Considerations on the nature of metaphorical meaning arising from a computational treatment of metaphor interpetation

A.M.Wallington, R.Agerri, J.A.Barnden, S.R.Glasbey and M.G.Lee

School of Computer Science,
University of Birmingham, UK)
`A.M.Wallington@cs.bham.ac.uk`

Abstract

This paper argues that there need not be a full correspondence between source and target domains when interpreting metaphors. Instead, inference is performed in the source domain, and conclusions transferred to the target. A description of a computer system, ATT-Meta, that partially implements these ideas is provided.

1 Introduction

It is now generally accepted, especially since the work of Lakoff and associates (e.g. [14,13,15]) that much of everyday discourse shows evidence of metaphor. Consequently, the question of how metaphor should be interpreted and what the semantic consequences are of using a metaphor is of major importance in determining how discourse should be interpreted.

Like Stern [19], we take the position that much of the interpretation of metaphor is highly context dependent and involves pragmatics. However, we believe that, for metaphor, pragmatics must be informed by theories of Artificial Intelligence and psychology. Thus we have some sympathy for Levinson's ([16] p.161) claim that:

"the interpretation of metaphor must rely on features of our general ability to reason analogically. ... It could be claimed that linguistic pragmatics alone should not be expected to provide such a general theory of analogy, without considerable help from psychological theory [1] ."

We depart from Levinson with respect to analogy, and in this paper, we shall challenge current theories of analogy (e.g. [7,8,11] and what might

---

[1] Levinson makes it clear that psychological theory includes Artificial Intelligence.

be termed correspondence theories of metaphor (e.g, Lakoff's Conceptual Metaphor Theory) in which a source domain is put in correspondence with a target domain. We shall attempt to show that there is far less parallelism between source and target than is often assumed and that the process of interpreting a metaphor often requires heavy use of inferencing in order to associate source domain entities for which there is no parallel target equivalent, what we term "Map-Transcending Entities" (MTEs), with information that is involved in source to target transfer. Now other AI approaches to metaphor also emphasise the role of source domain inference, Hobbs [10] and Narayanan [18] for example. And, Martin's MIDAS system includes a process of extending conventional source to target mappings [17]. However, apart from important technical differences between these systems and our own, we would wish to stress the implications extensive inferencing has for source-target parallelism and the repercussions this has for the semantics and pragmatics.

We do not yet have a fully developed semantics and pragmatics of metaphor (although see [9] for some preliminary suggestions based on Stern). However, what we do provide is an extensively developed (though informal) inference-based model of metaphor understanding that employs an event-based formalism similar to that of Hobbs [10], combined with a realization of this approach in a fully implemented system (ATT-Meta) that effects the type of reasoning that we claim is at the heart of much metaphor understanding (See [1,4,21]).

2 Correspondence approaches to interpreting metaphor

The work of Lakoff and Johnson e.g. [14,15] not only stressed the ubiquity of metaphor in everyday discourse, but also noted that many metaphorical utterances could be systematically related to each other, all appealing to different aspects of the same source domain and being used to describe the same target domain. In other words, what is involved in metaphor is the mapping of one cognitive domain into another. For example, Lakoff [13] notes that the following metaphors all involve a source domain of journeys being used to describe a target domain of the progress of a love affair: *Look how far we've come. It's been a long, bumpy road. We can't turn back now. We're at a crossroads. The relationship isn't going anywhere. We may have to go our separate ways. We're spinning our wheels. Our relationship is off the track. The marriage is on the rocks. We may have to bail out of this relationship.* To account for this generalization, Lakoff assumes a "LOVE-AS-JOURNEY" mapping, i.e. "a set of ontological correspondences that characterize epistemic correspondences by mapping knowledge about journeys onto knowledge about love" ([13] p207). Specifically, he assumes the following ontological correspondences:

THE LOVE-AS-JOURNEY MAPPING

i. The lovers correspond to travellers.
ii. The love relationship corresponds to the vehicle.

iii. The lovers' common goals correspond to their common destination.
iv. Difficulties in the relationship correspond to impediments to travel.

Lakoff does not spell out in any detail exactly how the epistemic correspondences function, and how inferences made in the source transfer to the target. He does however, claim that mappings "project source domain **inference patterns** onto target domain **inference patterns**" ([13] p245, emphasis added). However, we might turn to more formal work in analogy for a possible mechanism. Gentner (e.g. [7]) argues that complex systematic relations between source domain entities such as causal relations also transfer, whilst non-systematic relations such as attributes do not.

