed earlier, metaphorically-based

---

[9]We acknowledge that the model as specified does not account for the way people understand metaphors such as "Sally is a block of ice", in which the properties transferred from the source are themselves metaphorical. The metaphor transfers the property "cold" from ice to Mary, but this is a metaphor within a metaphor because we are refering to Mary's personality rather than her temperature. Metaphors occur in all shapes and sizes, and we have not addressed many of the subtler nuances of the phenomenon in this paper. We do believe, however, that the model can be elaborated to handle more sophisticated metaphors without revising the general framework we have presented.

general patterns of inference do not appear confined to naive reasoning in mundane situations. Gentner [20] and Johnson [24] have argued the significant role that metaphor plays in formulating scientific theories. In preliminary investigations, Larkin and Carbonell [27, 10] have isolated general inference patterns in scientific reasoning that transcend the traditional boundaries of a science. For instance, the notion of equilibrium (of forces on a rigid object, or of ion transfer in aqueous solutions, etc.) is, in essence, a more precise and general formulation of the balance metaphor. Reasoning based on recurring general inference patterns seems common to all aspects of human cognition. These patterns encapsulate sets of rules to be used in unison, and thereby bypass some of the combinatorial search problems that plague more traditional rule-based deductive inference systems. The inference patterns are frozen from experience and generalized to apply in many relevant domains.

At the present stage in the investigation, we are searching for general inference patterns and the metaphors that give rise to them, both in mundane and in scientific scenarios. As these patterns are discovered, they are cataloged according to the situational features that indicate their presence. The basic metaphor underlying each inference pattern is recorded along with exemplary linguistic manifestations. The internal structure of the inference patterns themselves are relatively simple to encode in an AI system. The difficulty arises in connecting them to the external world (i.e., establishing appropriate mappings) and in determining their conditions of applicability (which are more accurately represented as partial matches of the situations where apply, rather than as simple binary tests). For instance, it is difficult to formulate a general process capable of drawing the mapping between the "weight" of a hypothetical object and the corresponding aspect of the non-physical entity under consideration, so that the balance inference pattern my apply. It is equally difficult to determine the degree to which this or any other inference pattern can make a useful contribution to novel situations that bear sufficient similarity to past experience [12].

9. Conclusions

In this paper we have analyzed the role of metaphors in common sense reasoning. In particular, we showed how the balance metaphor exemplifies metaphorical inference, suggested that inference patterns valid for physical domains might provide the foundation upon which much of human common-sense reasoning rests, and provided the first steps toward a computationally-effective method for representing analogical mappings. However, since the current study is only in its initial stages, the hypothesis that metaphorical inference dominates human cognition retains the status of a conjecture, pending additional investigation. We would say that the weight of the evidence is as yet insufficient to tip the academic scales.

Our investigations to date suggest that intensified efforts to resolve the questions raised in this paper may prove fruitful, in addition to pursuing the following related research objectives:

- Develop an augmented representation language that handles analogical mappings as a natural operation. We intend to start from a fairly flexible, operational language such as SRL [39]. Using this language we intend to build and test a system that acquires new information from external metaphorical explanations.

- Continue to develop the MULTIPAR multi-strategy parsing system [21, 22] and incorporate within its evolving flexible parsing strategies a means of recognizing and processing metaphors along the lines mentioned in this paper.

- Examine the *extent* to which linguistic metaphors reflect underlying inference patterns. The existence of a number generally useful inference patterns based on underlying metaphors provides evidence against, but does not refute, the possibility that the vast majority of metaphors remain mere linguistic devices, as previously thought. In essence, the existence of a phenomenon does not necessarily imply its universal presence. This is a matter to be resolved by more comprehensive future investigation.

- Investigate the close connection between models of experiential learning and metaphorical inference. In fact, our earlier investigation of analogical reasoning patterns in learning problem solving strategies first suggested that the inference patterns that could be acquired from experience coincide with those underlying many common metaphors [12, 8].

