# Metaphor and Abduction

Jerry R. Hobbs
Artificial Intelligence Center
SRI International

*None but a geometrician or a blockhead can talk without a metaphor.*
Jean-Jacques Rousseau,
*Julie*, or *La Nouvelle Héloïse*

## Abstract

In this paper a recent approach to inference in text understanding based on abduction is applied to the problem of metaphor interpretation. The fundamental ideas in the "interpretation as abduction" approach are outlined. A succinct characterization of interpretation is given, along with a brief example and a description of the principal features of a weighted abduction scheme that is used. This approach is shown to lead to an elegant integration of syntax, semantics, and pragmatics. Three examples of metaphor interpretation are then analyzed within the abductive framework to determine what problems arise. The examples are a conventionalized metaphor schema, a standard category metaphor contextually interpreted, and a novel metaphor. The primary problem that arises for metaphor interpretation in the abductive framework is dealing with the fact that metaphors are not literally true. Two perspectives on this difficulty are offered.

## 1 Introduction

The use of inference in text understanding has been an important theme in computational linguistics for at least two decades, and there have been a number of inference-based approaches to metaphor interpretation (e.g., Carbonell, 1982; Hobbs, 1983; Indurkhya, 1987). In recent years, a particularly elegant formulation of the use of inference in text understanding has emerged, based on abduction, or inference to the best explanation (e.g., Charniak and Goldman, 1988; Hobbs et al., 1988, 1990; Norvig, 1987). A great many problems in interpretation, such as reference resolution, expansion of metonymies, and resolution of some lexical and syntactic ambiguities, have been shown to be subsumed under this new approach. This paper is an attempt to extend the abductive approach to cover metaphor interpretation as well, by recasting an inference-based approach to metaphor in light of one of the abductive frameworks. Such an effort ought to have the highest priority for any new approach to interpretation because, as Rousseau and many others since have pointed out, metaphor pervades discourse.

In Section 2 some basic assumptions about knowledge representation are laid out.

In Section 3 the fundamental ideas in the "interpretation as abduction" approach are outlined, though necessarily briefly. A succinct characterization of interpretation is given, along with a brief example. The principal features of a weighted abduction scheme are described and shown to lead to an elegant integration of syntax, semantics, and pragmatics.

In Section 4 three examples, analyzed in an older framework in Hobbs (1983), are reanalyzed within the abductive framework to determine what problems arise. The examples are a conventionalized metaphor schema, a standard category metaphor contextually interpreted, and a novel metaphor.

The primary problem that arises for metaphor interpretation in the abductive framework is dealing with the fact that metaphors are not literally true. In Section 5 two perspectives on this difficulty are offered.

## 2 Representation

In this paper formal logic is the representation language. This choice ought to be uncontroversial. There are six features any adequate representation language ought to have. Five of them are just the features that the syntax of formal logic provides. The sixth is a feature that many researchers working in a logical framework are attempting to provide. All but one of these is provided by neural nets as well, indicating that there is in fact agreement about requirements across a broad range of representational positions.[1]

[1]Neural nets are of course a computational device and not, strictly speaking, a representation language. However, the mode of computation they provide can be thought of as an inference mechanism. In many of their uses, the nodes are given propositional interpretations, and even in cases where their behavior is learned rather than programmed, intellectual curiosity often leads one to ask for a propositional characterization of what contribution to the total behavior of the system each node makes.

**1. Conjunction:** There needs to be a way for two propositions to have an additive effect. Conjunction provides this in formal logic. The additive effect of excitatory links provides this in neural nets.

**2. Implication and Modus Ponens:** One needs implicative rules, that allow one proposition to be concluded from another. Implication and modus ponens provide this in formal logic. In neural nets, the excitatory links between nodes provide it. Thus, when one node is activated, it will, under the right circumstances, cause other nodes to which it is connected to be activated as well.

**3. Negation and Recognizing "Obvious" Contraditions:** There needs to be a way of recognizing at least the most obvious contradictions, such as that a cat is not a dog. This is provided by negation in formal logic. It is provided by inhibitory links in neural nets.

**4. Predicate-Argument Relations:** There must be a way of expressing relations among entities, and of distinguishing which entity plays what role in the relation. This capability is provided in first-order predicate calculus. It is one of the principal features that frame-like representations seek to provide. Predicate-argument relations can be represented in neural nets without difficulty.

**5. Variables and Universal Instantiation:** Variables are required in some form or other, in order to express general knowledge, and there must be a way of binding variables to particular values. Formal logic provides a particularly clean way of doing this. Neural nets do not, and this is generally recognized as perhaps the greatest shortcoming of neural nets as a computational framework for simulating intelligent behavior.

There are some who would claim that, rather than using general knowledge, we understand new situations using previously acquired particular knowledge by a process of analogy or reminding. But in every analogy there is an implicit generalization, whether or not we acknowledge it explicitly. Thus, suppose we know about an old situation the following facts:

$$p_1(A),\ p_2(A),\ p_3(A),\ p_4(A),\ p_5(A),\ q(A)$$

Then we encounter a new situation in which the following are true:

$$p_1(B),\ p_2(B),\ p_4(B)$$

We hypothesize, by analogy, that $q(B)$ is also true. But this corresponds to using the implicit generalization

$$(\forall\, x)p_1(x) \,\wedge\, p_2(x) \,\wedge\, p_4(x) \,\supset\, q(x)$$

Such generalizations can be made defeasible in a number of ways, including using "et cetera" predications in an abductive framework, as described in Section 3.3.

**6. "Soft Corners":** There must be ways in which some inferences can be favored over others. There must be ways to accommodate inconsistent propositions. This is one of the primary strengths of neural net representations. Researchers working on model preference semantics for nonmonotonic logics and researchers working on probabilistic reasoning are searching for the appropriate means to give hard-edged logic softer corners. The scheme for weighted abduction presented here is another attempt to do just that.

Any system with these features can be used in three different ways—deductive- ly, inductively, and abductively. In deduction, from $(\forall\, x)p(x) \supset q(x)$ and $p(A)$, one concludes $q(A)$. In induction, from $p(A)$ and $q(A)$, or more likely, from a number of instances of $p(A)$ and $q(A)$, one concludes $(\forall\, x)p(x) \supset q(x)$. Abduction is the third possibility; from $(\forall\, x)p(x) \supset q(x)$ and $q(A)$, one concludes $p(A)$. One can think of $q(A)$ as the observable evidence, of $(\forall\, x)p(x) \supset q(x)$ as a general principle that could explain $q(A)$'s occurrence, and of $p(A)$ as the inferred, underlying cause or explanation of $q(A)$. Of course, this mode of inference is not valid; there may be many possible such $p(A)$'s. Other criteria of cost and utility are needed to choose among the possibilities.

In this paper we will use formal logic as our representation language, and we will use it abductively as our tool for interpreting natural language discourse.

# 3 Interpretation as Abduction

## 3.1 Characterizing Interpretation

Abductive inference is inference to the best explanation. The process of interpreting sentences in discourse can be viewed as the process of providing the best explanation of why the sentences would be true. More precisely, the process of interpreting a sentence can be described as follows:

To interpret a sentence:

(1) Prove the logical form of the sentence,
    together with the constraints that predicates impose on
        their arguments,
    allowing for coercions,
Merging redundancies where possible,
Making assumptions where necessary.