3   Map Transcending Entities

Let us return to Lakoff's list above of travel metaphors. Consider the statement that *we're spinning our wheels*. How might we infer from this that the love relationship is not progressing as it might? Plausibly, the following chain of inferences might be entertained. If wheels are referred to, then, defeasibly [2], a vehicle is involved. The spinning wheels are causing the vehicle not to move as it should. If the vehicle is not moving as it should then it is not moving towards its destination.

What can we make of this pattern of inferences? Both the vehicle and the destination have correspondents in the target domain of the love affair, namely the love relationship and the lovers' common goals. With these correspondences, we might assume that the source domain conclusion can be transferred to become a target domain conclusion. But, this is the transfer of it conclusion. It could never have been reached without the premise that something -the spinning wheels- was causing the vehicle not to proceed. So what is the target correspondent of the spinning wheels whose presence is needed in order to allow the entire inference pattern as opposed to just the conclusion to transfer from source to target? Nothing in the list of four ontological correspondences would seem to be appropriate.

If we look at the other examples listed, we find similar cases where the lack of a target domain correspondent would prevent crucial aspects of the inference pattern mapping from source to target. For example, what is the target domain correspondent of *the rocks* in the utterance *our marriage is on the rocks*? A similar point can be made about the location *off the track* in *our relationship is off the track*. It is not that these statements are uninterpretable. Both would seem to permit the same conclusion that was reached about the spinning wheels, namely that the relationship/vehicle is not progressing towards the destination and hence not towards the lovers' goal.

Note that it does not seem quite right to assume that *the spinning wheels, rocks* or *lack of tracks* are "impediments" in the sense of the third of Lakoff's

---

[2] We shall henceforth assume that inferences are defeasible unless otherwise stated.

correspondences. There may be an interpretation under which the *bumps* in the statement *its been a long bumpy road* refer to specific, listable, difficulties in the love relationship, and similarly, *the rocks* may refer, for some, to a particular event, but both could be used more generally. Whatever, if anything, *wheels* might refer to it is a vague and very unspecific target domain entity.

Let us now consider the utterance: *We're at a crossroads*. Stating that we are at some location on a road might be taken to implicate that we are going somewhere along the road and hence have a destination. This would map to a target domain "common goal" However, there is no specific correspondent given for crossroads. The source domain inference that if one is at a crossroads, then there is a choice of possible destinations cannot transfer to the target domain inference that there is a choice of possible relationship goals, if, as is stated, inference patterns are mapped from one domain to another.

Now, a lack of target correspondents for source domain entities is not peculiar to the LOVE AS A JOURNEY conceptual metaphor, but is generally the case. In section 3, we shall give a description of our approach to metaphorical reasoning and our computational system ATT-Meta that performs such reasoning. A sentence that we shall analyse in some detail is the following:

1 In the far reaches of her mind, Mary believed Kyle was having an affair.

We assume that there is a mapping between ideas such as the idea that Kyle was having an affair and physical, manipulable, objects. This reification of ideas has a wide currency. We shall also assume here that the mind is often viewed as a physical space. However, what does *the far reaches* refer to?

So what are we to make of these lacunae which prevent the step by step transfer of inferences about the source becoming inferences about the target? Note that the absence of target domain correspondents of source domain entities is not a mere technical problem in determining how source domain implications transfer and become target domain implications, nor should we assume that Lakoff's claim that entire **inference patterns** as opposed to certain **conclusions** transfer is just an imprecise form of words. If there are no target domain correspondents of source domain entities, then we must assume that there are entities in texts have no reference even indirectly via a target domain equivalent to anything in the real world. There are entities that can only be made sense of in terms of their literal meaning in the source domain; a somewhat similar conclusion to Davidson's [5] well known claim that metaphors have only a literal meaning.

One possibility is that the four correspondences given in the LOVE AS A JOURNEY mapping in section 2 are not intended to be exhaustive and the mapping would if properly specified include correspondences for the entities we have discussed. For example, the *crossroads* example might motivate adding the following correspondence to the mapping.

v. A choice of goals corresponds to a choice of destinations.

However, a number of objections can be made to this view. Many concep-

tual metaphors are open-ended allowing almost any entity associated with the source domain to be used when speaking of the target. Now, not all of these might be conventional. They might make use of what Lakoff and Johnson ([14] p 53) call "the unused part" of the source domain. For example, Lakoff [13] gives an example of a creative use of the conceptual metaphor LOVE IS A JOURNEY, the song lyric *we're riding in the fast lane on the freeway of love*. There is no correspondent listed for *fast lane* nor for a *freeway of love*. It would seem that no finite list of correspondents would ever be enough. This criticism would also defeat lexically based treatments of metaphor.

