CHAPTER 18

Metaphor and Artificial Intelligence

Why They Matter to Each Other

John A. Barnden

Introduction

Why is Artificial Intelligence concerned with metaphor, and what special contributions can AI offer to metaphor research? This chapter will indicate why AI needs to study metaphor and will outline what AI has been contributing to the illumination of metaphor, whether it is processed by artefacts or by the human mind.

Specific contributions AI research on metaphor that one can already point to, and that will be addressed to varying extents in this chapter, include the following: creation of detailed mechanisms for reasoning within the terms of the source-domain in a metaphor, in order expand the relevance of known source-target mappings; increased emphasis on uncertainty and gradedness in metaphorical reasoning; a richer view of overriding (source-over-target as well as target-over-source); mechanisms for exploiting context; important steps towards integration with metonymy interpretation; some emphasis on disanalogy and a limitation of the role of parallelism between source and target; the usefulness of reversed transfers (transfers from source domain to target domain); the importance of non-assertional metaphor; increased doubt about whether the notion of a "domain" is actually important and well-founded; and clarification of ways in which literal meaning can be involved in metaphor interpretation.

The plan of the chapter is as follows. The next section will make some observations about AI, explain why metaphor is important to applications-oriented aspects of AI, and indicate why, in general terms, AI can make distinctive contributions to the study of cognition as a whole, metaphor included. Then a new section will sketch five different, relatively recent AI research works on metaphor. This is to set the scene for the following section, which will discuss specific contributions of AI to metaphor research. The issues will be summarized in a brief concluding section. The chapter does not attempt to survey AI research on metaphor completely or to provide a history of this work, despite the fact that AI has long had an interest in metaphor (cf. e.g. Carbonell, 1980, 1982; Norvig, 1989; Russell, 1976, 1985; Way, 1991; Weber, 1989; Weiner, 1984; Wilks, 1978), and also simile (e.g. Winston, 1979) and analogy (see Hall, 1989, for a review). Readers interested in AI work not covered here may also wish to look at Martin's (1996) and Russell's (1986) reviews and the extensive review in Fass (1997, chap. 11). Also, we omit description of work on theoretical approaches to metaphor that while being interesting and important in themselves do not address processing issues to any large extent, such as the approaches of Asher and Lascarides (1995), Hintikka and Sandu (1990), Indurkhya (1991, 1992), van Genabith (2001), and Vogel (2001). For reasons of space we omit description of computational study of metaphor in corpora (e.g. Mason, 2004) despite some close connections to AI. The chapter makes some mention of metonymy because of the close connection of metaphor and metonymy and because, as we will see, some major AI work on metaphor also addresses metonymy.

Artificial Intelligence

AI has at least three separate, though interrelated, aims:

An "engineering" aim: To engineer, or provide computational principles and methods for engineering, *useful artefacts* that are arguably intelligent, without necessarily having any mechanistic similarity to human or animal minds/brains. The usefulness may be in an industrial domain or an everyday, practical domain, but may also be in other domains such as art or mathematical theorem proving.

A "psychological" aim: To devise computational principles, computationally detailed theories, or running computational systems that provide a basis for possible testable accounts of *cognition in human or animal minds/brains*.

A "general/philosophical" aim: To devise computational principles, computationally detailed theories, or running computational systems that serve as or suggest possible accounts of cognition *in general*, whether it be in human-made artefacts, in naturally occurring organisms, or in cognizing organisms yet to be discovered, or that illuminate *philosophical issues* such as the nature of mind, thought, intelligence, consciousness, perception, language, representation, learning, rationality, society, and so on . . . not forgetting computation itself.

On top of this multiplicity of aims, the word "intelligence" is usually taken very broadly in the field, to cover not only pure rational thought but also almost anything that could come under the heading of "cognition," "perception," "language use," "emotion" and so forth. Thus, the name "artificial intelligence" has always been somewhat of a *nom de plume*, with both parts of the name each hinting at only one aspect of the nature of the actual endeavour.

The three aims are often inextricably combined in a given piece of research. For one thing, an individual researcher may subscribe to more than one of the aims. But also, of course, developments in pursuit of any one of the aims could happen to inspire advances towards one of the others, and endeavours towards any one of the aims can proactively look for inspiration from research towards the others.

Before going on, it is useful to explain why metaphor is important for the Engineering aim of AI. Many intelligent artefacts that need to communicate well with people using human language will need to be able to cope with metaphor. Metaphor is prevalent in human linguistic discourse, even when it is just mundane conversation. Slightly more indirectly, some intelligent artefacts need to understand linguistic communication between people, for instance for the purpose of understanding newspaper articles written by people for other people. Indeed, metaphor is becoming an increasingly looming obstacle for Engineering AI, as attempts are made to bring better automated human-language processing into commercial products, to develop ever more advanced computer interfaces and virtual reality systems, to develop automated understanding and production of emotional expression given that this is often conveyed explicitly or implicitly by metaphor (Delfino & Manea, 2005; Emanatian, 1995; Fainsilber & Ortony, 1987; Fussell & Moss, 1998; Kövecses, 2000; Thomas, 1969; Yu, 1995), and also of gesture and sign language given that these forms of communication have strong metaphorical aspects (McNeill, 1992; P. P. Wilcox, 2004; S. Wilcox, 2004; Woll, 1985).

To return to the set of aims overall, their multiplicity, and their nature taken individually, cause problems in the evaluation of developments in AI. Engineering developments can clearly be evaluated on the basis of actual usefulness or promise of such, but the nature of evaluation is more difficult for the other aims. Evaluation can be on criteria such as coherence, simplicity, computational efficiency and so forth, and on whether the development in question does in principle achieve the intended cognitive ends, but beyond that the evaluation must be in the indirect, long-term, and subjective sense of the extent to which the development contributes eventually to other fields such as Philosophy or Psychology, or is at least perceived as embodying interesting and inspiring ideas for these fields. Since Psychology is currently the locus of intensive research on metaphor, it is worth stressing that within the Psychological aim there is not necessarily any goal to produce an immediately testable psychological theory. Rather, the aim is creatively to provide computationally well-founded and well-designed bases from which psychologists or others could proceed to develop testable theories.

I hope that in the descriptions of the three aims above the reader will have observed the hedging about whether the AI developments are actually "implemented" (that is, realized in the form of computer software or hardware). Hence the mention of *computational principles, methods and computationally detailed theories*, not just working computational systems. A product of AI research does not have to be a working computer program or piece of computer hardware. Rather, it can be a system description or formal logical account that is detailed and specific enough from which software or hardware could readily if laboriously be developed. It can also be a description of new types of representation, inferencing or other processing that could form part of an AI system (implemented or otherwise).

Such products of AI may be left without implementation not through neglect but rather because they can be assessed, to a useful degree, in terms of their coherence, effectiveness, efficiency, interest, distinctiveness, and so on without being implemented. Also, the act of creating the product can uncover problems and issues that would be unlikely to arise in less detailed and specific theorizing. Much of the point of creating even a *working* AI system is not so much to use it in practice but to serve just such ends as uncovering problems and gaps, studying the relationship to other proposed systems, and so on. In short, much of the point of developing a detailed computational account, implemented or not, is aid in the development of *principles, methods* and *theories* in more detail and with greater security than would otherwise be likely.

These explanations about AI could be paralleled to some extent by observations about Computer Science in general. Much research in Computer Science is not directly about producing working software of hardware. For example, much of the field is mathematical theory directed at the nature of computation, the complexity of algorithms, the abstract meaning of computer programs, and the well-founded design of programming languages and computer systems.

Given that the Psychological and General/Philosophical aims of AI impinge on the concerns of other disciplines the question arises as to whether AI research has anything special to offer to such disciplines over and above what they can do by themselves. There are several reasons for a positive answer. First, AI has special expertise in a wide variety of different forms of computation, in putting them on a proper, well-thought-out foundation and, importantly, in finding complicated combinations of or compromises between different forms. The hope is that a strong Computer Science background or context enables many AI researchers to come up with suggestions that are, in computational ways, more advanced, richer, more subtle, more complex, more formally coherent, and/or more extensively and securely developed than is generally possible in other disciplines, with their own demands and pressures concerning other matters such as proper experimental design.

Pressure towards developing effective compromises and combinations comes from the applications focus within the Engineering aim, and from the focus in all three aims of the production of working artefacts or at least *detailed* computational schemes and methods. These foci can also provide a useful "sanity check," helping for example to uncover unwelcome but difficult-to-discern interactions between parts of a theory, to avoid vagueness in descriptions of representations and processes, to avoid over-simplification, and to ensure greater coverage of underlying technical issues than in other fields.

AI Research on Metaphor: An Illustrative Review of Recent Work

In outlining the nature of AI above we looked at some general reasons why AI is in a position to make helpful contributions to the study of cognition, or, at least, why it is in a better position to make certain types of advance than other disciplines are. As for specific metaphor research issues on which AI is in a relatively good position to be helpful, we will examine some of them after reviewing, in this section, a handful of particular metaphor research works within AI.

Hobbs

Important work on metaphor in AI was done by Hobbs (1990, 1992). The ideas do not seem to have met with a substantial implementation effort, but Hobbs has devised a detailed computational account from which implementations could be developed reasonably readily as an extension to the implemented TACITUS system (Hobbs et al., 1993). We can divide the work into the following three strands:

1. *Unmodified-property transfer*: When X is metaphorically described as Y, this method can attribute to X a property P of Y, provided P also makes sense for X without modification. A simple example is interpreting "John is an elephant" to mean that John is clumsy, given that clumsiness is (let us assume) a property of elephants, and given that it can also be applied to people.
2. *Transfer by known mappings within inference*: This method uses known mappings between aspects of the source domain and aspects of the target domain. Importantly, unmapped aspects of the source domain can be used in a metaphor by virtue of their source-domain inferential connections to the source-domain elements that *are* mapped by known mappings. Also, the mappings are themselves cast as inference rules (see below). Thus, uses of mappings are just inference steps along with any other.
3. *Mapping discovery by analogy*: hypothesizing mappings between complex situations in source and target from scratch, by means of structural matching, in order to handle metaphor that is novel (to the understander).

All three strands are placed (in Hobbs, 1992) within a general inferential framework for natural language understanding, which, in particular, also handles metonymy. This framework has abduction as its guiding principle and its central means of inference. In essence, linguistic expressions are regarded as outward signs that are given rise to by underlying situations that are conveyed by the expressions, and the understander's task is to abductively move from the outward signs to the underlying situations. A crucial aspect of Hobbs's overall abductive approach is that it is, thereby, an approach founded on *uncertain* inference.