By the first line we mean "prove from the predicate calculus axioms in the knowledge base, the logical form that has been produced by syntactic analysis and semantic translation of the sentence." By "coercion" we mean the function or relation that maps from the explicit argument to the intended implicit argument in cases of metonymy, such as the mapping from "Shakespeare" to "the plays of Shakespeare" in the interpretation of the sentence "I like to read Shakespeare."

In a discourse situation, the speaker and hearer both have sets of private beliefs, and there is a large overlapping set of mutual beliefs. An utterance spans the boundary between mutual belief and the speaker's private beliefs. It is a bid to extend the area of mutual belief to include some private beliefs of the speaker's.[2] It is anchored referentially in mutual belief, and when we prove the logical form and the constraints, we are recognizing this referential anchor. This is the given information, the definite, the presupposed. Where it is necessary to make assumptions, the information comes from the speaker's private beliefs, and hence is the new information, the indefinite, the asserted. Merging redundancies is a way of getting a minimal, and thus a best, interpretation.

This approach to discourse interpretation has been implemented in the TACITUS system (Hobbs et al., 1990) at SRI International, using Stickel's Prolog Technology Theorem Prover (Stickel, 1989). The system has been employed in several moderate-scale applications.

[2]This is clearest in the case of assertions. But questions and commands can also be conceived of as primarily conveying information—about the speaker's wishes. In any case, most of what is required to interpret the three sentences,

John called the Boston office.
Did John call the Boston office?
John, call the Boston office.

is the same.

### 3.2 An Example

This characterization, elegant though it may be, would be of no interest if it did not lead to the solution of the discourse problems we need to have solved. A brief example will illustrate that it indeed does.

But first a notational convention that is used throughout this paper: we will take $p(x)$ to mean that $p$ is true of $x$, and $p'(e, x)$ to mean that $e$ is the eventuality or possible situation of $p$ being true of $x$. This eventuality may or may not exist in the real world. The unprimed and primed predicates are related by the axiom schema

$$(\forall x)p(x) \equiv (\exists e)p'(e, x) \wedge Rexists(e)$$

where $Rexists(e)$ says that the eventuality $e$ does in fact really exist. This notation, by reifying events and conditions, provides a way of specifying higher-order properties in first-order logic. This Davidsonian reification of eventualities (Davidson, 1967) is a common device in AI. See Hobbs (1985) for further explanation of the specific notation and ontological assumptions.

The example is

(2) The Boston office called.

This example illustrates three problems in "local pragmatics", the reference problem (What does "the Boston office" refer to?), the compound nominal interpretation problem (What is the implicit relation between Boston and the office?), and the metonymy problem (How can we coerce from the office to the person at the office who did the calling?).

Let us put these problems aside and interpret the sentence according to characterization (1). Ignoring tense, the logical form is something like

(3) $$(\exists e, x, o, b) call'(e, x) \wedge person(x) \wedge rel(x, o) \\ \wedge office(o) \wedge nn(b, o) \wedge Boston(b)$$

That is, there is a calling event $e$ by a person $x$ related somehow (possibly by identity) to the explicit subject of the sentence $o$, which is an office and bears some unspecified relation $nn$ to $b$ which is Boston. The requirement that $x$ be a person is the constraint that the predicate $call'$ imposes on its agent argument.

Suppose our knowledge base consists of the following facts: We know that there is a person John who works for $O$ which is an office in Boston $B$.

(4) $person(J)$, $work\text{-}for(J, O)$, $office(O)$, $in(O, B)$, $Boston(B)$

We also know that *work-for* is a possible coercion relation,

(5) $(\forall\, x, y) work\text{-}for(x, y) \supset rel(x, y)$

and that *in* is a possible implicit relation in compound nominals,

(6) $(\forall\, y, z) in(y, z) \supset nn(z, y)$

Then the proof of all but the first conjunct of (3) is straightforward, by backchaining on axioms (5) and (6) into the ground instances of (4). We thus assume $(\exists\, e) call'(e, J)$, which constitutes the new information in the sentence.

Notice now that all of our local pragmatics problems have been solved. "The Boston office" has been resolved to $O$. The implicit relation between Boston and the office has been determined to be the *in* relation. The metonymy has been resolved by coercing "The Boston office" into "John, who works for the Boston office."

This is of course a simple example, and the analysis has shown only the correct interpretation is *possible.* It is the function of the weighted abduction scheme, described in the next section, to choose this interpretation over other possibilities. A more detailed discussion of these issues and more complex examples and arguments are given in Hobbs et al. (1990).

The contrast between the abductive approach and earlier inference models can be described succinctly as follows. In the models proposed by Rieger (1974) and by Sperber and Wilson (1986) one forward-chains from the text and tries to maximize the implications. In the abductive model, one backchains from the text and tries to minimize the assumptions.

### 3.3 Weighted Abduction

Our scheme for weighted abduction (Stickel, 1988) has three features.

First, every conjunct in the logical form of the sentence is given an assumability cost, corresponding to the need for this conjunct to be proven if an adequate interpretation is to be achieved. Thus, conjuncts arising from definite noun phrases and selectional constraints have higher assumability costs than those arising from indefinite noun phrases and main verbs.

Second, this cost is passed back to the antecedents in Horn clauses by a system of weights. Axioms are stated in the form

(7) $P_1^{w_1} \wedge P_2^{w_2} \supset Q$

This says that $P_1$ and $P_2$ imply $Q$, but also that if the cost of assuming $Q$ is $c$, then the cost of assuming $P_1$ is $w_1 c$, and the cost of assuming $P_2$ is $w_2 c$. (Stickel (1988) generalizes this to arbitrary functions of $c$.)

Third, factoring or synthesis is allowed. That is, goal expressions may be unified, in which case the resulting expression is given the smaller of the costs of the input expressions. This feature leads to minimality through the exploitation of redundancy.

In rules like (7), we generally assign the weights so that $w_1 + w_2 > 1$. Thus, less specific assumptions ($Q$) are favored over more specific ones ($P_1 \wedge P_2$). But in

$$P_1^{.6} \wedge P_2^{.6} \supset Q$$

if $P_1$ has already been proved, it is cheaper to assume $P_2$ (for $.6c$) than to assume $Q$ (for $c$). $P_1$ has provided evidence for $Q$, and assuming the "balance" $P_2$ of the necessary evidence for $Q$ should be cheaper.

Factoring can also override less specific abduction. Suppose we have the axioms

$$P_1^{.6} \wedge P_2^{.6} \supset Q_1$$
$$P_2^{.6} \wedge P_3^{.6} \supset Q_2$$

and we wish to derive $Q_1 \wedge Q_2$, where each conjunct has an assumability cost of \$10. Assuming $Q_1 \wedge Q_2$ will then cost \$20, whereas assuming $P_1 \wedge P_2 \wedge P_3$ will cost only \$18, since the two instances of $P_2$ obtained by backchaining can be unified.

Thus, the abduction scheme allows us to adopt the careful policy of favoring less specific abduction while also allowing us to exploit the redundancy of texts for more specific interpretations.

Exactly how the weights and costs should be assigned is a matter of continuing research. It is discussed further in Hobbs et al. (1990). In the remainder of this paper, this issue will be ignored. Our concern will rather be to show that the correct interpretations of metaphors are *possible* in the abductive approach.