The last example contained very little that could plausibly correspond to target-domain entities; perhaps just the inference that a journey and thus a destination is involved. However, it might be conceded that whilst many modifier phrases often do not correspond, overall statement-like constituents of the source will correspond to statements in the target. However, consider the following example from real discourse of a particular metaphorical view running through several consecutive statement-like utterances:

"She was too confused to judge anything. If he'd done nothing else, he'd at least invaded her thoughts and tossed them around until there was only chaos in her head."

We shall assume that the clauses *he'd at least invaded her thoughts*, *[he had] tossed them around* and *there was only chaos in her head* all rely on the metaphorical views of IDEAS AS PHYSICAL OBJECTS and MIND AS PHYSICAL SPACE, and taken together partially convey, in target-domain terms, that the man, "he", had done something that had resulted in the thoughts of the woman, "her", **not functioning as they ought**. But is there any need to assume a specific target-domain parallel for *tossing physical objects around*? A similar point can be made about the statement *he'd ... invaded her thoughts*: the invading is just mentioned as a way of emphasizing that he had done something that had had a strong effect on her thoughts.

Now, it is possible that the mention of invading may have been used if the man had introduced thoughts into the woman's mind by saying things to her; the invading corresponding to introducing. It is then these thoughts that caused the tossing of the woman's pre-existing thoughts. However, it is possible to imagine the above discourse segment occurring in a context where it is clear that the man had not communicated anything to her.

In short, even a sentence, the traditional unit of analysis of truth, might have no meaning other than in terms of the source domain. In the rest of this paper we shall describe our approach and implemented computational model, ATT-Meta, for reasoning with metaphorical utterances that contain MTEs.

4   Inferencing and ATT-Meta

Recall that our proposed fifth correspondence in the previous section had a 'choice' in the target corresponding to a 'choice' in the source. If we look at the

fourth correspondence, we find something similar. A target domain 'difficulty' corresponds to a type of difficulty or 'impediment' in the source. Note also that Lakoff argues a-propos the 'fast lane' song lyric that driving in the fast lane is exciting and that the excitement transfers to the target. And we could give examples in which other emotional states transfer from the source to the target. In the 'invasion' passage, the conclusion that the woman's mind was not functioning properly was likewise transferred. What all these cases have in common is that they involve essentially 'invariant' transfers from the source to the target. These are of a very different nature from the cross-domain correspondences between say journeys and love affairs. We argue that such transfers are affected by what we term "View-Neutral Mapping Adjuncts" and argue that they apply universally, at least by default, regardless of what cross-domain mappings are in play and our system has made a start both at investigating what VNMAs are required and at formalising some of them. (See [2,3,21] for more details.)

In order to sketch our approach let us return to the Anne/Kyle example:

1 In the far reaches of her mind, Anne believed Kyle was having an affair[.]

and compare it to the following variant:

2 Anne had in her mind the belief that Kyle was having an affair.

We assume that both sentences utilize the conceptual metaphors (a term which we replace with the more neutral "metaphorical view" in our work): (A) IDEAS AS PHYSICAL OBJECTS (B) MIND AS PHYSICAL SPACE.

We assume that one correspondence included under view (A) is the following: "Conscious mental processing corresponds to physical manipulation." (See [12] for motivation). We also assume that with activities such as processing/manipulating, which one can have the ability to perform to a greater or lesser degree, DEGREE is a VNMA and maps over from source to target in an invariant manner. Thus a very low ability to mentally process an idea corresponds to a very low ability to physically manipulate an object.

We assume that accompanying metaphorical view (B) are two ancillary assumptions. Firstly, that the conscious self of the mind's possessor is metaphorically cast as a person physically located in (a central part of) the mind-region. Secondly, that when a cognitive state (such as believing) is cast as located in a physical sub-region of a mind, then the idea or whatever that is the object of the state is also to be thought of as physically located in that sub-region. As we suggested in the previous section, we assume that there is no known correspondent for the far reaches; it is a map-transcending entity.

So how does the informational contribution of (1) and (2) differ? Plausibly, what (2) principally conveys to the reader is that Anne has the ability to operate in a conscious mental way on the idea that Kyle was having an affair. In brief: Anne is aware of the affair. By contrast, what (1) seems to convey is that the ability to operate holds only to a very low degree. In brief: Anne had very little conscious awareness of the affair.