In the first strand, Unmodified-Property Usage, Hobbs has an appealing, context-driven view of how the properties are selected in a given case. As he says, "John is an elephant" cannot be precisely interpreted outside of context (Hobbs, 1990, 59). But he claims that, given suitable context, coherence considerations can lead to a precise interpretation. Thus, Hobbs asserts that "Mary is graceful, but John is an elephant" suggests the interpretation that John has a property that contrasts with gracefulness. If it is known that elephants are clumsy, and this is the only elephant property that is an opposite of gracefulness, then the clumsiness interpretation is secured. This context-driven approach to the choice of properties to transfer contrasts with approaches that use selection principles relying on, for example, context-insensitive notions of salience of properties (as in Ortony, 1979).

In the third strand – mapping discovery by analogy – Hobbs proposes something not much different in broad outline from other researchers (e.g. Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1983) who propose analogical structure-matching as the way to deal with (some) metaphor. Arguably the second strand, transfer by known mappings within inference, is the most interesting of the three. Hobbs provides as a prime example the use of spatial metaphor in Computer Science. One talks of a variable in a computer program as being "at", a number, say 100, as a way of saying that the variable's value is 100. Hobbs proposes therefore that a communicating agent that is familiar with this way of talking (or thinking) could have an inferential rule that can be glossed in English as

> *IF in some situation a variable's value is V THEN in that situation the variable is* [spatially-]at *V.*

Thus, this rule embodies a *known mapping link* between the source domain of space and the target domain of computer-science entities. The rule has the same status as any other inferential rule, and can be used at any convenient point during an overall process of inference-based understanding. It may look strange that the rule has the IF/THEN going from target to source rather than source to target. This is because rules are used abductively in Hobbs's approach: the variable's being spatially-at V leads to the abductive hypothesis that the variable's value is V.

The displayed rule acquires an indefinite amount of extra power in the following way. One talks of a variable "going" from say 100 to 200, as a way of stating a value change; of a variable being "between" two numbers; of a variable "keeping one step behind" another; and so forth: productively using an indefinite large part of the domain of space. Hobbs argues that such talk can be handled without the need to have separate mappings for "go," "between," and so on: rather, it suffices to use inferential connections within the source domain such as one between *going* and *being-spatially-at* and, thereby to be able to connect *going* to the mapping displayed above from being-spatially-at to having-as-value. Thus, a variable's "going" from 100 to 200 is ultimately interpreted as a change from a situation of having value 100 to a situation of having value 200.

Finally, Hobbs (1990) regards metaphor as crossing over between different domains, but fully accepts that domains have fuzzy boundaries and that the notion of domain is difficult. He therefore propounds that the exact scope of the notion of metaphor is theory-relative, in depending on decisions about what domains exist: there is no objective, theory-free fact of the matter about the boundaries of metaphor. In any case, Hobbs's actual computational approach does not impose or operationally rely upon any domain divisions at all. Therefore, mapping rules could in principle link concepts that, intuitively, are arbitrarily close.

Martin

An implemented computer program, called MIDAS, for metaphor interpretation and generation was produced by James Martin (1990, 2000). The acronym MIDAS stands for Metaphor Interpretation, Denotation, and Acquisition System. MIDAS was designed in part as a supplement to the Unix Consultant system, a computer program for automatically answering users' questions about the Unix operating system.

MIDAS incorporates knowledge of a set of a roughly Lakovian conceptual metaphors. The specific set included can be changed and is not itself the interesting side of the system. We will assume here, for expositional purposes, that MIDAS knows the conceptual metaphor USING A COMPUTER PROCESS IS BEING PHYSICALLY INSIDE A REGION. The system's knowledge base consists of a network of concepts. Among the concepts are the concept of using a computer process and the concept of being inside a region. These two concepts are linked by a "metaphor map." The metaphor map will be notated here in the following way, although the real structure is much more complex:

being-inside-a-region ↔ *using-a-computer-process*

Also, the two concepts have roles (or "slots") within them. Correspondingly there are two additional metaphor maps, this time crossing between roles:

the-enclosing-region ↔ *the-used-process*
the-thing-enclosed ↔ *the-process-user.*

As a result of knowing the conceptual metaphor, the system can easily understand a statement such as "I am in Emacs" to mean that the speaker is using Emacs, given that the word "in" accesses the being-inside-a-region concept. The literal interpretation that the speaker is physically in Emacs is rejected, because Emacs is not represented in the system as being a region. By contrast, the metaphorical interpretation is accepted because Emacs is represented as being a computer process. Importantly, though, the literal interpretation does not need to be rejected *before* the metaphorical one is accepted. The system tries to apply the possibly relevant conceptual metaphors it knows of, irrespective of whether the literal interpretation is acceptable.

MIDAS can also interpret metaphorical utterances that do not immediately fit its known mappings. The process of handling such utterances is handled by the MES (Metaphor Extension System) component of MIDAS. It uses two different techniques: *similarity-extension* and *core-extension*. Suppose the system knows that conversations are similar to computer processes, in the sense that they are both special cases of a more general concept of a process. Then the system can interpret the sentence *"I am in a conversation"* by using its known conceptual metaphor USING A COMPUTER PROCESS IS BEING PHYSICALLY INSIDE A REGION. Because of the known similarity between COMPUTER PROCESSES and CONVERSATIONS, the system has a mechanism for coming up with the new conceptual metaphor BEING ENGAGED IN A CONVERSATION IS BEING PHYSICALLY INSIDE A REGION.

This similarity-extension method is powerful, but core-extension is yet more so. The system can interpret the sentence "How do I get into Emacs?" just on the basis of knowing the conceptual metaphor USING A COMPUTER PROCESS IS BEING PHYSICALLY INSIDE A REGION and knowing some simple things about regions. The system is unable to find an acceptable interpretation using that conceptual metaphor directly. However, through knowing about a **result** relationship between the concept of *moving-into* (accessed by the phrase "get into") to the concept of being *physically-in*, and knowing that a usage of a process by a user is a *result* of the user invoking that process, the system can conjecture that the speaker is asking, in effect, "How do I invoke Emacs?" The system will create a new conceptual metaphor INVOKING A COMPUTER PROCESS IS PHYSICALLY MOVING INTO A REGION. The term "core-extension" is used because the concepts involved, such as moving-into and being-physically-in, must be "core-related." This somewhat complex notion covers only rather direct relationships such as the *result* relationship involved above.

Martin seeks to avoid having a literal-first account and thereby to obey the "total time constraint" (Gerrig, 1989) that conventional metaphors should take no longer to process than superficially similar literal language. MIDAS certainly avoids being literal-first in the sense that it avoids the need to *reject* literal interpretations before considering metaphorical ones. However, it does need to *construct* literal interpretations before considering metaphorical ones.

As Fass (1997, 316) points out, MIDAS is to be applauded for being able to prefer a metaphorical reading of "McEnroe killed Connors" to a literal reading, even though the latter is itself semantically acceptable. It turns out that the scoring mechanisms in the system, which knows that McEnroe and Connors are sportsmen, cause it to regard a SPORTIVE DEFEATING AS KILLING interpretation as more tightly fitting the sentence than a literal interpretation does, because sports-defeat requires its role-fillers to be competitors whereas killing has a much less specific requirement.

Martin does not make any use of the notion of a domain in his account of MIDAS, and there are no explicit domain divisions in MIDAS. Metaphor maps can in principle join arbitrarily close concepts, and what metaphor amounts to for the system is therefore entirely dependent on what maps happen to be included and how existing conceptual metaphors can be extended.

Fass

A second major implemented AI system for metaphor processing is that of Dan Fass (1997), indirectly related to the research of Wilks (1978). Fass's system is called meta5 (punningly, a step beyond meta*phor*). The system proceeds entirely by discovering analogies between source and target structures from scratch, with the process being guided by a relevance criterion explained below. It should be mentioned at once that the analogies discovered are of a very simple sort. However, the processing needed to discover them can be complex and subtle. Also, the system is unusual in measuring the degree of *dis*analogy between source and target structures, and using this measure in rating the aptness of the metaphor.

One standard-bearing example of meta5's processing is provided by

(4) My car drinks gasoline

taken from Wilks (1978). The system can interpret this as meaning "My car uses gasoline" essentially by finding an analogical match between the prior knowledge the system has that animals drink drinkable stuff and the prior knowledge that cars in general use gasoline. As a consequence, in constructing the internal meaning representation of the sentence, a *use* word-sense is employed as the right sense for the verb "drink" in the sentence.

In somewhat more detail, we can explain the process as follows, assuming the system only has one lexical sense for the verb "drink," namely the normal sense of an animal imbibing a liquid. We notate this sense here as *drink*. That the agent must be an animal and the patient must be a liquid is encoded as "preferences" (or "selection restrictions") in the permanent representation of the lexical sense in the system. The system finds, though, that the actual agent according to the sentence, the car, is not an animal. At some point the system will therefore look for a possible metaphorical way of interpreting the car-drink relationship in the sentence. It does this by seeing whether its knowledge about animals contains an item that is *relevant* to the sentence. The approach here is simple: from the sentence it takes only the *drink* word-sense, notes this sense's preference for an animal agent, and sees whether in the knowledge about animals there is information that they take part in a relationship that is either drinking or a word-sense-wise ancestor of drinking. Indeed, the system finds the knowledge item that *animals drink drinkable-stuff*. No other knowledge item for *animal* is relevant.

The system then looks for knowledge items within its prior knowledge of cars that match that animal knowledge item. It finds that the following matching item: *cars use gasoline*. It determines that there is a match because the *use* word sense and the *drink* word-sense are "sisters": they both have the same immediate parent sense, namely *expend*. Equally, the senses *drinkable-stuff* and *gasoline* are sisters, with *liquid* as parent. Such a pair of sister relationships between two knowledge items is necessary for them

P1: JZP
9780521841061c18 CUUS110/Gibbs 978 0 521 84106 1 April 5, 2008 9:43

to match. The system has now found a metaphorical relationship between "car" and "drinks" in the sentence, and can build a sentence meaning representation tantamount to "My car uses gasoline."

The system also looks at its non-relevant knowledge items of *animal* and *car*, in the above sense of relevance, and measures both how many other matching knowledge items there are and how many knowledge items for each of those two word-senses are not matched by a knowledge item for the other. The extra matches contribute to the strength of the metaphor, but the difference counts are inspired by the claim of Tourangeau and Sternberg (1982) that the greater the conceptual distance between source and target the more apt the metaphor. The counts can be used to choose between competing metaphorical interpretations, in other examples.

A point on which meta5 can be criticized, and is indeed criticized by Fass (1997, sect. 10.3.1.1), is that there is no coordination between a metaphorical relation found between the patient and verb ("car" and "drink") and a metaphorical or other relation found between verb and patient ("drink" and "gasoline"). Thus, the system does not look holistically at the sentence in determining the presence of analogies. This creates a problem with a sentence such as "My car drinks coffee," which Fass wishes his system to regard as anomalous and not metaphorical, and therefore not to settle on a metaphorical relation between the car and the drinking. Fass suggests a detailed way in which his problem could be fixed.