In the current TACITUS implementation, whenever an assumption is made, it is checked for consistency. The extension of the abductive approach to metaphor interpretation suggests that this check should be soft rather than hard. Inconsistent assumptions should be allowed if they will result in an otherwise very good interpretation. This is the topic of Section 5.

It might seem that since we use only backchaining to find a proof and a set of assumptions, we cannot use superset information. However, the fact that we can make assumptions enables us to turn axioms around. In general, an axiom of the form

$$\text{species} \supset \text{genus}$$

can be converted into a biconditional axiom of the form

$$\text{genus} \wedge \text{differentiae} \equiv \text{species}$$

Often we will not be able to prove the differentiae, and in many cases we cannot even spell them out. But in our abductive scheme, this does not matter; they can simply be assumed. In fact, we need not state them explicitly at all. We can simply introduce a predicate that stands for all the remaining properties. It will never be provable, but it will be assumable. Thus, in addition to having axioms like that referred to in Section 4.3,

$$(\forall\, y) elephant(y) \supset clumsy(y)$$

we may have axioms like

$$(\forall\, x) clumsy(y)^{.4} \wedge etc_1(y)^{.8} \supset elephant(y)$$

where the weights are distributed roughly according to the "semantic contribution" of each conjunct in the antecedent to the consequent.[3] Then, even though we are strictly backchaining in search for an explanation, the fact that something is clumsy can still be used as (perhaps weak) evidence for its being an elephant, since we can assume the "et cetera" predication $etc_1(x)$ for a certain cost.

This device may seem ad hoc, especially in this paper since it is used in only two axioms where it is essential to get the correct interpretation to go through. On the contrary, however, we view the device as implementing a fairly general solution to the problems of nonmonotonicity in commonsense

[3] Hobbs et al. (1990) addresses the issue of what "semantic contribution" means in probabilistic terms.

reasoning and vagueness of meaning in natural language, very similar to the use of abnormality predicates in circumscriptive logic (McCarthy, 1987). Whereas, in circumscriptive logic, one typically specificies a partial ordering of abnormality predicates in accordance with which they are minimized, in the weighted abduction framework, one uses a somewhat more flexible system of costs.

There is no particular difficulty in specifying a semantics for the "et cetera" predicates. Formally, $etc_1$ in the axiom above can be taken to denote the set of all things that are either not clumsy or are clumsy elephants. Intuitively, $etc_1$ conveys all the information one would need to know beyond clumsiness to conclude something is an elephant. As with nearly every predicate in an axiomatization of commonsense knowledge, it is hopeless to spell out necessary and sufficient conditions for an "et cetera" predicate. In fact, the use of such predicates in general is due largely to a recognition of this fact about commonsense knowledge.

The "et cetera" predicates could be used as the abnormality predicates are in circumscriptive logic, with separate axioms spelling out conditions under which they would hold. However, in the view adopted here, more detailed conditions would be spelled out by expanding axioms of the form

$$(\forall\, x)p_1(x) \wedge etc_1(x) \supset q(x)$$

to axioms of the form

$$(\forall\, x)p_1(x) \wedge p_2(x) \wedge etc_1(x) \supset q(x)$$

An "et cetera" predicate would appear only in the antecedent of a single axiom and never in a consequent. Thus, the "et cetera" predications are only place-holders for assumption costs. They are never proved. They are only assumed.

They constitute one of the principal devices for giving our logic "soft corners". We would expect them to pervade the knowledge base. Virtually any time there is an axiom relating a species to a genus, there should be a corresponding axiom, incorporating an "et cetera" predication, expressing the inverse relation.

### 3.4 The Integrated Framework

The idea of interpretation as abduction can be combined with the older idea of parsing as deduction (Kowalski, 1980, pp. 52-53; Pereira and Warren, 1983). Consider a grammar written in Prolog style just big enough to handle sentence (2).

(7) $(\forall\, i, j, k) np(i, j) \wedge v(j, k) \supset s(i, k)$
(8) $(\forall\, i, j, k, l) det(i, j) \wedge n(j, k) \wedge n(k, l) \supset np(i, l)$

That is, if we have a noun phrase from "inter-word point" $i$ to point $j$ and a verb from $j$ to $k$, then we have a sentence from $i$ to $k$, and similarly for rule (8).

We can integrate this with our abductive framework by moving the various pieces of expression (3) into these rules for syntax, as follows:

(9) $(\forall\, i, j, k, e, x, y, p) np(i, j, y) \wedge v(j, k, p) \wedge p'(e, x)$
$\wedge\, Req(p, x) \wedge rel(x, y) \supset s(i, k, e)$

That is, if we have a noun phrase from $i$ to $j$ referring to $y$ and a verb from $j$ to $k$ denoting predicate $p$, if there is an eventuality $e$ which is the condition of $p$ being true of some entity $x$ (corresponding to $call'(e, x)$ in (3)), if $x$ satisfies the selectional requirement $p$ imposes on its argument (corresponding to $person(x)$), and if $x$ is somehow related to, or coercible from, $y$, then there is an *interpretable* sentence from $i$ to $k$ describing eventuality $e$.

(10) $(\forall\, i, j, k, l) det(i, j, the) \wedge n(j, k, w_1) \wedge n(k, l, w_2)$
$\wedge\, w_1(z) \wedge w_2(y) \wedge nn(z, y) \supset np(i, l, y)$

That is, if there is the determiner "the" from $i$ to $j$, a noun from $j$ to $k$ denoting predicate $w_1$, and another noun from $k$ to $l$ denoting predicate $w_2$, if there is a $z$ that $w_1$ is true of and a $y$ that $w_2$ is true of, and if there is an $nn$ relation between $z$ and $y$, then there is an *interpretable* noun phrase from $i$ to $l$ denoting $y$.

These rules incorporate the syntax in the literals like $v(j, k, p)$, the pragmatics in the literals like $p'(e, x)$, and the compositional semantics in the way in which the pragmatics literals are constructed out of the information provided by the syntax literals.

To parse with a grammar in the Prolog style, we prove $s(0, N)$ where $N$ is the number of words in the sentence. To parse and interpret in the integrated framework, we prove $(\exists\, e) s(0, N, e)$.

One of the appeals of declarative formalisms is that they can be used equally for interpretation and generation. This is true of our framework

$$\begin{tikzcd}
\underline{\textbf{Source Domain}} & \underline{\textbf{Target Domain}} \\
\text{Complex Source Concept} \arrow[r, "3"] & \text{Complex Target Concept} \\
\text{Basic Source Concept} \arrow[u, "2"'] & \text{Basic Target Concept} \arrow[u, "4"'] \arrow[l, "1"']
\end{tikzcd}$$

Figure 1: Analogical Processes Underlying Metaphor.

as well. To generate a sentence describing a given eventuality is to prove $(\exists\, n) s(0, n, E)$. In generation, it is the terminal nodes in the grammar, atoms like $det(i, j, the)$, that are assumed. Thus, in generation, assumptions correspond to actions.

This approach can be extended upward to include the recognition of coherence relations in discourse and downward to include the spellings or pronunciations of words (Hobbs et al., 1990).