- Exploit the human ability for experientially-based metaphorical reasoning in order to enhance the educational process. In fact, Sleeman and others have independently used the *balance metaphor* to help teach algebra to young or learning disabled children. Briefly, a scale is viewed as an equation, where the quantities on the right and left hand sides must balance. Algebraic manipulations correspond to adding or deleting equal amounts of weight from both sides of the scale, hence preserving balance. First, the child is taught to use the scale with color-coded boxes or different (integral) weights. Then, the transfer to numbers in simple algebraic equations is performed. Preliminary results indicate that children learn faster and better when they are able to use explicitly this general inference pattern. We foresee other applications of this and other metaphorical inference patterns in facilitating instruction of more abstract concepts. The teacher must make the mapping explicit to the student in domains alien to his or her past experience. As discussed earlier, establishing and instantiating the appropriate mapping is also the most problematical phase from a computational standpoint, and therefore should correspond to the most difficult step in the learning process.

Clearly, the possible research directions suggested by our initial investigations far outstrip our resources to pursue them in parallel. Hence, we will focus first on the basic representation and parsing issues central to a computational model of metaphorical reasoning.

10. Bibliography

1.   Black, M., *Models and Metaphors,* Cornell University Press, 1962.

2.   Bobrow, D. G. and Winograd, T., "An Overview of KRL, a Knowledge Representation Language," *Cognitive Science,* Vol. 1, No. 1, 1977, pp. 3-46.

3.   Brachman, R. J., *A Structural Paradigm for Representing Knowledge,* PhD dissertation, Harvard University, 1977, Also BBN report #3888

4.   Brachman, R. J., "On the Epistemological Status of Semantic Networks," in *Associative Networks,* N. V. Findler, ed., New York: Academic Press, 1979.

5.   Burstein, M. H., "Concept Formation Through the Interaction of Multiple Models," *Proceedings of the Third Annual Conference of the Cognitive Science Society,* 1981.

6.   Carbonell, J. G., "Default Reasoning and Inheritance Mechanisms on Type Hierarchies," *SIGART, SIGPLAN, SIGMOD Joint volume on Data Abstraction,* 1980.

7.   Carbonell, J. G., "A Computational Model of Problem Solving by Analogy," *Proceedings of the Seventh International Joint Conference on Artificial Intelligence,* August 1981, pp. 147-152.

8.   Carbonell, J. G., "Metaphor: An Inescapable Phenomenon in Natural Language Comprehension," in *Strategies for Natural Language Processing,* W. Lehnert and M. Ringle, eds., New Jersey: Erlbaum, 1982, pp. 415-434.

9.   Carbonell, J. G., "Invariance Hierarchies in Metaphor Interpretation," *Proceedings of the Third Meeting of the Cognitive Science Society,* August 1981, pp. 292-295.

10.   Carbonell, J. G., Larkin, J. H. and Reif, F., "Towards a General Scientific Reasoning Engine," *Proceedings of the Eighth International Joint Conference on Artificial Intelligence,* 1983, (Submitted)

11.   Carbonell, J. G., "Experiential Learning in Analogical Problem Solving," *Proceedings of the Second Meeting of the American Association for Artificial Intelligence,* Pittsburgh, PA, 1982.

12.   Carbonell, J. G., "Learning by Analogy: Formulating and Generalizing Plans from Past Experience," in *Machine Learning, An Artificial Intelligence Approach,* R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., Tioga Press, Palo Alto, CA, 1983.

13.   Charniak, E., Riesbeck, C. and McDermott, D., *Artificial Intelligence Programming,* Lawrence Erlbaum Associates, 1980.

14.   Chase, W. G. and Simon, H. A., "Perception in Chess," *Cognitive Psychology,* Vol. 4, 1974, pp. 55-81.

15.   Clements, J., "Analogical Reasoning Patterns in Expert Problem Solving," *Proceedings of the Fourth Annual Conference of the Cognitive Science Society,* 1982.