The incremental semantic construction approach in the (unfixed) system is in itself interesting because it means that the system does not first even *construct* a literal interpretation of the *whole sentence* before investigating metaphoricity, let alone *reject* a literal interpretation. But it is important to note that in the investigation by the system of a *part* of the sentence, such as "my car" together with "drinks," the system does adopt a fully literal-first approach: a metaphorical relation is only sought if an acceptable literal interpretation cannot be found for that part. Although it can be argued that this is a wrong approach even for sentence-parts, it does the service of showing us that the question of processing order in metaphorical sentence interpretation is much more complex than that of how literal and metaphorical interpretations of the *whole* sentence are ordered.

The system includes a complex numerical scoring mechanism to choose between competing interpretations of sentence parts as it goes along. This is largely based on lengths of paths in the semantic network. Fass (1997, sect. 10.2.2) has implemented a system extension in which the match scoring aspects of the system are enriched. The enrichment adds a diagnostic-salience measure on knowledge items that is dependent on how much inheritance was in involved in finding them: e.g. that a car has a definite physical boundary is inherited from further away in the semantic network than that a car has wheels, and is therefore less salient. Differences of salience could then be used to refine the comparative evaluation of discovered analogies.

On the other hand, there are major problems with the simplistic requirement that metaphorical analogies require sister relationships between cell components. For instance, it appears that the metaphorical interpretation above could not be found if, instead of *gasoline* being a direct descendant of *liquid*, there were a *liquid-fuel* sense interposed. However, given that the system already includes complex distance-based scoring, it would be straightforward to adjust the system to allow generalized cousin relationships rather than sister relationships, and to downplay or discount relationships that involved excessively long paths.

Finally, meta5 is interesting in being a fully implemented system that performs complex metonymic understanding as well as metaphorical understanding. It has knowledge of some conventional metonymic relationships such as ARTIST FOR ART PRODUCT and can therefore interpret sentences such as "John reads Shakespeare." Indeed, the system can handle arbitrarily long chains of metonymy. A limitation of the system is that metonymic interpretation is tried strictly before metaphorical, restricting the possibilities of interaction. The system can nevertheless obtain some forms of mixed metaphorical/metonymic interpretation.

There is no notion of domain in the design of the system, and word-senses are not sorted by domain. Indeed, as the sister relationship (above) is the core of analogy in meta5, metaphorical relationships can be between structures that are conceptually arbitrarily close up to sisterhood. Gasoline could have kerosene as a sister.

Finally, Iverson and Helmreich (1992) implemented a system, Metallel, that can be viewed as a substantially modified version of meta5, correcting some of its deficiencies. The system is ably summarized by Fass (1997, sect 10.1). Metallel views metonymy and metaphor as being on a par, rather than metonymy having precedence as in meta5. Once Metallel has found some potential available metonymic and metaphorical interpretations by a somewhat loose form of path search, it selects between them on the basis of a "grounding" process, which incorporates a type of analogical matching much like meta5's but that takes into account the whole sentence, not just parts of it in succession as meta5 does.

Barnden: ATT-Meta, Map-Transcendence and Pretence

The present author has implemented an approach, called ATT-Meta, for performing a type of reasoning that is arguably often necessary for metaphor interpretation. The approach is described in Barnden (1998, 2001a), Barnden, Glasbey, Lee, and Wallington (2004), Barnden, Helmreich, Iverson, and Stein (1994), Barnden and Lee (1999, 2001), and Lee and Barnden (2001a). The implemented ATT-Meta program is only a reasoning system and does not take linguistic strings as input, but, rather, logical forms derived from sentences by initial processing. For now the reader can take these logical forms to encode the literal meanings of the sentences, but we will refine this point below.

The metaphorical utterances of main interest in the ATT-Meta project are those that are conceptually related to known conceptual metaphors but that transcend them by involving source-domain elements not directly handled by the mappings in those metaphors. In ATT-Meta parlance these utterances are *map-transcending*. For instance, going back to the Hobbs examples, the sentence "N leaps from 1 to 100" is map-transcending for an understander if he/she/it only knows a *physically-leap* lexical sense for the verb "leap" but does not know a mapping for that sense into the target domain of variables and values, even though he/she/it does knows a mapping from, say, *spatially-at to have-as-value*. Similarly, if an understander knows a metaphorical mapping from *physically-in to using-a-process* (see Martin case) but has no mapping for *physically-enter*, then the sentence "How do I enter Emacs?" is map-transcending.

Clearly, map-transcendence is a fuzzy concept that is relative to particular understanders and particular conceptual metaphors the understander knows, and to our intuitive perceptions as to what is conceptually related to what (e.g. physically-leaping to being-spatially-at). Nevertheless, it is a useful intuitive characterization of a phenomenon that lies along a broad sector of the spectrum between conventional metaphor on the one hand and, on the other hand, entirely novel metaphor where no relevant mapping is known at all. Map-transcendence is strongly related to the phenomenon of unused parts of the source domain as discussed in Lakoff & Johnson (1980).

Very broadly speaking, ATT-Meta's approach is similar to Hobbs's second strand (Transfer by Known Mappings within Inference): ATT-Meta is based on rules encapsulating known metaphorical correspondences such as between *physically-at* and *has-as-value*, and on an integrated inferential framework which, in particular, allows arbitrarily rich source-domain reasoning to connect sentence components to source-domain concepts that can be mapped by known mappings. So, both systems can infer that a variable N *has value* 100 from any sentence couched in spatial terms that implies that N is *physically-at* 100, as long as the systems have the necessary knowledge about physical space to infer that N is physically-at 100 from the sentence. The inference can be arbitrarily indirect and complex in principle. To make the point, a vivid example would be a sentence such as "N started a circuitous route towards 100 but didn't complete the journey until after M fell to 0." This implies, among other things, that N (at some point) had value 100.

However, there is a fundamental difference of approach, as well as many technical differences of representation and reasoning, between ATT-Meta and Hobbs's scheme. The difference is that ATT-Meta avoids placing internal propositions such as *N is physically-at 100*, which are not statements about reality, on a par with statements such as *N has value 100*, which are. Hobbs's approach does maintain them on a par: there is nothing in his internal representation to say that the former proposition is merely a metaphorical pretence or fiction.

Instead, ATT-Meta creates a special computational "mental space" in which such propositions and inference arising from them are kept aside from propositions and reasoning about reality. We call this space a *metaphorical pretence cocoon*. Thus, the internal proposition *N physically-leaps from 1 to 100* arising directly from the sentence "N leaps from 1 to 100" is placed in the cocoon, and the inference result that (say) *N is spatially-at 100 afterwards*, together with the inference chain itself, lies within the cocoon. A metaphorical mapping rule that takes *spatially-at* to *has-as-value* can then give the result that, in reality, N has value 100 afterwards.

By clearly marking some propositions as being pretences, the use of a cocoon ensures that the system is not misled by the propositions directly derived from metaphorical utterances, that is, propositions like *N physically-leaps from 1 to 100*. Notice that in the case of "McEnroe killed Connors," the understander needs to be clear that the directly derived proposition *McEnroe biologically killed Connors* is not a statement about reality. But, in addition, if the understander knows that McEnroe definitely did not biologically kill Connors in reality, we do not want to let that information defeat the pretend information that McEnroe did biologically kill Connors. Thus, pretence cocoons prevent pretences from infecting reality but equally protect the integrity of pretences.

The use of cocoons has another benefit. Lee and Barnden (2001a) studied mixed metaphor of various types, and show how ATT-Meta deals with them. The main distinction studied was between serial mixing (commonly called chaining), where A is viewed as B and B is viewed as C, and parallel mixing, where A is used simultaneously as B and as C (see also Wilks, Barnden, & Wang, 1991). Serial mixing is viewed as having the B material in a cocoon directly embedded in the reality space, whereas the C material as in a cocoon embedded within the B cocoon. Thus, there is a pretence within a pretence. In parallel mixing, on the other hand, the B and C material is either combined in a single cocoon or is in two separate cocoons both directly embedded within the reality space. Thus, we have two pretences either side by side or blended with each other. There are unresolved issues about how to decide between these two possibilities, but in any case different dispositions of pretence cocoons allow important differences between types of mixing of metaphor to be reflected in the processing.

We have indicated that what is initially inserted in the pretence cocoon in the case of "N leaps from 1 to 100" is the proposition *N physically-leaps from 1 to 100*, and what is inserted in the case of "McEnroe killed Connors" is *McEnroe biologically killed Connors*. This reflects a general assumption in the ATT-Meta approach that what is inserted in the cocoon is a "direct" meaning of the metaphorical sentence (or of some metaphorical sentence-component such as a clause). A direct meaning is a logical form derived compositionally from the "direct" senses of lexical units in sentences. A direct sense is just any sense listed for the lexical unit in the understander's lexicon, so that it is directly accessible from the lexical unit. In particular, we have been assuming that the verbs "leap" and "kill" have as direct senses the concepts of *physically leap and biologically kill* respectively.

Clearly, a given lexical unit could actually have more than one direct sense, and indeed some of the direct senses could be metaphorical or special in some other way. We simply embrace such possibilities, saying that if, for instance, "leap" had something like *change-value* as a direct sense, then "N leaps from 1 to 100" could be understood without use of the inferential pretence mechanism outlined above, although in principle the mechanism could still be redundantly used as well. Equally, a direct sense may be figurative in some way but still lead to the construction of a proposition in the pretence cocoon. For instance, suppose the word "star" has *astronomical-star* and *prominent-movie-actor* as its only direct senses, and that we regard the latter as a figurative sense. Then "Mike is a star of the department" could be understood via the pretence mechanism using *Mike is a prominent movie actor in the department* in the cocoon. (Another option could be to use the *astronomical-star* sense.)

Thus, in the ATT-Meta approach, the pretence mechanism is potentially useful if direct meanings of sentence lead by within-pretence reasoning to within-pretence propositions that can be mapped by known mapping rules. It is irrelevant whether the direct meaning is dubbed as "literal" or not. We may or may not wish to regard *physically leap* as a literal sense of "leap" and *prominent-movie-actor* as a literal sense of "star", but such terminological decisions have no bearing in themselves on whether the pretence mechanism could be fruitful.

Another fundamental reason for not relying on a notion of literal meaning arises from serial mixing (A as B as C). In such a case, some of the phrasing in the utterance refers to the C domain, and this can cause material to arise in the B domain by C-to-B transfer. Therefore, B-to-A transfers may be working on non-literal material derived by transfer from C. For this reason alone, it is misguided to think of metaphorical mapping as a matter of transforming literal meanings. The consequences of this point have hardly been explored in metaphor research.

Insofar as direct meanings of sentences *can* often be regarded as literal meanings, ATT-Meta is in the class of systems that rely on constructing a literal meaning first (not necessarily from a whole sentence, but perhaps from a component such as a prepositional phrase or clause). Still, there is no reliance on *rejecting* that literal meaning before proceeding to metaphorical processing.