The integrated approach suggests a shift of perspective. Initially, the problem of interpretation was viewed as being given certain observable facts, namely, the logical form of the sentence, and finding a proof that demonstrates why they are true. In this section, we no longer set out to prove the observable facts. Rather we set out to prove that we are viewing a coherent situation; it is built into the rules specifying what situations are coherent that an explanation must be found for the observable facts.

# 4 Interpreting Metaphors by Abduction

## 4.1 The Schema for Metaphor

The basic schema for metaphor (and analogy) is that shown in Figure 1.

There are two domains of knowledge, a source domain that is generally very well understood, expressed as a highly elaborated set of axioms, and a target domain, that is generally less well understood.[4] We wish to reason

[4]The term "domain" carries no theoretical weight in this framework. It is just a way of speaking about a group of axioms intuitively perceived to be about the same topic.

or describe something in the target domain. Rather than doing so directly, we map a basic concept in the target domain into a corresponding basic concept in the source domain. We reason or describe in the source domain, with its richer vocabulary and set of axioms, yielding a complex concept in the source domain. Then we map the result back into the target domain, thereby expressing a complex concept there.

Interpreting a metaphor is a matter of reversing this process. We are given a complex concept in the target domain, expressed in the vocabulary of the source domain. The problem is to discover what this expression means by determining how it is composed out of basic concepts in the target domain. To do this, we decompose the complex concept into basic concepts in the source domain, and then undo the analogical mapping to determine the meaning in the target domain.

A computational account of metaphor must specify precisely how each of the arrows in this commuting diagram is realized in a formal system. Our answer is essentially as follows. The relation between domains is taken to be simply identity. Predicates from the source domain will simply be predicated of entities from the target domain. This of course brings with it problems of logical consistency, and how to deal with that is the subject of Section 5. The relations between basic and complex concepts will be those implicational relations encoded in the axioms. Interpreting a metaphor by abduction will then be a matter of backchaining along arrows 3, 2, and 1 to an account in terms of the basic concepts in the target domain.

We show how this works for three examples, a conventionalized metaphor schema, a standard category metaphor whose interpretation depends on context, and a novel metaphor. That these three kinds of cases can be handled in the abductive approach ought to be suggestive of the power of the approach in general.

### 4.2 A Conventionalized Metaphor Schema

The first metaphor to be examined is

(11) The variable N goes from 1 to 100.

Here, the target domain, computer science, is being modeled in terms of the domain of spatial or perhaps more abstract topological relations. This metaphor rests on the core metaphor that identifies a variable having a value with an entity being located *at* some place. This conventionalized identification can be expressed by the following axiom:[5]

(12) $(\forall\, e, x, y) variable(x) \wedge value'(e, y, x) \supset at'(e, x, y)$

That is, if $x$ is a variable and $e$ is the condition of $y$'s being its value, then $e$ is also the condition of $x$ being *at* $y$. The predicate *at* is thus not merely a spatial term but more generally a topological relation locating a figure against a ground.

The relation between the complex predicate *go* and more basic predicates is expressed by the following axiom, capturing the fact that a change in location is a going event:

(13) $(\forall\, e, e_1, e_2, x, y, z) change'(e, e_1, e_2) \wedge at'(e_1, x, y) \wedge at'(e_2, x, z)$
$\supset go'(e, x, y, z)$

That is, if $e$ is a change from state $e_1$ to state $e_2$ where $e_1$ is the state of $x$ being at $y$ and $e_2$ is the state of $x$ being at $z$, then $e$ is a going by $x$ from $y$ to $z$.

Now consider the example. Its logical form is

$$(\exists\, e_0) go'(e_0, N, 1, 100) \wedge variable(N)$$

This is a statement in the target domain, computer science. But we treat it as though it were a statement in the source domain and use source domain axiom (13) to decompose the complex concept *go* into the more basic concepts of *change* and *at*. We then use axiom (12) to interpret the *at* relation. The two atoms $variable(N)$ generated in this way are unified with the identical atom from the logical form, and that condition, the change, and the two value relations are assumed, yielding the minimal interpretation. We thereby have interpreted sentence (11) as asserting a change in value for the variable N. This process is illustrated in Figure 2.

### 4.3 A Category Metaphor

The next metaphor to be examined is

[5]In fact, the existence of such an axiom in the knowledge base is precisely what it means in this framework for the metaphor to be conventional.

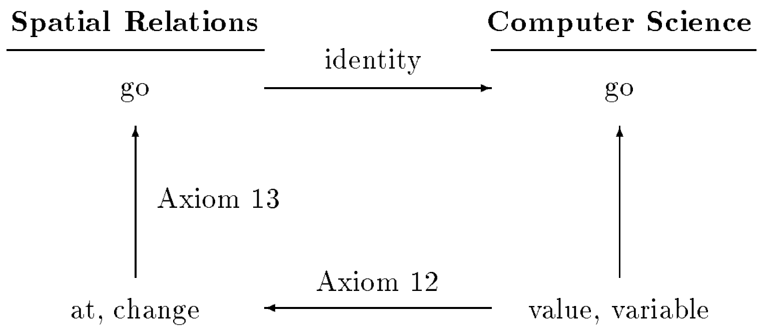


Figure 2: Inference Processes in a Conventionalized Metaphor Schema.

(14) John is an elephant.

A number of suggestions have been made about the appropriate inferences to draw in cases such as this. Ortony et al. (1978) said that it is high salience properties that should be transferred, such as size in the case of elephants. Glucksberg and Keysar (1990) say it is diagnostic properties; that is, in (14), we look for some property of elephants for which an elephant is the prototypical exemplar, such as large size. Carbonell (1982) has argued that abstract properties, rather than physical properties, should be transferred; thus, "has a trunk" should not be transferred. Gentner (1983) has argued that relations (predicates with two or more arguments) are more frequently transferred than monadic properties.

One difficulty with all these suggestions is that they do not depend on context, whereas we know that interpretation always depends on context. Consider the following sentence:

(15) Mary is graceful, but John is an elephant.

The most reasonable interpretation is that John is clumsy. This is not an especially high salience property of elephants. It is not clear that elephants are prototypical exemplars of clumsiness. Clumsiness seems to be intermediate between an abstract and a physical property. And it is not a relation.

The non-abductive analysis of this example was relatively clean in Hobbs (1983). There was an axiom that said elephants are clumsy:[6]

$$(\forall\, y) elephant(y) \supset clumsy(y)$$

That inference was selected because it led to the recognition of a contrast relationship between the two clauses, as signalled by the conjunction "but".

In the abductive approach, the analysis is complicated somewhat by the fact that we can only backchain. This axiom must be rewritten as

$$(16) \quad (\forall\, y) clumsy(y) \wedge etc_2(y) \supset elephant(y)$$

That is, if something is clumsy and some other unspecified properties hold, then it is an elephant.