16.   Doyle, J., "A Truth Maintenance System," *Artificial Intelligence,* Vol. 12, No. 3, 1979, pp. 231-272.

17.   Fahlman, S. E., *NETL: A System for Representing and Using Real World Knowledge,* MIT Press, 1979.

18.   Fox, M. S., "On Inheritance in Knowledge Representation," *Proceedings of the Sixth International Joint Conference on Artificial Intelligence,* 1979, pp. 282-284.

19.   Garey, M. and Johnson, D., *Computers and Intractability,* W. H. Freeman, San Francisco, 1979.

20.   Gentner, D., "The Structure of Analogical Models in Science," Tech. report 4451, Bolt Beranek and Newman, 1980.

21.   Hayes, P. J., and Carbonell, J. G., "Multi-Strategy Construction-Specific Parsing for Flexible Data Base Query and Update," *Proceedings of the Seventh International Joint Conference on Artificial Intelligence,* August 1981, pp. 432-439.

22.   Hayes, P. J. and Carbonell, J. G., "Multi-Strategy Parsing and it Role in Robust Man-Machine Communication," Tech. report CMU-CS-81-118, Carnegie-Mellon University, Computer Science Department, May 1981.

23.   Hobbs, J. R., "Metaphor, Metaphor Schemata, and Selective Inference," Tech. report 204, SRI International, 1979.

24.   Johnson, M., "Metaphorical Reasoning", Unpublished manuscript

25.   Kolodner, J. L., *Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model,* PhD dissertation, Yale University, Nov. 1980.

26.   Lakoff, G. and Johnson, M., *Metaphors We Live By,* Chicago University Press, 1980.

27.   Larkin, J. H. and Carbonell, J. G., "General Patterns of Scientific Inference: A Basis for Robust and Extensible Instructional Systems", Proposal to the Office of Naval Research

28.   Lebowitz, M., *Generalization and Memory in an Integrated Understanding System,* PhD dissertation, Yale University, Oct. 1980.

29.   Moore, J. and Newell, A., "How can MERLIN Understand?," in *Knowledge and Cognition,* L. Gregg, ed., Hillsdale, NJ: Erlbaum Assoc., 1974, pp. 253-285.

30.   Ortony, A., "The Role of Similarity in Similes and Metaphors," in *Metaphor and Thought,* A. Ortony, ed., Cambridge University Press, 1979.

31.   Ortony, A. (Ed.), *Metaphor and Thought,* Cambridge University Press, 1979.

32.   Schank, R. C., "Reminding and Memory Organization: An Introduction to MOPS," Tech. report 170, Yale University Comp. Sci. Dept., 1979.

33.   Schank, R. C., "Language and Memory," *Cognitive Science,* Vol. 4, No. 3, 1980, pp. 243-284.

34.   Schank, R. C., *Dynamic Memory,* Cambridge University Press, 1982.

35.   Searle, J.R., "Metaphor," in *Metaphor and Thought,* A. Ortony, ed., Cambridge University Press, 1979.

36.   Wilks, Y., "Knowledge Structures and Language Boundaries," *Proceedings of the Fifth International Joint Conference on Artificial Intelligence,* 1977, pp. 151-157.

37.   Winston, P. H., "Learning and Reasoning by Analogy," *Comm. ACM,* Vol. 23, No. 12, 1979, pp. 689-703.

38.   Winston, P.H., "Learning by Creating and Justifying Transfer Frames," *Artificial Intelligence,* Vol. 10, No. 2, 1978, pp. 147-172.

39.   Wright, K. and Fox, M., "The SRL Users Manual," Tech. report, Robotics Institute, Carnegie-

Metaphor and Common Sense Reasoning

Mellon University, 1983.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CMU-CS-83-110	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) METAPHOR AND COMMON-SENSE REASONING		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Jaime Carbonell Steven Minton		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0661 N00014-82-C-50767
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Department Pittsburgh, PA 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE March 5, 1983
		13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)