Before proceeding further in this description of ATT-Meta we also must explain that its reasoning is entirely query-directed. Query-directed reasoning – more usually called goal-directed reasoning – is a powerful technique much used in AI (see e.g. Russell & Norvig, 2002). In this form of reasoning, the process of reasoning starts with a query – an externally supplied or internally arising question as to whether something holds. Queries are compared to known propositions and/or used to generate further queries by some means. In a rule-based system as ATT-Meta, queries are compared to the result parts of rules, and then new queries arise from the condition parts. For example, in the case of a rule that says if someone is a student then he or she is presumably poor, a query as to whether John is poor would give rise to a subquery as to whether John is a student.

The system's metaphor-based reasoning is thoroughly integrated into a general-purpose rule-based framework for uncertain reasoning using qualitative uncertainty measures. ATT-Meta's reasoning both in source-domain terms and in target-domain terms is generally uncertain. Rules and propositions are annotated with qualitative certainty levels. There is a heuristic conflict-resolution mechanism that attempts to adjudicate between conflicting lines of reasoning, by considering their relative specificity.

We are now ready to look in more detail at an example. Consider:

"In the far reaches of her mind, Anne believed that Kyle was having an affair."

This is slightly adapted from a real-discourse example (Gross, 1994). We assume ATT-Meta is given knowledge of conceptual metaphors MIND AS PHYSICAL SPACE and IDEAS AS PHYSICAL OBJECTS. We also assume that "far reaches" only has a spatial sense for the system and that the notion is not mapped to the mental domain by any conceptual metaphor known to the system. The most important mapping known to ATT-Meta is the following, and is part of ATT-Meta's knowledge of IDEAS AS PHYSICAL OBJECTS:

> *degree of (in)ability of an agent's conscious self to operate physically on an idea that is a physical object*, in the pretence cocoon, *corresponds to degree of (in)ability of the agent to operate in a conscious mental way on the idea*, in the reality space.

A given metaphorical mapping link such as this is implicit in a set of *transfer rules* that we will not detail here.

In the example as we run it using ATT-Meta, the system is given an initial target-domain query (IQ) that is, roughly speaking, of the form *To what exact degree is Anne able to consciously operate mentally on the idea that Kyle had an affair?* In Barnden and Lee (2001) we justify this a reasonable query that could arise out of the surrounding context. The query is *reverse-transferred* from target terms to source terms via the above mapping to become a query of form *To what degree is Anne's conscious self able to operate physically on the idea?*

ATT-Meta can then reason that that degree of physical operability is very low, using the source-domain information gleaned from the mention of "far reaches" in the utterance and from commonsense knowledge about physical spaces and objects. Once this very low degree is established in the source domain, it is forward-transferred via the mapping to give a very low degree of conscious mental operability as the answer to the initial query (IQ). The program's reasoning for this example is treated in more detail in Barnden and Lee (2001). A variety of other examples are also computationally treated in that report and Barnden (2001c), Barnden et al. (2002), and Lee and Barnden (2001b).

We must note a largely unimplemented aspect of the ATT-Meta approach: "view-neutral mapping adjuncts" (VNMAs) (Barnden & Lee, 2001; Barnden et al., 2003). With partial inspiration from Carbonell (1982)'s AI work on metaphor, we view certain aspects of source domain information such as attitudes, value judgments, beliefs, functions, rates, gradedness, uncertainty and event structure to carry over to the target domain by default (the results can be overridden). For instance:

- We assume that the ordering of events and their qualitative rates and durations carry over by default, whatever the nature of the particular metaphorical mapping being used, thus avoiding the need for individual mapping rules to deal with them.
- If an agent A in the pretence has an attitude X (mental or emotional) to a proposition P, and A and P correspond, respectively, to an agent B and a proposition Q in reality, then B has attitude X to Q.
- As for gradedness, if a property P in a pretence corresponds to a property Q in reality, then a degree of holding of P should map to the same degree of holding of Q (unless there is additional evidence about Q).

We have produced an experimental implementation that handles rates and durations as VNMAs, but much work remains to be done on other VNMAs. In particular, gradedness is currently handled directly in individual rules – notice the degrees in the metaphorical correspondence used above. In place of this handling, we would like to have instead simpler mapping rules that do not mention degree, relying on a separate, general mechanism for the degree transfer.

Finally, the ATT-Meta approach does not rely on domain distinctions, even theoretically, let alone enshrine them in some way in the implemented system. Although in

P1: JZP
9780521841061c18 CUUS110/Gibbs 978 0 521 84106 1 April 5, 2008 9:43

this article we generally adopt the common practice of saying that metaphor transfers information from a source domain to a target domain, the ATT-Meta approach has a different stance: metaphor is a matter of transferring from a *pretence* to *reality* (or to a surrounding pretence, in the case of serial mixing). Notice that in the mapping rule set out above, reference is made to pretence and reality, not to domains. It does not matter what domains the information used in the pretence comes from, and this means that it does not matter how we may intuitively circumscribe the source and target domains in the metaphor. In particular, it does not matter how close, difficult to distinguish, or overlapping those domains are. In practice, it will often be the case that we can theoretically identify a source domain in which the direct meaning of he sentence lies, and that inferences from this meaning also lie within that domain. However, this has no bearing on the course of processing, and the reasoning within the pretence is not limited by any consideration of domains.

Narayanan

Srini Narayanan has implemented a metaphor-understanding system (Narayanan, 1997, 1999) that has mostly been applied to interpreting metaphorical statements about economic policy, where the source domain is that of everyday physical movement activities such as walking, as in the headline "Liberalization plan stumbling." However, it would appear reasonably straightforward to apply a modified version of the system to other source and target domains, and Narayanan (1999) mentions using a health-based source domain.

The system has been applied to many utterances about economics from newspaper articles, and has powerful facilities for addressing subtle aspects of such utterances. However, much as in the case of ATT-Meta, the system does not take sentences as such as input, but rather simple feature-value representations that could result from initial processing of sentences or other discourse fragments. The system is based on knowing a set of conceptual metaphor maps such as ACTING IS MOVING, OBSTACLES ARE DIFFICULTIES, and FAILING IS FALLING.

Examples of fragments successfully handled include "Liberalization plan stumbling," "European Giant falls sick," "taking a cautious step in the right direction" and "Economic reform is like crossing a river by feeling for the stones." Narayanan is especially concerned to deal with aspect, i.e. the internal temporal structure of events. The system can deal with, for instance, the intermittent nature of an action such as rubbing, the aspect conveyed by the perfect tense, and aspect conveyed in phrases such as "start to pull out," "on the verge of" and "back on track."

Both the source domain and the target domain are represented as fixed network structures, of rather different types. The target domain representation is a "belief network" (Pearl, 1986), in which nodes stand for economic variables needed for depicting the economic situations of interest. The variables include economic actors (example value: Indian Government), economic policy (example value: capitalism), status of a policy, gross domestic product, geographical location, rate of progress, level of difficulty (e.g. of implementing a policy), and goals of actors. Each node is repeated across a small sequence of time slices (up to four), so that for instance there is a policy node for time 1, a policy node for time 2, and so on. Nodes are linked together to represent probabilistic relationships between variables. For instance, the links allow the conditional probability of policy being such-and-such at time 2 given that it is so-and-so at time 1 and a policy failure happens at time 1. When the belief network is used for inference, particular probability values at nodes are fixed on the basis of input and metaphorical transfer, and then the links cause posterior probabilities for particular variable values at nodes to be calculated. In this way, the network can probabilistically model a complex unfolding economic situation.

The source-domain representation is, roughly speaking, a type of marker passing network in which (the main type of) nodes represent states that can occur in activities such as walking, falling and getting up. Links between these nodes show how states can (stochastically) be caused by predecessor states, and markers passing along these links simulate the progress of activities.

The state nodes in the source domain include a subset that serve as the inputs to the system's metaphorical maps. For instance, the DIFFICULTIES ARE OBSTACLES map responds to the presence of a marker in the *bump* node in the source-domain network and contributes to the setting of the probability level at the *difficulty* node in the target network. One type of map, "parameter" maps, handles gradedness. For instance, velocity in the source domain is mapped to rate of progress of a policy in the target domain, or distance travelled in walking to degree of completion of an economic plan.

The processing within the source-domain network allows rich examples of map-transcendence to be handled. For instance, consider any discourse fragment that mentions an economic policy approaching a cliff edge. Recall that *falling* maps over to *failing*. Provided that the source-domain network has the right structures to predict falling from walking to the cliff's edge, the system can infer the target domain conclusion that the economic policy will fail.

Clearly, the system makes strong use of source-domain inference, if we regard the mental simulation of activities within the source domain as inference. Furthermore, it is uncertain inference, because of the stochastic nature of marker passing between state nodes. It is clear also from the above that the system places great weight on gradedness.

As for the role of literal meaning, consider the sentence "Economic reform is like crossing a river by feeling for the stones." This will be input to the sentence in the form of a setting of the source-domain network that depicts a fictional entity, corresponding to economic reform, crossing a river, and so on. In this sense, the system constructs a whole literal interpretation first. However, the system does not itself evaluate whether economic reform can itself cross a river, so, as with Hobbs's approach, MIDAS and ATT-Meta, there is no sense in which the system itself *rejects* a literal meaning before computing a metaphorical one.

The system is, clearly, strongly founded on domain distinctions, which are explicit in the structure of the system. Given the intuitive, qualitative distance between economics and bodily movement, this might not, superficially, appear to be a problem. However, various types of extension or enrichment of the system could soon run into problems. For one thing, mental processes are important both for physical activities in the world (e.g. reasoning about what to do at a crossroads) and in the economic domain, and this is already weakly evident in Narayanan's work. A more detailed treatment of mental processing in the two domains would require separate and differently organized network structures to handle mental states, whereas intuitively the two domains simply overlap on the matter of mental processes, which themselves could just as much be viewed as forming a domain.

Veale: The Sapper System

Tony Veale (Veale, 1998; Veale & Keene, 1997) has constructed Sapper, an implemented hybrid sym-bolic/connectionist model for finding structural analogies. It is based on a semantic network framework in which nodes stand for concepts and between which activation values can flow. The work on Sapper appears to be largely separate from Veale's work on a "conceptual scaffolding" theory of metaphor (Veale & Keene, 1992).

Sapper does not take linguistic input as such, but rather attempts to find a metaphorical mapping between any two concepts S and T in its network that are from different domains, for instance *composer* and *[military] general*. In this example, the system comes up with a rich metaphorical mapping, involving component correspondences such as *orchestra* corresponds to *army*, *musician* to *soldier* and *musical-instrument* to;

musket. In this way it is similar in orientation to analogy-finding systems in Cognitive Psychology, such as SME and ACME. Indeed, Veale has shown in much detail, both theoretical and experimental, that his system can find analogies similar to those found by SME and ACME, while performing less processing.

Sapper has a long-term "bridge"-forming aspect and a short-term structure-matching aspect. The former is done in advance of any analogy-finding, and finds potential analogical correspondences between concepts. It does so by means of purely symbolic processing over the semantic network, based on certain simple heuristics (a "Triangulation" rule and a "Squaring" rule). Such a potential correspondence is called a bridge and is implemented as a special link between the nodes.