We will need to introduce a further complication as well, since we will have to refer explicitly to the properties of clumsiness, elephanthood, and grace. Axiom (16) must be rewritten as follows:

$$\begin{aligned}(17) \quad & (\forall\, e_3, y) clumsy'(e_3, y) \wedge etc_2(e_3, y) \\ & \qquad \supset (\exists\, e_2) elephant'(e_2, y) \wedge gen(e_3, e_2)\end{aligned}$$

That is, if $e_3$ is the condition of $y$'s being clumsy and some other unspecified things are true of $e_3$ and $y$, then there is a condition $e_2$ of $y$'s being an elephant. Furthermore, there is a very tight relation between $e_3$ and $e_2$: $y$ is an elephant *by virtue of* its being clumsy and the other things being true. We encode this relation with the predicate *gen*, since it is similar to the "generates" relation common in the philosophical literature.[7]

In Hobbs (1983) the interpretation of (15) was driven by the recognition of a coherence relation between the clauses. In many cases in the abductive approach, especially where a conjunction occurs explicitly, this can be subsumed under the general characterization of interpretation. In (15), the "but" relation is part of the information conveyed by the text, and consequently part of what needs to be explained.[8] We can say that a "but"

[6] Elephants of course are not clumsy, but according to our conventional stereotype, they are. This property is therefore in our "knowledge" base and hence available in metaphor interpretation. Searle (1979) made this point, with respect to gorillas' being "fierce, nasty, prone to violence, and so forth."

[7] The analyses of a large number of phenomena in discourse require an appeal to this "generates" relation between eventualities.

[8] The same interpretation is available for the sentence

relation holds between two eventualities $e_1$ and $e_2$ if they are contradictory properties $p$ and $\neg p$ of two entities $x$ and $y$ that are similar by virtue of sharing some other property $q$:

$$(\forall\, p, q, x, y, e_1, e_2, e_4)p'(e_1, x) \wedge not'(e_2, e_4) \wedge p'(e_4, y) \\ \wedge\, q(x) \wedge q(y) \supset but(e_1, e_2)$$

This however is too strong. It may be that the contrast is between not $e_1$ and $e_2$ but between eventualities related to $e_1$ and $e_2$. In the case of example (15), the contrast is between $e_1$ and an eventuality related to $e_2$, so for this example we will rewrite the above axiom as follows:

(18) $$(\forall\, p, q, x, y, e_1, e_2, e_3, e_4)p'(e_1, x) \wedge not'(e_3, e_4) \wedge p'(e_4, y) \\ \wedge\, gen(e_3, e_2) \wedge q(x) \wedge q(y) \supset but(e_1, e_2)$$

That is, a "but" relation holds between $e_1$ and $e_2$ if there is a $p$ such that $e_1$ is $p$'s being true of some $x$, and there is an $e_3$ that generates $e_2$ that is the negation of an $e_4$ which is $p$'s being true of some $y$, and there is some $q$ true of $x$ and $y$. (This axiom is second-order, but not seriously so, if we restrict the instantiations of the predicate variables to predicate constants.)

Next we need an axiom relating clumsiness and grace.

(19) $$(\forall\, e_3, e_4, y)not'(e_3, e_4) \wedge graceful'(e_4, y) \supset clumsy'(e_3, y)$$

That is, if $e_3$ is the condition of $e_4$ not being true, where $e_4$ is the condition of $y$'s being graceful, then $e_3$ is the condition of $y$'s being clumsy.

Suppose we also know that Mary and John are people:

$$person(M),\ person(J)$$

Now we are ready to interpret sentence (15). Its logical form is

$$(\exists\, e_1, e_2)graceful'(e_1, M) \wedge elephant'(e_2, J) \wedge but(e_1, e_2)$$

---

Mary is graceful; John is an elephant.

This is because the mere adjacency of the clauses conveys information—namely, that the two situations are somehow related. One possible relation is contrast, which is characterized roughly as "but" is. But other relations are theoretically possible as well, leading to the sense we have that it is less certain that the interpretation of the second clause is "John is clumsy". The next clause might be "Mary can dance on his back," in which case the second clause would not be in contrast with the first but background for the third, and John would be a real rather than a metaphorical elephant.

$$
\begin{tikzcd}
\textbf{Animals} & & \textbf{People} \\
\text{elephant} \arrow[rr, "\text{identity}"] & & \text{“elephant”} \\
\text{clumsy} \arrow[u, "\text{Axiom 17}"'] & & \text{clumsy} \arrow[ll, "\text{identity}"'] \arrow[u]
\end{tikzcd}
$$

Figure 3: Inference Processes in a Category Metaphor.

We can then backchain on axiom (17) from "elephant" to "clumsy", assume $etc_2(e_3, J)$, backchain on axiom (19) from "clumsy" to "not graceful", and assume $not'(e_3, e_4)$ and $graceful'(e_4, J)$. We also assume $graceful'(e_1, M)$. Then we have a proof of $but(e_1, e_2)$, using axiom (18), with $p$ instantiated to $graceful$ and $q$ instantiated to $person$.

We have thereby interpreted the metaphor. Figure 3 illustrates the interpretation of "elephant", although it was the requirement to explain the "but" relation that drove the interpretation.

This account is somewhat more complex than that given in Hobbs (1983), but every complication is independently motivated. By subsuming metaphor interpretation under a general account of interpretation, we more than justify the moderate increase in complexity.

### 4.4 A Novel Metaphor

The last metaphor to be examined occurred in *Newsweek* in a quote by an American Congressman complaining that the bills the Congress passes are too easy for the President to veto.

(20) We insist on serving up these veto pitches that come over the plate the size of a pumpkin.

This metaphor evokes a mapping between a schema for the passage of a bill into law by the American government and a schema for baseball. The Congress schema says that Congress sends a bill to the President, and

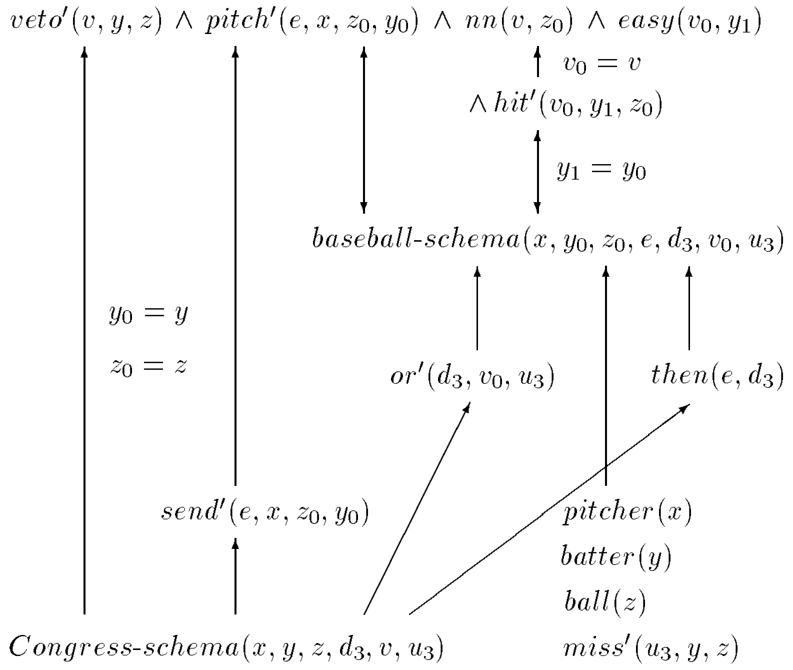


Figure 4: Abductive Proof for Novel Metaphor.

then the President signs or vetoes it. The baseball schema says that the pitcher sends a ball to the batter, and then the hitter either hits or misses it or hits a foul ball. The mapping identifies the Congress with the pitcher, the President with the batter, and the bill with the ball. Vetoing the bill corresponds to hitting the ball.