Thus, the situation described by the 'far reaches' utterance is cast as being one where Anne's conscious self is a person in a central part of Anne's mind-region, and the idea that Kyle was having an affair is in the far reaches of the mind-region. Now, let us assume that the understander's common sense knowledge of physical space and physical objects includes the following:

* things in the far reaches of a region are usually distant from things in a central part (distance being relative to the scale of the whole region).
* if a person is physically distant from a physical object then the person usually has only a very low degree of ability to manipulate that object physically.

Thus, the understander can reason, within the terms of the source domains of the metaphorical views (PHYSICAL SPACE and PHYSICAL OBJECTS), that, probably, Anne's conscious self has only [3] a very low degree of ability to physically manipulate the idea that Kyle was having an affair.

This conclusion can become the target-domain conclusion that Anne has only a very low degree of ability to operate in a conscious mental way on the idea that Kyle was having an affair, by virtue of the correspondence between physical manipulation and conscious mental processing that was assumed as an aspect of the IDEAS AS PHYSICAL OBJECTS mapping, and by virtue of the VNMA invariantly mapping the very low degree from source to target.

In our approach source-target correspondences are implicit in transfer rules. In the case of the correspondences just mentioned, English glosses of the relevant rules include:

```
* IF in reality X is a person and K is an idea
* AND K is being viewed as a physical object
* AND person X's conscious self is being viewed as being able to
  operate physically on K to at least degree D
* THEN presumably in reality X can mentally operate consciously
  on K to degree at least D.
```

This rule allows one aspect of the source-domain conclusion to lead to the target-domain conclusion that Anne can mentally operate consciously on the Kyle-affair idea to degree at least "very low".

In sum, our approach involves the following main types of processing:

* Construction of a representation of the direct, source-domain meaning of the utterance, i.e. the meaning it has by taking only the source-domain senses of the metaphorically-used words/phrases in the utterance. This meaning consists of one or more propositions.
* In some cases, application of ancillary assumptions associated with the relevant metaphorical views to create further propositions in source-domain terms.
* Usually, performance of source-domain reasoning on the basis of the direct

---

[3] A very low degree of ability might implicate that Anne does not have a higher degree, but does not entail it. Hence our addition of 'only'.

source-domain meaning, the products of ancillary assumptions, and general knowledge relevant to the source domain meaning.

* Source-to-target transfer acts by application of transfer rules (and VNMAs).

This listing does not imply any particular temporal ordering of the types of processing. Indeed in ATT-Meta the reasoning actually works backwards from reasoning queries posed internally within the system and can involve any intertwining and ordering of instances of the above types of reasoning.

An important feature of our approach that we have not yet mentioned is that it encapsulates the source-domain reasoning based on the literal meaning of the utterance within a special computational context we call a pretence cocoon. Metaphorical transfer acts based on rules such as those above operate between the inside of the pretence cocoon and the reality-context outside. Thus, for the Anne/Kyle example, the understander pretends, within the cocoon, that Anne's mind really is a physical space and that the believing really does occur in the far reaches of this space. Consequences of this are inferred in the pretence cocoon, possibly by substantial amounts of reasoning, using ancillary assumptions and knowledge about physical objects and space. The conclusions reached may then be able to be transmuted, via transfer rules forming part of the relevant metaphorical views, into propositions in the reality environment. However, we ought to stress that many different lines of reasoning will be explored, many ultimately proving unsuccessful.

We should also stress that when a pretence cocoon is created, it is not tagged as having to do with any particular metaphorical view. Only by having the utterance's direct source-domain meaning placed within it, such as the mind having far-reaches, can an inference be made that that the particular metaphorical view MIND AS PHYSICAL SPACE with its associated correspondences is being used. Thus, even the question of the metaphorical views involved in an utterance results from a possibly extensive web of inferences.

Finally note that although Anne's mind is categorized in the pretence as a physical region, this is in addition to its being categorized there as a mind. (Thus, a pretence cocoon is reminiscent of a blend space in Blending Theory: [6].) Given the existence of suitable knowledge rules, such as that a mind is not a physical region, we can get conflicting propositions arising within the pretence, because in general it is wrong to prevent rules about the target domain operating within the pretence. In the present case we would get both strong support for the mind being a physical region and for its not being a physical region. The ATT-Meta system implements conflict-resolution mechanisms that deal with reasoning conflicts in general, and that embody a small number of general principles about conflict resolution in metaphor [1,20]. In the present case, the mechanisms ensure that the proposition that Anne's mind is a physical region wins over the proposition that it is not.