Analogy-finding per se in a particular case, such as for *composer* and *general*, consists of the short-term structure-matching aspect. This aspect exploits the long-term bridges via activation-spread in a way to be described shortly, and thereby constructs overall, coherent mappings containing component correspondences such as *orchestra* to *army* in the example above.

Structure-matching works in outline as follows, given two nodes S and T, thought of as the source and target nodes respectively. Activation is sent out from S and T, to a prespecified distance ("horizon") in the network. If the two waves of activation meet at a bridge between two nodes S and T, respectively, then the system sees if there is a chain of links from S to S that is isomorphic to a chain of links from T to T. That is, the two chains consist of links of the same types in the same directions. Then for each pair of corresponding nodes on the chains the system considers them to be mapped to each other, and takes the overall mapping thus defined by the chains to be a partial interpretation of the T-is-S metaphor. Now the system takes the "richest" partial interpretation found by this method, and considers the remaining ones in descending order of richness, attempting to combine them consistently with the richest one. The final result is Sapper's overall metaphorical interpretation of T-is-S.

The theory behind Sapper places important, explicit weight on domains, and Domain distinctions are used in the structure-matching process. A domain in Sapper is relative to a given "root" node. The domain for the node is the region of the semantic network that is reachable from the node via network links in a particular way. However, Veale does not appear to address the difficulties arise with source and target domains that intuitively overlap, which would require that activation flow during structure-matching not be domain-confined as he assumes it to be. For instance, drums are used in bands in armies, not just in ordinary orchestras.

It appears that the processing in Sapper is entirely symmetrical between source and target, so that for instance "a composer is a general" creates the same the same metaphorical correspondences as "a general is a composer." This may look as though it goes against claims in the metaphor literature (e.g. Ortony, 1979, 197) about the asymmetry of metaphor. However, it is not difficult to bias the processing in Sapper in ways that would asymmetrically affect the activation flow and thus ensure asymmetrical results. Also, Barnden (2001d) argues that asymmetry is a more subtle and delicate matter than it is usually portrayed as being; for example, the true asymmetry between S-is-T and T-is-S can reside in which particular mapping links are used in interacting with the overall discourse rather than with whether the links themselves differ between S-is-T and T-is-S. Indeed, on his website Veale describes how Sapper does structural transfer, in a way roughly similar to other analogy systems. Structure on the source side that is not paralleled on the target side can be transferred as "candidate inferences" to the target side. Structural transfer from source to target involves different pieces of domain information from those involved in transfer from target to source, even when the same metaphorical linkages are involved.

Sapper could be said to perform source-domain inference in using activation flow within that domain. Activation levels represent gradedness, for instance the degree to which the property denoted by the node holds. The levels therefore do not represent degrees of certainty, as they do in many connectionist systems.

Discussion: Contributions of AI to Metaphor Research

Here we examine some specific issues on which AI is being helpful to metaphor research. We will draw heavily on the preceding review of particular AI approaches, but will also make additional observations.

Mundaneness

Non-AI research such as that of Lakoff (Lakoff, 1993; Lakoff & Johnson, 1980) and of many researchers in Corpus Linguistics and Applied Linguistics has shown us that metaphor is an aspect of ordinary, everyday language, not just of literary or other heightened forms. AI is in a peculiar position to add both to the appreciation of the variety and complexity of metaphor as it arises in practical discourse and to the question of how to process real metaphor in practical contexts, because of the inclusion within AI of applications-oriented research. One of the AI systems reviewed above (MIDAS, by James Martin) concentrated on metaphor arising in question-and-answer sessions between users and an automated Unix help system. Narayanan's research used the domain of economics as an application area. A research project led by the present author, not reviewed above but drawing upon the ATT-Meta research, is looking at the metaphorical expression of affect (emotion, value judgments, etc.) in the context of an e-drama system that supports virtual dramatic improvisation by users sitting at computer terminals (Zhang, Barnden, & Hendley, 2005). Improvisations can be on any topic, but the system has in particular been used for improvisations concerning school bullying and embarrassing illnesses.

Non-Assertional Metaphor

One consequence of looking at applications is as follows. In describing MIDAS we cited a metaphorical question as an example – "How do I get into Emacs?" It is remarkable, though not generally remarked upon, that the vast bulk of writing on metaphor has concentrated on assertions. Yet, metaphor is just as appropriate in questions, commands, and so on, as it is in assertions, and often occurs in non-assertions in real discourse. Non-assertional metaphor raises special issues. Questions and commands are not centrally about conveying new information about the target or making the understander appreciate the target in a special way, yet existing theorizing on the meanings or connotations of metaphorical utterances presupposes that some new information or special view of the target is being communicated. In particular, whereas with an assertional metaphorical utterance an incompatibility between one potential interpretation and the target domain may indicate that the interpretation is incorrect, in the case of a metaphorical question the incompatibility may mean simply that a negative response is needed or the speaker has an incorrect supposition about the target domain, so that an answer could be directed at countering this. It could turn out that particular existing theories based on assertion could be smoothly generalized to deal with non-assertional metaphor, but the issue needs at least to be explicitly addressed.

Details of Mappings

Much work on metaphor outside AI has specified particular mappings between sources and targets. The mappings are often backed up by discursive accounts of how they could help in the understanding of particular example utterances or types of utterance. However, without their being embedded in a detailed computational system it is difficult to determine whether, on the one hand, the mappings really do achieve all the effects they are credited with, and whether, on the other hand, they successfully avoid interacting to produce unwanted side-effects. In other words, mappings proposed in non-AI literature on metaphor are typically only vaguely evaluated as to coverage, coherence and effectiveness. In contrast, systems such as MIDAS and ATT-Meta provide a framework within which to do extensive experimentation with alternative sets of mappings.

Source-Domain Reasoning and Pretence Reasoning

Several of the reviewed AI systems (by Hobbs, Martin, Narayanan) make crucial use of online source-domain inference: inference that is in terms of the source-domain subject matter and that is made at the time of trying to understand a metaphorical utterance. Source-domain reasoning was also briefly advocated in the seminal work of Carbonell (1982) on metaphor in AI. The ATT-Meta system is centred on the closely related notion of within-pretence reasoning.

Now, source-domain inference has arisen quite frequently in the non-AI literature. For example, comments in Lakoff (1993) and Lakoff and Turner (1989, 62, 64, 94) suggest the use of source-domain inference. The discussion of metaphorical inference patterns in Turner (1987) appears to allow for online source-domain inference. The work of Ruiz de Mendoza Ibáñez (1999) on interactions between metonymy and metaphor includes mention of metonymy occurring within the source domain of a metaphor, and this amounts to a type of online source-domain inference. As for online within-pretence reasoning, Levin's (1988) work on metaphor in literature implies the use of it, and van Dijk (1980) provides a tentative account of metaphor in terms of counterfactuals. The "blending" ("conceptual integration" approach) in Cognitive Linguistics (Fauconnier & Turner 1998), when applied to metaphor, makes inference within the blend-space central. A blend-space is similar to a pretence cocoon in ATT-Meta, though the latter concept is more computationally specific while being unconstrained by notions of domain.

But the study of source-domain reasoning and within-pretence reasoning in AI research on metaphor has givenflesh to and clarified the somewhat schematic and limited discussion of it in the non-AI literature. What AI can distinctively contribute is detailed, effective mechanisms for performing it. Complex technical matters of representation, reasoning and evidence-comparison are involved here, especially when uncertainty and gradedness are brought in.

The reason for the intense attention to source-domain and within-pretence reasoning in AI may be that, in concentrating on real examples of metaphor in mundane contexts, the researchers concerned have been affected by the fact that truly novel metaphor is far from being predominant in metaphor in real discourse, and have concentrated on the rich, open-ended exploitation of already-known mappings. Source-domain or within-pretence inference enables the map-transcending aspects of the utterance – the aspects not directly handled by known mappings – to be linked to the aspects that are so handled. Map-transcendence is a central problem of metaphor that has not been adequately treated, although the Hobbs, Martin, Narayanan, and Barnden approaches are important developments.

Economizing on Parallelism, and Use of Disanalogies

Hobbs, Narayanan, and Barnden all recognize that much or all of what one needs to get out of a map-transcending metaphorical utterance can often or perhaps usually be got *without finding target-domain correspondents for the map-transcending items*. This stance is against the idea that the fundamental task in metaphor understanding is to establish new mappings, indeed, to establish as much parallelism as possible between the two domains. Rather, the three approaches seek to exploit as far as possible the already known mappings. In particular, Barnden, Helmreich, Iverson, and

Stein (1996) explicitly championed the thesis that it is often misguided to think that map-transcending source-domain elements should be expected to have a parallel in the target, let alone to think that it is profitable to look for it. For example, it seems excessive to expect the "dim recesses" mentioned in "The idea was in the dim recesses of Tony's mind" to actually correspond to any identifiable components of the mind in reality, rather than serving *merely to connote physical inaccessibility within the metaphorical pretence*. On the other hand, there are certainly situations where one needs to find some target-domain correspondents. The question of which these situations are is an outstanding research issue, on which a start is made in Barnden and Lee (2001).

Relatedly, the benefits of attending to disanalogies between source and target in metaphor deserve more study. Fass's system (meta5) is unusual, and unique among the systems reviewed, in regarding disanalogies between source and target as a source of useful information.

Dissolving Metaphorical Transfers into the Overall Processing

The Hobbs and Barnden approaches achieve great flexibility in allowing target-domain (or within-reality) reasoning steps, source-domain (or within-pretence) reasoning steps, and metaphorical transfer steps to be arbitrarily mixed together in a completely uniform and task-dependent way. This flexibility is a contribution to conceptions of how the different types of processing in metaphor can fit together. Most discussions of metaphor appear to assume that transfer steps occur in some special phase of processing.

The flexibility of mixing is aided by casting mappings as inference rules that are applied just in the same way as other rules. Usually in metaphor research, whether in AI or elsewhere, mappings are a different sort of entity, which inhibits even the realization that a uniform treatment would be liberating and beneficial.

Context and Extent

It is often pointed out that the information conveyed by a metaphorical utterance can be highly sensitive to context, and a considerable amount of psychological experimentation and philosophical theorizing has addressed this (e.g. Giora, 1997; Leezenberg, 1995; Stern, 2000). Context is important for the understanding of much non-metaphorical language as well, but metaphor heightens its effect.

The sentence "Mike is a rock" is highly indeterminate as to what it might convey, absent any specific context. Perhaps the speaker is intending to convey that Mike can be relied upon. However, in "Mike's friends are very upset by criticism, but he's a rock" the contribution of "rock" is much more definite. It is probably not getting at Mike's reliability: the sentence is arguably saying that Mike is highly tolerant of criticism, and if so it is presumably exploiting a correspondence between invulnerability of rocks to physical assault and tolerance of criticism by people.