To clarify the exposition of this example, variables in axioms and the logical form that will be unified in the final proof will be represented with the same letter, differently subscripted. Unsubscripted variables and the subscripts 0 and 1 will be used in the logical form. The subscript 2 is used in axiom (21), 3 in axiom (22), and 4 in axiom (23). Thus, $x$ is Congress in the logical form, $x_2$ is Congress in axiom (21), $x_3$ is the pitcher in axiom (22), and $x_4$ is the sender/pitcher in axiom(23).

A diagram of the proof is presented in Figure 4.

In the abductive framework, a schema is represented as an axiom that has in the antecedent a "schema predication", with all of the role fillers as arguments, and in the consequent the various properties of the role fillers.

A schema is often used in the interpretation of a text because assuming the single schema predication explains so much of the content of the text; it is one way of arriving at a minimal abductive proof. Several different schemas can be used simultaneously by simply assuming or proving their schema predications, with their variables instantiated in the appropriate ways.

The Congress schema would thus be encoded in the following axiom:

$$\begin{aligned}(21)\quad & (\forall\, z_1, e_1, d_1, v_1, u_1) Congress\text{-}schema(x_2, y_2, z_2, e_2, d_2, v_2, u_2) \\ & \quad \equiv Congress(x_2) \wedge President(y_2) \wedge bill(z_2) \\ & \quad\quad \wedge send'(e_2, x_2, z_2, y_2) \wedge or'(d_2, v_2, u_2) \wedge veto'(v_2, y_2, z_2) \\ & \quad\quad \wedge sign'(u_2, y_2, z_2) \wedge then(e_2, d_2)\end{aligned}$$

That is, there is a Congress schema situation involving the listed variables, if and only if $x_2$ is Congress, $y_2$ is the President, $z_2$ is a bill, $e_2$ is a sending by Congress of the bill to the President, $d_2$ is the disjunction of the vetoing $v_2$ by the President of the bill and the signing $u_2$ by the President of the bill, and $d_2$ happens after $e_2$.

The baseball schema would be encoded in the following axiom:

$$\begin{aligned}(22)\quad & (\forall\, x_3, y_3, z_3, e_3, d_3, v_3, u_3) baseball\text{-}schema(x_3, y_3, z_3, e_3, d_3, v_3, u_3) \\ & \quad \equiv pitcher(x_3) \wedge batter(y_3) \wedge ball(z_3) \\ & \quad\quad \wedge pitch'(e_3, x_3, z_3, y_3) \wedge or'(d_3, v_3, u_3) \wedge hit'(v_3, y_3, z_3) \\ & \quad\quad \wedge miss'(u_3, y_3, z_3) \wedge then(e_3, d_3)\end{aligned}$$

That is, there is a baseball schema situation involving the listed variables, if and only if $x_3$ is the pitcher, $y_3$ is the batter, $z_3$ is the ball, $e_3$ is a pitching by the pitcher of the ball to the batter, $d_3$ is the disjunction of the hitting $v_3$ by the batter of the ball and the missing $u_3$ by the batter of the ball, and $d_3$ happens after $e_3$.

Both schema axioms have been written as biconditionals, but only axiom (22) will be used in both directions in this example.

One further axiom is required, relating sending and pitching.

$$\begin{aligned}(23)\quad & (\forall\, e_3, x_3, y_3, z_3) send'(e_3, x_3, z_3, y_3) \wedge etc_3(e_3, x_3, y_3, z_3) \\ & \quad\quad \supset\ pitch'(e_3, x_3, z_3, y_3)\end{aligned}$$

That is, if $e_3$ is a sending by $x_3$ of $z_3$ to $y_3$ and some other unspecified conditions hold for these entities, then $e_3$ is a pitching by $x_3$ of $z_3$ to $y_3$.

To focus on the metaphor, we will simplify the example somewhat. We assume that "that come over the plate the size of a pumpkin" has already been interpreted to mean "that are easy to hit". We will also assume that "we" has been resolved to Congress and that we therefore know $Congress(x)$. We assume that the fact that the control verb "serving up" is applied to "pitches" has enabled us to determine that it is Congress that is pitching. Then the logical form of the portion of the sentence that concerns us would be as follows, where all the variables are existentially quantified:

$$\ldots \wedge veto'(v, y, z) \wedge pitch'(e, x, z_0, y_0) \wedge nn(v, z_0) \wedge easy(v_0, y_1) \\ \wedge hit'(v_0, y_1, z_0) \wedge \ldots$$

That is, there is a vetoing event $v$ by someone $y$ of something $z$, there is a pitching event $e$ by Congress $x$ of something $z_0$ to someone $y_0$, the vetoing $v$ is related somehow ($nn$) to the thing $z_0$ that is pitched, and the hitting $v_0$ by someone $y_1$ of $z_0$ is easy for $y_1$.

This expression needs to be proved abductively. The conjuncts $pitch'(e, x, z_0, y_0)$ and $hit'(v_0, y_1, z_0)$ can be proved by assuming the baseball schema, instantiated as follows:

$$baseball\text{-}schema(x, y_0, z_0, e, d_3, v_0, u_3)$$

so that $y_0$ and $y_1$ (the one to whom something is pitched and the one who finds it easy to hit) are identified.

But now the baseball schema axiom can be used in the opposite direction, to backchain into the following predications:

$$pitch'(e, x, z_0, y_0),\ or'(d_3, v_0, u_3),\ then(e, d_3)$$

By assuming the "et cetera" predication of axiom (23), we can backchain from the first of these into

$$send'(e, x, z\, y_4)$$

But now if we assume the remaining conjuncts on the right-hand side of axiom (22),

$$pitcher(x),\ batter(y_0),\ ball(z_0),\ miss'(u_3, y_0, z_0)$$

then we can prove the baseball schema from the Congress schema, instantiated as follows:

$$Congress\text{-}schema(x, y_0, z_0, e, d_3, v_0, u_3)$$

$$
\begin{tikzcd}
\underline{\textbf{Baseball}} & \underline{\textbf{Congress}} \\
\text{pitch, hit} \arrow[r, "\text{identity}"] & \text{veto} \\
\text{Baseball Schema} \arrow[u, "\text{Axiom 22}"'] & \text{Congress Schema} \arrow[u, "\text{Axiom 21}"'] \arrow[l, "\text{Axioms 21, 23, and 22}"']
\end{tikzcd}
$$

Figure 5: Inference Processes in a Novel Metaphor.

Thus, Congress is identified as a pitcher, the President as a batter, and so forth.

We can prove $veto'(v, y, z)$ in the logical form from this instantiation of the Congress schema, yielding

$$Congress\text{-}schema(x, y, z, e, d_3, v, u_3)$$

so that the implicit agent $y$ of the vetoing gets identified with the President/batter $y_1$, the vetoing $v$ gets identified with the hitting $v_0$, and the thing vetoed $z$ is identified with the thing pitched $z_0$.

Finally, $nn(v, z)$ must be instantiated to some relationship between the vetoing $v$ and the thing pitched $z$, and either $veto'(v, y, z)$ or $hit'(v, y, z)$ will do. In fact, maximum redundancy is achieved by assuming these two relations are the same, and this is one reason the vetoing and hitting are identified, rather than the vetoing and the missing.