5   Conclusion

We have provided a brief outline of some of the ideas in our implemented, inference-based approach to metaphor. Much more detail, including the application to other examples, can be found elsewhere [1,2,4].

The main point has been the use of inference to connect source-domain aspects that are raised by an utterance but not handled by known metaphorical mappings to source-domain aspects in mappings that the understander does know, and particularly to knowledge of what invariant aspects of metaphorical utterances are likely to transfer. By this means, the approach can deal with open-ended extensions of metaphorical beyond what can be readily dealt with by known mappings by themselves, without the need for creating mappings for the unmapped source-domain aspects.

We thus radically downplay source/target parallelism in metaphor in favour of inference, and place great weight on the thesis that metaphors often introduce source-domain aspects that do not need any correspondents in the target domain (let alone already have any): their only purpose is to support useful lines of source-domain inference that connect to known mappings. One of the interesting semantic issues raised is that these unmapped aspects do not by themselves have any meaning in target-domain terms, and it would be a mistake to try to specify such meaning.

6   acknowledgements

References

[1] Barnden, J.A. (2001) Uncertainty and conflict handling in the ATT-Meta context-based system for metaphorical reasoning. In, V. Akman, P. Bouquet, R. Thomason and R.A. Young (Eds), *Procs. Third International Conference on Modeling and Using Context. Lecture Notes in Artificial Intelligence*, Vol. 2116. Berlin: Springer, 15-29.

[2] Barnden, J.A. and Lee, M.G., (2001). *Understanding open-ended usages of familiar conceptual metaphors: An approach and artificial intelligence system.* Technical Report CSRP-01-05, School of Computer Science, University of Birmingham.

[3] Barnden, J.A., Glasbey, S.R., Lee M.G. and Wallington, A.M. (2003). Domain-transcending mappings in a system for metaphorical reasoning. In *Proceedings of the Research Note Sessions of the 10th Conference of EACL.*

[4] Barnden, J.A., Glasbey, S.R., Lee, M.G. and Wallington, A.M. (2004), Varieties and directions of inter-domain influence in metaphor. *Metaphor and Symbol*

19(1), 1–30.

[5] Davidson, D. (1979). What metaphors mean. In, S. Sacks (Ed.), *On Metaphor*. U. Chicago Press, 29-45.

[6] Fauconnier, G and Turner, M. (2002). *The Way We Think: Conceptual Blending and the Minds Hidden Complexities*. NY: Basic Books.

[7] Gentner, G. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155–170.

[8] Gentner, D., Falkenhainer, B. and Skorstad, J. (1988). Viewing metaphor as analogy. In D.H. Helman (Ed.), *Analogical reasoning. Dordrecht*. Kluwer.

[9] Glasbey, S.R and Barnden, J.A. (submitted). Towards a situation-based discourse semantics for metaphor. Submitted to the journal *Research on Language and Computation*.

[10] Hobbs, J.R. (1990) *Literature and Cognition* CSLI Lecture Notes, Center for the Study of Language and Information, Stanford University.

[11] Holyoak, K J. and Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13(3), 295-355.

[12] Jaekel, O. (1995). The Metaphorical Concept of Mind, in J.R. Taylor and R.E. MacLaury (eds), *Language and the Cognitive Construal of the World*. Berlin New York, Mouton de Gruyter. 197–229.

[13] Lakoff, G. (1993). The contemporary theory of metaphor. In A Ortony (Ed.), *Metaphor and Thought*, 2nd ed. Cambridge, UK: Cambridge University Press.

[14] Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

[15] Lakoff, G. and Johnson, M. (1999). *Philosophy in the Flesh*. NY: Basic Books.

[16] Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

[17] Martin, J. H. (1990). *A Computational Model of Metaphor Interpretation*. NY: Academic Press.

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

[19] Stern, J. (2000). *Metaphor in Context*. Cambridge, MA and London, UK: Bradford Books, MIT Press.

[20] Wallington, A.M and Barnden, J.A. (2004). Uncertainty in Metaphorical Reasoning. In *Procs of the Workshop on Computational Models of Natural Argument (CMNA) at ECAI 2004*. August 2004, Valencia, Spain.

[21] Wallington, A.M., Barnden, J.A. Glasbey S.R. and Lee M. G. (2006). *Metaphorical reasoning with an economical set of mappings*. Delta, 22:1.