In this example the disambiguating context about Mike's friends and criticism is near to the metaphorical clause, but in other cases the necessary contextual information might arise from further afield, and might have to be derived from the surrounding passage or other information by subtle or knowledge-intensive processes of inference. Thus, a full approach to metaphor must deal with possibly complex, extensive passages of discourse, and complex inference.

Although AI work on metaphor has yet to address context fully, some of the systems reviewed above give context a crucial guiding role and are at least in a position to accommodate its effects smoothly. Hobbs and Barnden place much weight on reasoning goals derived from context as a crucial driver of what metaphorical interpretations are drawn, and their approaches are unusual amongst detailed metaphor-processing schemes in this respect. Contextual-goal drivenness is a powerful tool not only against the often-noted indeterminacy of metaphorical meaning (see e.g. Stern, 2000) but also against the problem of inappropriate or irrelevant aspects of the source domain getting in the way (such as the shape of a pig's tail when classifying a person as a pig). In a contextual-goal driven approach, those irrelevant aspects will simply tend not to be queried.

As we have made clear, many authors outside AI have discussed the importance of context. What AI can contribute is detailed, computationally tested mechanisms by which it can be brought to bear.

Uncertainty

The information gained from metaphor is generally uncertain. The indeterminateness of the import of "Mike is a rock" without a sufficiently specific context is itself a type of uncertainty. But even with the context shown above, we cannot be *certain* that Mike is tolerant of criticism (according to the speaker). Perhaps, after all, the speaker is intending to convey that Mike can be relied upon to give support to his colleagues when they are upset by criticism.

But even if an interpretation in terms of Mike's tolerance to criticism is correct, we cannot be certain about the degree of tolerance: perhaps the speaker is merely trying to say that Mike has a normal level of tolerance, in contrast his colleagues' marked lack of tolerance. After all, different types of rock have different degrees of vulnerability to physical assault, and, without further information, it can merely be a presumption that a rock has a high degree of invulnerability.

Therefore as well as the uncertainty arising between there being qualitatively different possible interpretations (e.g. one appealing to reliability and one appealing to tolerance), there is also uncertainty arising from within the source domain itself. Another example of the latter phenomenon would arise from talking about someone "burying" an idea in his mind. In the physical world, once something is buried it (at best) only *normally* stays buried. There can therefore be no *certainty* that the idea will not "pop up" again.

Most work on metaphor outside AI sidesteps detailed considerations of uncertainty, although systems such as SME and ACME, where there are scoring mechanisms, do provide some support for a restricted type of uncertainty handling. Amongst our reviewed AI systems, Narayanan, Hobbs, and Barnden all allow the system's source-domain reasoning and target-domain reasoning to be uncertain. Uncertainty is important for making the overall processing do justice to people's use of metaphor, but unfortunately greatly complicates the technical nature of the computational framework.

Source/Target Overrides

The uncertainty issue also reveals the importance of often allowing information transferred from source to target *to override information about the target*. This possibility is under-studied in metaphor research, because usually the information about a target domain is cast simplistically in the form of certainties which cannot be overridden. This practice has led to researchers, outside AI and within, almost exclusively concentrating on the fact that target-domain information must sometimes override what comes from the source. Of course, this is indeed appropriate in many cases: since it is certain that France and Germany are not cognitive agents and are therefore incapable of love, metaphorically casting the relationship of those countries within the EU as a "marriage" (Musolff, 2004) should not lead to the result that they love each other in reality.

But, if a piece of target-domain knowledge is not certain, but let us say merely a default, there is no reason in principle why the information should not be overridden by transfers form the source. Thus, "SnakeByte Technologies nursed its competitor RabbitWare Inc. back to health" would override the default that competing companies do not normally deliberately help each other. The utterance "In the far reaches of her mind, Anne believed that Kyle had been unfaithful" defeats the normal presumption that people's thoughts about their spouses' possible affairs are central and conscious ones. It may even be that one important function of metaphor is to convey situations that are exceptions to target-domain defaults. The exception-expressing function of metaphor may be especially significant given that exceptional situations are less likely to be easily expressible using the resources native to the target domain.

It appears that only in the context of the ATT-Meta system has the process of source-over-target overriding been studied in computational detail, though see Indurkhya (1992, 85–86) for other comments on the importance of such overriding. In ATT-Meta, both directions of override are possible, depending on the fine detail of the reasoning lines involved in particular cases.

Gradedness

The rock example above brings out the importance of matters of gradedness (degree) in metaphor. It is gradations, not black-and-white propositions, that metaphor is often getting at, a point that deserves greater emphasis in metaphor research. The interpretation suggested for "Mike's friends get very hurt by criticism, but he's a rock" was not the bald proposition that Mike is tolerant of criticism but that he is *highly* so. Equally, the sentence "The memory was hidden far back in the labyrinth of John's memory" plausibly does not convey that the memory was completely inaccessible to John but rather that it was *highly* inaccessible, or very difficult to access. A range of specific examples of gradedness in metaphor interpretation can be found in Barnden (2001b, 2001c).

Once gradedness and uncertainty are considered it also becomes evident that a metaphorical utterance may not necessarily introduce totally new information but may rather change the degree of holding, and/or the certainty, of some existing piece of information. Gibbs and Tendahl (2006) discuss this under the heading of the "strengthening" of (and the opposite: contradiction of) existing assumptions, in the light of considerations of metaphor in Relevance Theory (Carston, 2002; Sperber & Wilson, 1995). In the rock example, other evidence may already have established that Mike *may* be *somewhat* insensitive to criticism, so the sentence is both strengthening the *may* to *presumably* and strengthening the *somewhat* to *highly*. Note also that such strengthening goes beyond the notion that metaphor can draw attention to or increase the salience of (Ortony, 1979) pieces of information about the target domain. We are talking instead about adjusting pieces of information about the target domain.

It cannot be claimed that AI or any other field has developed generally accepted, comprehensive methods for handling gradedness. Nevertheless, Narayanan and Barnden place weight on the handling of gradedness and the transfer of graded information from source to target. Perhaps as important than the actual handling of gradedness in some recent AI metaphor systems is the sheer fact that the pressure in AI towards considering the details of processing practical examples in realistic contexts makes one more readily appreciate the central role that gradedness plays in metaphor (going beyond the obvious role of gradedness in scale-based conceptual metaphors such as MORE IS UP).

Domain Distinctions

Metaphor is frequently characterized as a matter of mappings or transfers between different "domains," often to make a contrast with metonymy, which is often claimed to operate within a single domain. On the other hand, some authors have questioned the usefulness of the domain notion or the degree of distinctness that is required between the two domains in a metaphor (see e.g. Dirven & Pörings, 2002; Kittay, 1989). For simplicity of discussion we have mostly used the notion of domain uncritically in this article. It is certainly true that in much metaphor there is an intuitive sense in which the source and target are qualitatively very different. The question is whether real sense can be made

P1: JZP
9780521841061c18 CUUS110/Gibbs 978 0 521 84106 1 April 5, 2008 9:43

of this and whether it matters to metaphor processing anyway.

The present author has found in his own AI work on the ATT-Meta approach that the detail and clarity required for well-founded computational implementation to be a major factor in his coming to doubt the usefulness of the concept of "domain" in studying metaphor (and metonymy). In trying to make decisions about what domains particular pieces of knowledge should be assigned to he came to realize what a hopeless and arbitrary task it was. The resulting despair was relieved by an ultimate realization that having domain distinctions was not operationally useful in any case.

The nature of the other systems in the review above also throws doubt on the usefulness of the notion. Only Veale and Narayanan actually have domains affect how their systems are structured and how the processing works. Hobbs does believe that metaphor is a matter of mapping between qualitatively disparate domains, but this stance has no operational effect in his system. In contrast, Barnden regards this disparateness as merely being a common case and is happy for the two sides of a metaphor to be arbitrarily close in their qualitative nature. Metaphors such as "Thatcher was Britain's Reagan" are common, and have source and target domains that are broadly similar in subject matter. For an example with even less qualitative distance between the two sides, one's neighbour's teenage children can act as a metaphor for one's own: if one has a daughter Jenny and the neighbours have a son Jonathan who behaves similarly to Jenny, then one could say "Jenny is our family's Jonathan." Of course, it is open to someone to say that the Jenny family is qualitatively different from the Jonathan family, and that they are therefore different domains, but this is post hoc rationalization with no operational significance.

Despite the closeness between target and source in the Jenny/Jonathan example, the metaphorical utterance appears quite apt to the present author. If this impression is shared with others, it may appear to conflict with the evidence adduced by Tourangeau and Sternberg (1982) that the greater the conceptual distance between source and target the more apt the metaphor. However, note that the linguistic form of the metaphorical utterance and the presence of context are important factors. A bald statement that "Jenny is Jonathan" without much context might well not come over as apt.

Apart from considerations of overall qualitative closeness, there is often a considerable amount of overlap between the intuitive source and target domains in metaphor even when they otherwise differ a great deal. We noted some overlap between the economics (target) and health (source) domain in the Narayanan discussion – and we could also have pointed out that health services are part of the economy – and between the orchestra and army domains in the Veale discussion. With reference to the Fass discussion, the domain of cars involves the domain of animals because cars can carry people and other animals.

It is quite possible to maintain a fiction that domains do real work in metaphor as long as one only deals schematically with some isolated examples, and does not try to come up with a unified and processually detailed approach to metaphor that can work on a wide variety of metaphors on the basis of the same overall knowledge base.

Relationship to Metonymy

The relationship of metaphor to metonymy is highly contentious and complex (Dirven & Pörings, 2002; Fass, 1997). It has proved difficult to distinguish clearly between the two phenomena, and they may be at ends of a spectrum within which many compromises are possible. Particular discourse examples are often hard to classify as to whether they exhibit metonymy or metaphor. Also, metaphor and metonymy often co-occur in richly interactive ways in discourse. However, there has been little work on processing accounts that handle both phenomena. As it happens, two of the AI approaches reviewed above – those of Hobbs and Fass – pay much attention to metonymy as well as to metaphor, and allow certain types of interaction. They complement work such as that of Ruiz de Mendoza Ibáñez (1999) and Goossens (1990) outside AI. Hobbs's approach is perhaps especially noteworthy in that, as in the case of metaphor, it embeds metonymy as just one type of inference within the system's inferencing as a whole (Hobbs et al., 1993). Therefore, in principle, arbitrarily complex and diverse mixes of metaphor and metonymy should be able to be handled, and it is likely that compromises between metaphor and metonymy are possible.

If domains are abandoned as a well-founded underpinning for metaphor, then metaphor cannot be distinguished from metonymy on the usual ground of between-domain moves versus within-domain moves. Thus, any profound effect that metaphor research in AI and other disciplines may ultimately have on the fate of domains must be matched on by profound effect on the metaphor/metonymy relationship.