Thus, by assuming the Congress schema, assuming what was needed to infer a pitching from a sending, and assuming the missing pieces of the baseball schema, we derived the baseball schema and identified the corresponding actions, thereby interpreting the novel metaphor. The two schemas together then accounted for the propositional content of the sentence.

This is illustrated in Figure 5. We assume the Congress Schema, and this allows us, via axiom (21), to explain "veto". Axioms (21), (23), and (22), the last used from right to left, establish the baseball schema, from which we can infer "pitch" and "hit", by axiom (22) used from left to right. This results in the identification of the hitting and the vetoing.

## 5 Metaphor and Truth

There is a problem in our account of how examples (15) and (20) are interpreted. They require the assumption of propositions that are not true. John is not an elephant, and the bill is not a ball.[9] Here I will sketch two possible solutions to this problem.

**1. Predicate Coercion:** When a predicate is applied to an argument for which it is not appropriate, there are two possible interpretive moves. We can decide that the argument actually refers to something other than it denotes explicitly—we can coerce the argument into something related to it. This is metonymy. Or we can decide that the predicate actually denotes a property other than the one it denotes explicitly—we can coerce the predicate to a related predicate. Metaphor is one example of this. Thus, we are likely to interpret

> Dan Quayle couldn't read Plato.

metonymically. We coerce from Plato to the dialogues written by Plato. By contrast, we are likely to interpret

> George Bush couldn't read Saddam Hussein.

metaphorically. We coerce the predicate, taking "read" to mean "understand".[10]

In metonymy, one applies a coercion function $k$ to the argument, transforming $p(x)$ into $p(k(x))$, in order to achieve congruence between the predicate and the argument. In our functionless notation, this becomes

$$p(y) \wedge Req(p, y) \wedge rel(y, x)$$

where $Req(p, y)$ is the requirement that $p$ imposes on its argument, and $rel(y, x)$ expresses the coercion relation between $y$ and $x$.

By analogy, this would suggest that metaphor be handled by applying a coercion function to the predicate, transforming $p(x)$ into $k(p)(x)$. In our functionless notation, this becomes

$$q(x) \wedge Req(q, x) \wedge Rel(q, p)$$

[9] Example(11) does not run into this problem only because *at* has already by convention been abstracted into a topological relation more general than simple spatial location.

[10] A process of predicate coercion is also proposed by Nunberg (1991) and by Bobrow et al. (1991) for different but compatible purposes.

The predicate $q$ would be new. The question arises as to what $q$ means. The meaning—even the denotation—of a predicate is determined or at least constrained by the axioms it occurs in. To construct a new predicate $q$ we would have to specify the axioms in which it occurs. In abductive interpretation, a subset of the axioms involving $p$ are selected as relevant to the sentence, i.e., those that lead to the best explanation of the content of the sentence. If we substitute $q$ for $p$ in just these axioms, we have thereby specified the axioms for $q$ and hence delimited its meaning. We are in effect achieving what selective inferencing did in Hobbs (1983). The uncertainties in this process correspond to the uncertainties we sometimes experience in interpreting metaphors; they are one of the sources of the metaphor's power.

In this approach *elephant*, in the context of example (15), is coerced into another predicate, which is implicationally related to *clumsy* and perhaps to *large*, but does not imply the existence of a trunk.

**2. Interpretation and Judgment:** In the abductive approach to interpretation, we make assumptions when we are unable to prove something (for less cost). But these assumptions can play a number of different roles. In Section 3.2 and in many cases in Section 4, assumptions are used to accept new information. More generally, however, they can be used to accommodate speakers, whatever they say. Suppose someone says,

John called Mary a Marxist, and she insulted him too.

To interpret this sentence, we must assume that it is bad to be a Marxist, or at least that the speaker believes it, even though we may not believe it ourselves. But we will not normally change our beliefs as a result of interpreting this sentence.

One can also use assumptions to interpret one's way past a mistake. Suppose the approach of Section 3.4 were carried all the way down to the level of individual letters. We would have the following rule to tell us how the pronoun "it" is spelled:

$$(\forall\, i, j, k) I(i, j) \wedge T(j, k) \supset pro(i, k, it)$$

That is, if there is the letter "i" between point $i$ and point $j$ and the letter "t" between point $j$ and point $k$, then there is the pronoun "it" between point $i$ and point $k$. If we then encountered the string

If is easy to please John.

we could interpret it by assuming that the second letter is in fact a "t" and not an "f". This may yield the best interpretation of the entire set of

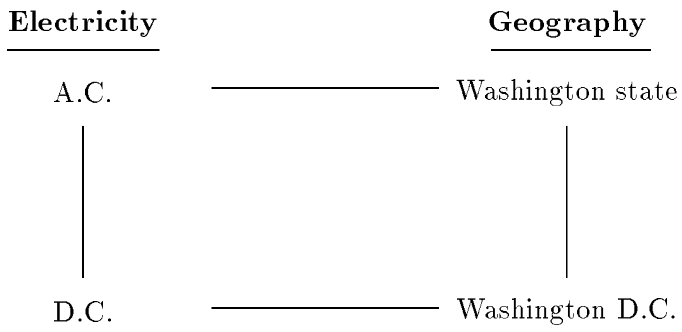


Figure 6: Cognitive Processes in a Local Convention.

observables in the text, even though it is flatly contradicted by one of the observables.

Assumptions can also be used for adopting and using local conventions. Consider the following exchange (due to John DuBois.)

A: I'm going to Washington tomorrow.
B: D.C. or A.C.?
A: A.C.

In this dialogue "A.C." has come to mean "state", and assuming this equivalence is necessary for interpreting it, as illustrated in Figure 6. But that does not mean this identification will persist to other dialogues.

Finally, in an approach to generation presented in Hobbs et al. (1990) based on the integration described in Section 3.4, assumptions correspond to actions, namely, the uttering of words.

We have said that to interpret a text is to find the best explanation for why it *would be* true, not why it *is* true. Deciding whether something *is* true is a logically (though not necessarily chronologically) separate process, one that we can call *judgment*.

We can take metaphor interpretation as working in much the same way. We make certain assumptions in order to interpret the metaphor, such as that John is an elephant and that the bill is a ball, and then in a logically separate judgement step, we decide which of our assumptions we are in fact prepared to believe.

Let us carry this approach one more step both toward formalization and toward embedding it in a larger framework. In Hobbs et al. (1990), it is suggested that a rational agent can be seen as going through the world, continuously trying to prove abductively the proposition "The situation I am in is coherent and I will thrive in it".[11] The first clause generates interpretation, the second action. One kind of coherent situation, involving both interpretation and action is a turn in a conversation. Here there is a speaker $S$, a hearer $H$, and an utterance $u$. The utterance is an action on the part of $S$ that serves in the achievement of $S$'s goals. The utterance has an interpretation $\phi$, which we may think of as a set of propositions. The hearer makes some kind of judgment about the information contained in $\phi$. This can all be expressed by the rule

$$(\forall\, u, \phi, H, S) Serve\text{-}Goal(u, S) \wedge Interp(u, \phi) \wedge Judge(H, \phi)$$
$$\supset Turn\text{-}in\text{-}Conversation(u, S, H)$$

That is, if an utterance $u$ serves a goal of the speaker $S$, the interpretation of $u$ is $\phi$, and the hearer $H$ judges $\phi$, then there is a turn in a conversation in which $S$ utters $u$ to $H$.