The Literal: Its Nature and Use

Strongly related to the domains issue is a theme that appears throughout the field of metaphor, and continues to be a matter of debate in the field (Gibbs & Tendahl, in press): the role, if any, of the literal meaning of metaphorical utterances or words in them in deriving their metaphorical meaning.

Of the systems reviewed, only Fass's (meta5) has any use for the idea of having to *reject* a literal interpretation before considering a metaphorical one, and even in his case the incremental semantic processing (while problematic in itself) means that the rejection is by sentence-part rather than by whole sentence. See Lytinen, Burridge, and Kirtner (1992) for another system with a related incremental quality.

The approaches of Hobbs, Martin, and Narayanan do rely on *constructing* a literal interpretation of a metaphorical sentence, or sentence-like subunit such as a clause. Barnden's approach is similar in this respect, though there it is a "direct" meaning that is constructed, with the question of whether it is necessarily to be called the literal interpretation being left as a terminological side issue. It should not be feared that there is necessarily any conflict between these approaches and psychological experimental results about metaphor processing being about as fast as, or sometimes faster than, literal-language processing under certain conditions (see e.g. Gibbs & Tendahl, in press, for a discussion of such results). This speediness does not of itself show that literal meanings are not being computed. The evidence on these matters from psychological experiment is mixed, because it is bound up with the nature of the context of the metaphorical utterance and the novelty or otherwise of its metaphorical elements: context could by itself suggest part or all of the meaning, and a piece of familiar metaphorical terminology could have its target-domain meaning listed in a lexicon. Also, the type of literal (or direct) meaning that is constructed in the aid of metaphor understanding is plausibly less fully fledged than that needed in cases where the linguistic string really should be interpreted literally. In the latter case, the literal meaning itself needs to involve integration with the context, whereas in the metaphor case it is instead the metaphorical meaning that needs to be fully integrated with context. It is possible all that the metaphorical processing is adding is the occasional hop from a complex source-domain (or pretence) scenario into a target-domain (or reality) scenario, and the time for such hops could be swamped by the time needed for all the other processing going on, such as anaphor resolution and semantic/pragmatic inferencing of many other types. AI can contribute here in clarifying the overall computations needed and how they can be imaginatively structured and optimized.

Finally, note that serial mixing (chaining) of metaphor complicates the role of literal meaning in metaphor, as noted in the discussion of ATT-Meta. What is transferred online in metaphor can already be a product of online metaphorical transfer.

P1: JZP
9780521841061c18 CUUS110/Gibbs 978 0 521 84106 1 April 5, 2008 9:43

Transfer of Attitudes and Value Judgments

A metaphorical utterance often conveys or instigates a mental or emotional attitude or a value judgment about the target subject matter. This is perhaps especially prevalent in metaphor used in political discourse (see e.g. Musolff, 2004). The attitude or judgment can be on the part of some person mentioned in the discourse, or it can be on the part of the speaker/hearer. For instance, talking about somebody's mind as if it were a "cess-pit" may be intended to make the hearer have an emotional revulsion to, or negative value judgment of, the ideas of that person. On the other hand, saying that "The problem crushed Mike into the ground" primarily conveys something about Mike's emotions, although of course it can also engender the meta-emotion of sorrow over Mike's fate.

Although attitudes such as emotions and value judgments are of widely recognized importance for metaphor, it is important to have detailed accounts of how exactly they may be processed in metaphor understanding. The processing of attitudes interacts heavily with ordinary inferencing, rather than being an isolatable matter. In addition, emotions and value judgments are intrinsically graded, so the theme in this subsection interacts strongly with the general gradeness issue we identified above.

The description of the ATT-Meta project mentioned that mechanisms are being developed in that project for transferring attitudes and value judgments from source to target by default, whatever the particular conceptual metaphor involved, obviating the need for special mechanisms per conceptual metaphor.

Connections to Reasoning about Beliefs

Little research into metaphor has taken into account the fact that if a hearer wishes to understand what a speaker means by a particular metaphorical utterance, it is the speaker's beliefs about the target and source domains that are important, not the hearer's. In effect, the metaphorical processing should occur within the speaker's "belief space" (as perceived by the hearer). Relatedly, metaphor can occur within the complement clauses of mental state verbs, as in "Mary believes that SnakeByte nursed RabbitWare back to health." One interpretation of such a sentence is that the metaphorical conception of the target is Mary's own (or rather, Mary's own, as viewed by the speaker), not (directly) the speaker's. In this case, metaphorical processing should be embedded within a belief space for Mary (within a belief space of the speaker). Stern (2000) and van Dijk (1980) are rare in metaphor research to have addressed those issues, albeit only in an abstract way.

The issue is important in the ATT-Meta project. As well as handling metaphor, the ATT-Meta system can perform reasoning about agents' beliefs and reasoning. Methods are being developed for processing metaphor within the context of a specific agent's beliefs rather within the system's own view of reality. This involves embedding a pretence cocoon within a belief space for the agent.

Conversely, in personification metaphor, it can be necessary to reasoning about the beliefs and reasoning of the entity that is metaphorically viewed as a person. This involves, in ATT-Meta terms, embedding a belief space within a pretence cocoon.

Reversed Transfers

The ATT-Meta approach is unusual in advocating that "reverse transfers" – transfers of information from target to source domain (more properly, reality to pretence) – are useful in metaphor understanding. One reason is the reverse transfer of reasoning queries that arise (notionally) from context. A query in target-domain terms can be reversed-transferred to become query in source-domain terms, and an example was given in the review of ATT-Meta above. This and two other reasons for doing reverse transfers are discussed at length in Barnden et al. (2004). One of them is based on an argument that, in the case of a conceptual metaphor being used in a distributed way across multiple utterances, it may be easier and more effective to form a coherent scenario in source-domain terms than to do so directly in target-domain terms by translating each metaphorical utterance into target-domain terms. This approach can instead involve "metaphorizing" the literal sentences in the relevant discourse segment: translating the information in them into source-domain terms. We present this possibility as a potentially fruitful topic for future research into metaphor.

Conclusion

AI is not just about the engineering of "intelligent" artefacts for useful purposes but also about mapping out the space of possible principles and mechanisms of cognition, whether artificial or natural. For the Engineering aim, metaphor is an important challenge, and AI can draw here on insights on the problem from many other disciplines. Conversely, through its non-Engineering aims, various features of AI – its partial applications focus, its input from Computer Science, its need or ambition to produce detailed processing accounts – put AI in a good position to help metaphor research. The help can consist of facilitating certain types of advance, identifying certain types of neglected problem, or effecting salutory changes of emphasis. This is not to say that these advances, problem identifications, and emphasis shifts could not arise from other disciplines, but just that AI is especially well-placed to generate them.

Specific helpful things that one can point to already as coming out of AI research on metaphor – whether they are advances, problem identifications or emphasis shifts – include the working out of detailed mechanisms for source-domain reasoning, the detailed elaboration of the alternative notion of within-pretence reasoning for metaphor, the casting of mappings as inference rules, the emphasis on and inclusion of gradedness in metaphor interpretation, mechanisms for exploiting context, the thorough inclusion of uncertainty into metaphorical reasoning, a richer view of overriding (source-over-target as well as target-over-source), important steps towards integration with metonymy interpretation, some emphasis on disanalogy, the usefulness of reversed transfers, steps towards mechanisms for handling the default transfer of attitudes and value judgments, the importance of non-assertional metaphor, enriched doubt about domains, and clarification and specification of ways in which literal meaning can be involved in metaphor interpretation.

All these matters require much further research, within AI and outside. But let us celebrate the fact that metaphor is, *par excellence*, an area for truly interdisciplinary investigation!

Acknowledgments

This chapter draws on research supported by grants GR/M64208 and EP/C538943/1 from the Engineering and Physical Sciences Research Council of the United Kingdom. I am indebted to the metaphor research group in my department – Sheila Glasbey, Mark Lee, and especially Alan Wallington – for their help with the development of this article, and to Ray Gibbs for comments and guidance.

References

Asher, N., & Lascarides, A. (1995, March). Metaphor in discourse. In *Proceedings of the AAAI Spring Symposium Series: Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity, and Generativity* (pp. 3–7). Stanford, CA.

Barnden, J. A. (1998). Combining uncertain belief reasoning and uncertain metaphor-based reasoning. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society* (pp. 114–119). Mahwah, NJ: Lawrence Erlbaum Associates.

Barnden, J. A. (2001a). Uncertainty and con?ict handling in the ATT-Meta context-based system for metaphorical reasoning. In V. Akman,

P. Bouquet, R. Thomason, & R. A. Young (Eds.), *Proceedings of the Third International Conference on Modeling and Using Context: Vol. 2116. Lecture Notes in Artificial Intelligence* (pp. 15–29). Berlin: Springer.

Barnden, J. A. (2001b). *Application of the ATT-Meta metaphor-understanding approach to selected examples from Goatly* (Technical Report CSRP-01–01). Birmingham, UK: School of Computer Science, University of Birmingham.

Barnden, J. A. (2001c). *Application of the ATT-Meta metaphor-understanding approach to various examples in the ATT-Meta project databank* (Technical Report CSRP-01–02). Birmingham, UK: School of Computer Science, University of Birmingham.

Barnden, J. A. (2001d). The utility of reversed transfers in metaphor. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Meeting of the Cognitive Science Society* (pp. 57–62). Mahwah, NJ: Lawrence Erlbaum Associates.

Barnden, J. A., Glasbey, S. R., Lee, M. G., & Wallington, A. M. (2002). *Application of the ATT-Meta metaphor-understanding system to examples of the metaphorical view of* TEACHERS AS MIDWIVES (Technical Report CSRP-02–10). Birmingham, UK: School of Computer Science, University of Birmingham.

Barnden, J. A., Glasbey, S. R., Lee, M. G., & Wallington, A. M. (2003). Domain-transcending mappings in a system for metaphorical reasoning. In *Conference Companion to the 10th Conference of the European Chapter of the Association for Computational Linguistics* (EACL 2003) (pp. 57–61). Association for Computational Linguistics.

Barnden, J. A., Glasbey, S. R., Lee, M. G., & Wallington, A. M. (2004). Varieties and directions of interdomain influence in metaphor. *Metaphor and Symbol, 19*(1), 1–30.

Barnden, J. A., Helmreich, S., Iverson, E., & Stein, G. C. (1994). An integrated implementation of simulative, uncertain and metaphorical reasoning about mental states. In J. Doyle, E. Sandewall, & P. Torasso (Eds.), *Principles of knowledge representation and reasoning: Proceedings of the Fourth International Conference* (pp. 27–38). San Mateo, CA: Morgan Kaufmann.

Barnden, J. A., Helmreich, S., Iverson, E., & Stein, G. C. (1996). Artificial intelligence and metaphors of mind: within-vehicle reasoning and its benefits. *Metaphor and Symbolic Activity, 11*(2), 101–123.

Barnden, J. A., & Lee, M. G. (1999). An implemented context system that combines belief reasoning, metaphor-based reasoning and uncertainty handling. In P. Bouquet, P. Brezillon, & L. Serafini (Eds.), *Second International and Interdisciplinary Conference on Modeling and Using Context: Vol. 1688 Lecture Notes in Artificial Intelligence* (pp. 28–41). Berlin: Springer.

Barnden, J. A., & Lee, M. G. (2001). *Understanding usages of conceptual metaphors: An approach and artificial intelligence system* (Technical Report CSRP-01–05). Birmingham, UK: School of Computer Science, University of Birmingham.

Carbonell, J. G. (1980). Metaphor – key to extensible semantic analysis. In *Proceedings of the 18th Annual Meeting of the Association for Computational Linguistics* (pp. 17–21).

Carbonell, J. G. (1982). Metaphor: An inescapable phenomenon in natural-language comprehension. In W. Lehnert & M. Ringle (Eds.), *Strategies for natural language processing* (pp. 415–434). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carston, R. (2002). Thoughts and utterances: The pragmatics of explicit communication. Oxford: Blackwell.

Delfino, M., & Manea, S. (2005, April). Figurative language expressing emotion and motivation in a web based learning environment. In *Proceedings of the Symposium on Agents That Want and Like: Motivational and Emotional Roots of Cognition and Action* (pp. 37–40), at AISB'05 Convention, University of Hertfordshire, UK. Brighton, UK: Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Dirven, R., & Pörings, R. (Eds.). (2002). *Metaphor and metonymy in comparison and contrast*. Berlin: Mouton de Gruyter.

Emanatian, E. (1995). Metaphor and the experience of emotion: The value of cross-cultural perspectives. *Metaphor and Symbolic Activity, 10*(3), 163–182.

Fainsilber, L., & Ortony, A. (1987). Metaphorical uses of language in the expression of emotions. *Metaphor and Symbolic Activity, 2*(4), 239–250.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*(1), 1–63.

Fass, D. (1997). *Processing metaphor and metonymy*. Greenwich, CT: Ablex.

Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science, 22*(2), 133–187.

Fussell, S. R., & Moss, M. M. (1998). Figurative language in descriptions of emotional states. In S. R. Fussell & R. J. Kreuz (Eds.), *Social and Cognitive Approaches to Interpersonal Communication*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 95–119.

Gerrig, R. J. (1989). Empirical constraints on computational theories of metaphor: Comments on Indurkhya. *Cognitive Science, 13*(2), 235–241.

Gibbs, R., & Tendahl, M. (2006). Cognitive effort and effects in metaphor comprehension: Relevance theory and psycholinguistics. *Mind & Language, 21*, 379–403.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8*(3), 183–206.

Goossens, L. (1990). Metaphtonymy: the interaction of metaphor and metonymy in expressions for linguistic action. *Cognitive Linguistics, 1*, 323–340.

Gross, L. (1994). Facing up to the dreadful dangers of denial [U.S. ed.]. *Cosmopolitan, 216*(3).

Hall, R. P. (1989). Computational approaches to analogical reasoning: a comparative analysis. *Artificial Intelligence, 39*, 39–120.

Hintikka, J., & Sandu, G. (1990). Metaphor and the varieties of lexical meaning. *Dialectica, 44*(1–2), 55–78.

Hobbs, J. R. (1990). *Literature and cognition* (CSLI Lecture Notes No. 21). Stanford, CA: Stanford University.

Hobbs, J. R. (1992). Metaphor and abduction. In A. Ortony, J. Slack, & O. Stock (Eds.), *Communication from an Artificial Intelligence perspective: Theoretical and applied issues* (pp. 35–58). Berlin: Springer-Verlag.

Hobbs, J. R., Stickel, M. E., Appelt, D. E., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence, 63*, 69–142.

Indurkhya, B. (1991). Modes of metaphor. *Metaphor and Symbolic Activity, 6*(1), 1–27.

Indurkhya, B. (1992). *Metaphor and cognition: An interactionist approach*. Dordrecht: Kluwer.

Iverson, E., & Helmreich, S. (1992). Metallel: An integrated approach to non-literal phrase interpretation. *Computational Intelligence, 8*(3), 477–493.

Kittay, E. F. (1989). *Metaphor: Its cognitive force and linguistic structure* (Paperback ed.). Oxford: Clarendon Press.

Kövecses, Z. (2000). *Metaphor and emotion: Language, culture, and body in human feeling*. New York: Cambridge University Press.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.).(pp. 202–251) Cambridge: Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.

Lee, M. G., & Barnden, J. A. (2001a). Reasoning about mixed metaphors with an implemented AI system. *Metaphor and Symbol, 16*(1&2), 29–42.

Lee, M. G., & Barnden, J. A. (2001b). Mental metaphors from the master metaphor list: Empirical examples and the application of the ATT-Meta system (Technical Report CSRP-01–03). Birmingham, UK: School of Computer Science, University of Birmingham.

Leezenberg, M. (1995). *Contexts of metaphor*. ILLC Dissertation Series, 1995–17, Institute for Language, Logic and Computation, University of Amsterdam, The Netherlands.

Levin, S. R. (1988). *Metaphoric worlds*. New Haven, CT: Yale University Press.

Lytinen, S. L., Burridge, R. R., & Kirtner, J. D. (1992). The role of literal meaning in the comprehension of non-literal constructions. *Computational Intelligence, 8*(3), 416–432.

Martin, J. H. (1990). *A computational model of metaphor interpretation*. San Diego, CA: Academic Press.

Martin, J. H. (1996). Computational approaches to figurative language. *Metaphor and Symbolic Activity, 11*, 85–100.

Martin, J. H. (2000). Representing UNIX domain metaphors. *Artificial Intelligence Review, 14*(4–5), 377–401.

Mason, Z. J. (2004). CorMet: A computational, corpus-based conventional metaphor extraction system. *Computational Linguistics, 30*(1), 23–44.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

Musolff, A. (2004). *Metaphor and political discourse: Analogical reasoning in debates about Europe*. Basingstoke, UK: Palgrave Macmillan.

Narayanan, S. (1997). *KARMA: Knowledge-based action representations for metaphor and aspect*. Unpublished PhD thesis, Computer Science Division, EECS Department, University of California, Berkeley.

Narayanan, S. (1999). Moving right along: A computational model of metaphoric reasoning about events. In *Proceedings of the National Conference on Artificial Intelligence* (AAAI '99, pp. 121–128). AAAI Press.

Norvig, P. (1989). Marker passing as a weak method of text inferencing. *Cognitive Science, 13*(4), 569–620.

Ortony, A. (1979). The role of similarity in similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought* (pp. 186–201). Cambridge: Cambridge University Press.

Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence,* 29(3), 241–288.

Ruiz de Mendoza Ibáñez, F. J. (1999). *From semantic underdetermination via metaphor and metonymy to conceptual interaction* (Series A: General & Theoretical Papers No. 492). LAUD Linguistic Agency.

Russell, S. W. (1976). Computer understanding of metaphorically used verbs. *American Journal Computational Linguistics, 2*, 15–28.

Russell, S. W. (1985). Conceptual analysis of partial metaphor. In L. Steels & J. Campbell (Eds.), *Progress in Artificial Intelligence* (pp. 193–201). Chichester, UK: Ellis Horwood.

Russell, S. W. (1986). Information and experience in metaphor: A perspective from computer analysis. *Metaphor and Symbolic Activity, 1*(4), 227–270.

Russell, S., & Norvig, P. (2002). *Artificial Intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: Blackwell.

Stern, J. (2000). *Metaphor in context*. Cambridge, MA: Bradford Books<th>/<th>MIT Press.

Thomas, O. (1969). *Metaphor and related subjects*. New York: Random House.

Tourangeau, R., & Sternberg, R. J. (1982). Understanding and appreciating metaphors. *Cognition, 11*, 203–244.

Turner, M. (1987). *Death is the mother of beauty: Mind, metaphor, criticism*. Chicago: University of Chicago Press.

van Dijk, T. A. (1980). Formal semantics of metaphorical discourse. In M. K. L. Ching, M. C. Haley, & R. F. Lunsford (Eds.), *Linguistic perspectives on literature* (pp. 115–138). London: Routledge and Kegan Paul.

van Genabith, J. (2001). Metaphors, logic, and type theory. *Metaphor and Symbol, 16*(1&2), 43–57.

Veale, T. (1998). "Just in Time" analogical mapping, an iterative-deepening approach to structure-mapping. In *Proceedings of the Thirteenth European Conference on Artificial Intelligence* (ECAI '98). New York: John Wiley.

Veale, T., & Keane, M. T. (1992). Conceptual scaffolding: a spatially founded meaning representation for metaphor comprehension. *Computational Intelligence,* 8(3), 494–519.

Veale, T., & Keane, M. T. (1997). The competence of sub-optimal structure mapping on 'hard' analogies. In *Proceedings of the International Joint Conference on Artificial Intelligence*. Nagoya, Japan.

Vogel, C. (2001). Dynamic semantics for metaphor. *Metaphor and Symbol, 16*(1&2), 59–74.

Way, E. C. (1991). *Knowledge representation and metaphor*. Dordrecht: Kluwer.

Weber, S. H. (1989). Figurative adjective-noun interpretation in a structured connectionist network. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 204–211). Hillsdale, NJ: Lawrence Erlbaum Associates.

Weiner, J. (1984). A knowledge representation approach to understanding metaphors. *Computational Linguistics, 10*(1), 1–14.

Wilcox, P. P. (2004). A cognitive key: Metonymic and metaphorical mappings in ASL. *Cognitive Linguistics, 15*(2), 197–222.

Wilcox, S. (2004). Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed language. *Cognitive Linguistics, 15*(2), 119–147.

Wilks, Y. (1978). Making preferences more active. *Artificial Intelligence, 11*, 197–223.

Wilks, Y., Barnden, J., & Wang, J. (1991, August). Your metaphor or mine: Belief ascription and metaphor interpretation. In *Proceedings of the 12 th International. Joint Conference on Artificial Intelligence*, Sydney, Australia (pp. 945–950). San Mateo, CA: Morgan Kaufmann.

Winston, P. H. (1979). Learning by creating and justifying transfer frames. In P. H. Winston & D. Brown (Eds.), *Artificial Intelligence: An MIT*

Perspective (Vol. 1, pp. 347–374). Cambridge, MA: MIT Press.

Woll, B. (1985).Visual imagery and metaphor in British Sign Language. In W. Paprott é & R. Dirven (Eds.), *The ubiquity of metaphor: Metaphor in language and thought* (pp. 601–628). Amsterdam: John Benjamins.

Yu, N. (1995). Metaphorical expressions of anger and happiness in English and Chinese. *Metaphor and Symbolic Activity, 10*(2), 59–92.

Zhang, L., Barnden, J. A., & Hendley, R. J. (2005, July). Affect detection and metaphor in e-drama: The first stage. In *Proceedings of the Workshop on Narrative Learning Environments, 12th International Conference on Artificial Intelligence in Education* (AIED-2005), Amsterdam.