A small set of axioms enable backchaining from $Interp(u, \phi)$ into the whole abductive framework of interpretation described in this paper, via a "grammar" of the sort described in Section 3.4. One may think of this as the entry into the *informational* aspect of discourse.

Other axioms having $Serve\text{-}Goal(u, S)$ as their consequent would tap into the whole *intentional* aspect of discourse, as elucidated in the work of Cohen and Perrault (1979) and many others. Thus, there might be an axiom that says that if $H$'s believing $\phi$ serves a goal of $S$, then $u$ serves a goal of $S$.

Because the conjuncts $Interp(u, \phi)$ and $Serve\text{-}Goal(u, S)$ share variables, the informational and intentional aspects can influence each other. Thus, what might otherwise be the best interpretation of an utterance could be rejected if there is no way to relate it to the speaker's goals.

Finally, a first cut at an expansion of $Judge(H, \phi)$ might go as follows. To judge $\phi$ one must judge each proposition $P$ in $\phi$. There are three possibilities for $P$. $P$ may already be mutually known, the given, in which case there is nothing to do. $P$ may be inconsistent with what is already known, in which case it is judged false and rejected. Otherwise, $P$ will be entered into the

[11] Actually, one could see the first clause as a precondition for the second.

knowledge base, as mutual knowledge. This of course is oversimplified. In fact, the conjunct $Judge(H, \phi)$ taps into the whole domain of belief revision.

In this account, it would be perfectly normal in the course of interpretation to assume a proposition that is known to be false. The judgment as to whether it should become a permanent belief is part of a logically separate step.

The predicate coercion solution to the problem of metaphor and truth has the advantage of giving an analogous treatment to metaphor and metonymy. Its disadvantage is that it involves a significant increase in notational complexity. The judgment solution has the advantage of requiring nothing that is not already required in a larger framework for discourse interpretation and generation, but of course means that the details of that framework must be worked out.

## Acknowledgements

I have profited from discussions on this work with Douglas Appelt, Joanna Moore, and the participants of the NATO workshop on Computational Models of Communication in Trento, Italy, November 1990. I would also like to acknowledge three anonymous reviewers of an earlier version of this paper, the uniformity of whose largely inappropriate criticisms led me to suspect shortcomings in the exposition. This research was funded by the Defense Advanced Research Projects Agency under Office of Naval Research contracts N00014-85-C-0013 and N00014-90-C-0220, and by a gift from the Systems Development Foundation.

## References

[1] Bobrow, Robert, Robert Ingria, and David Stallard, 1991. "The 'Mapping Unit' Approach to Subcategorization", *Proceedings of the Fourth DARPA Workshop on Speech and Natural Language*, Asilomar, California.

[2] Carbonell, Jaime, 1982. "Metaphor: An Inescapable Phenomenon in Natural-Language Comprehension" In W. Lehnert and M. Ringle (Eds.), *Strategies for Natural Language Processing*, pp. 415-434, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

[3] Charniak, Eugene, and Robert Goldman, 1988. "A Logic for Semantic Interpretation", *Proceedings, 26th Annual Meeting of the Association for Computational Linguistics*, pp. 87-94, Buffalo, New York, June 1988.

[4] Cohen, Philip, and C. Raymond Perrault, 1979. "Elements of a Plan-based Theory of Speech Acts", *Cognitive Science*, Vol. 3, No. 3, pp. 177-212.

[5] Davidson, Donald, 1967. "The Logical Form of Action Sentences", in N. Rescher, ed., *The Logic of Decision and Action*, pp. 81-95, University of Pittsburgh Press, Pittsburgh, Pennsylvania.

[6] Gentner, Dedre, 1983. "Structure-Mapping: A Theoretical Framework for Analogy", *Cognitive Science*, vol. 7, pp. 150-170.

[7] Glucksberg, Sam, and Boaz Keysar, 1990. "Understanding Metaphorical Comparisons: Beyond Similarity", *Psychological Review*, Vol. 97, No. 1, pp. 3-18.

[8] Hobbs, Jerry R., 1983. "Metaphor Interpretation as Selective Inferencing: Cognitive Processes in Understanding Metaphor", *Empirical Studies in the Arts*, Vol. 1, No. 1, pp. 17-34, and Vol. 1, No. 2, pp. 125-142.

[9] Hobbs, Jerry R. 1985. "Ontological Promiscuity." *Proceedings, 23rd Annual Meeting of the Association for Computational Linguistics*, pp. 61-69. Chicago, Illinois, July 1985.

[10] Hobbs, Jerry R., Mark Stickel, Paul Martin, and Douglas Edwards, 1988. "Interpretation as Abduction", *Proceedings, 26th Annual Meeting of the Association for Computational Linguistics*, pp. 95-103, Buffalo, New York, June 1988.

[11] Hobbs, Jerry R., Mark Stickel, Douglas Appelt, and Paul Martin, 1990. "Interpretation as Abduction", SRI Technical Note 499, SRI International, Menlo Park, California. December 1990.

[12] Indurkhya, Bipin, 1987. "Approximate Semantic Transference: A Computational Theory of Metaphors and Analogies", *Cognitive Science*, vol. 11, no. 4, pp. 445-480, October-December 1987.

[13] Kowalski, Robert, 1980. *Logic for Problem Solving*, North Holland, New York.

[14] McCarthy, John, 1987. "Circumscription: A Form of Nonmonotonic Reasoning", in M. Ginsberg, ed., *Readings in Nonmonotonic Reasoning*, pp. 145-152, Morgan Kaufmann Publishers, Inc., Los Altos, California.

[15] Norvig, Peter, 1987. "Inference in Text Understanding", *Proceedings, AAAI-87, Sixth National Conference on Artificial Intelligence*, Seattle, Washington, July 1987.

[16] Nunberg, Geoffrey, 1991. "Interpretations of Indexicals and Descriptions", manuscript.

[17] Ortony, Andrew, R. Reynolds, and J. Arter, 1978. "Metaphor: Theoretical and Empirical Research", *Psychological Bulletin*, Vol. 85, No. 5, pp. 919-943.

[18] Pereira, Fernando C. N., and David Warren, 1983. "Parsing as Deduction", *Proceedings of the 21st Annual Meeting, Association for Computational Linguistics*, pp. 137-144. Cambridge, Massachusetts, June 1983.

[19] Rieger, Charles J., III., 1974. "Conceptual Memory: A Theory and Computer Program for Processing the Meaning Content of Natural Language Utterances", Memo AIM-233, Stanford Artificial Intelligence Laboratory, Stanford University.

[20] Searle, John, 1979. "Metaphor", In A. Ortony (Ed.), *Metaphor and Thought*, pp. 92-123, Cambridge University Press, Cambridge, England.

[21] Sperber, Dan, and Deirdre Wilson, 1986. *Relevance: Communication and Cognition*, Harvard University Press, Cambridge, Massachusetts.

[22] Stickel, Mark E., 1988. "A Prolog-like Inference System for Computing Minimum-Cost Abductive Explanations in Natural-Language Interpretation", *Proceedings of the International Computer Science Conference-88*, pp. 343–350, Hong Kong, December 1988.

[23] Stickel, Mark E. 1989. "A Prolog Technology Theorem Prover: A New Exposition and Implementation in Prolog", Technical Note No. 464, SRI International, Menlo Park, California.