UNCERTAINTY IN METAPHORICAL REASONING

Alan Wallington and **John Barnden** [1]

Abstract. We present an approach and an associated computer program (ATT-Meta) which can interpret novel uses of familiar metaphors. ATT-Meta performs sophisticated reasoning with uncertainty of a kind that is required for reasoning about the content of metaphorical utterances. ATT-Meta's handling of uncertainty is qualitative and handles potential conflicts between different lines of reasoning or arguments. It also uses a particular approach to specificity, which involves a complex examination of the complete argument structures supporting the conflicting hypotheses/arguments, as a powerful tool for comparing arguments.

1 INTRODUCTION

A claim of much recent work in cognitive linguistics is that metaphor plays a fundamental role in our ways of viewing and conceptualizing the world, especially in the conceptualization of abstract notions such as thought, emotion, purpose, scientific theory (e.g. [9]; [7]). It is argued that participants in a language community share sets of fixed mapping links between source and target domains, using knowledge of a source domain as if it were in reality knowledge about the target. We commonly talk for example of *getting an idea into our mind*, even thought we know very well that minds are not containers and that ideas are not objects that may occupy such containers. So much is quite widely accepted (see [5] for an alternative). Furthermore, speakers may *extend* and *elaborate* on these source-to-target mapping links in a creative manner (see [10]). But, this creativity creates a problem when we try to model the process of metaphor understanding, especially for models based on the structure-mapping approach to analogy (e.g. [4], [3]) since it suggests that new mapping links need to be created on the fly whenever an elaboration of an existing metaphor introduces a new source domain concept for which there is no existing mapping. Searching for new mapping links is a computationally intensive process [14]. We have devised an alternative approach to metaphor interpretation and developed an associated computer program ATT-Meta (e.g. [1]). Instead of searching for analogues of what we call 'map-transcending' elements, we argue that metaphor interpretation requires only an 'economical' set of transfer rules mapping between source and target domains. Metaphors are extended through the use of, often extensive, source domain reasoning, using exactly the same type of inference rules that would normally be used in non-metaphorical discourse about the source domain.

If we are right, then a task for metaphor theory should include designing systems that can reason. Now, in the real, commonsense, imperfect, world, reasoning is non-monotonic and there are conflicting arguments. The addition of more specific information about a particular domain will frequently force the retraction of prior conclusions. Tentative inferences and conclusions are made even though knowledge will be imperfect or incomplete. Systems that can model such uncertain reasoning are highly desirable in many NLP applications and become indispensable when the language used may be metaphorical or not, for almost all metaphor can be seen as involving a conflict between what is known to be true and what is claimed in the metaphorical utterance. And, much of natural language is metaphorical.

The need for uncertain reasoning systems to deal with natural language and in particular with metaphor constrains the particular approach to uncertain reasoning that we take. It is hard to see where the necessary numbers for prior and conditional probabilities would come from (at least within a realistic time frame) that would inform probabilistic approaches to uncertain reasoning. Yet qualifiers denoting the uncertainty of a proposition abound in natural language. Consider adverbial qualifiers such as "usually". Humans are able to react quickly and without conscious effort to such qualifiers without demanding any numerical measures. Consequently, our system uses qualitative measures of certainty [11].

Thus our approach and associated computer program (ATT-Meta), which we have demonstrated on a variety of examples taken from real discourse, performs reasoning with uncertainty of a kind that is firstly qualitative and secondly argument-based [12]. In other words, hypotheses or reasoning queries are tagged with qualitative certainty levels allowing an argument for a proposition with a high certainty level to win out over an argument for its complement with a lower certainty level. And furthermore, whilst the reasoning is essentially logical consisting of 'if-then' type rules, there is not one chain of reasoning that proves or disproves a hypothesis. Instead, there will be arguments for and against the hypothesis. Indeed, our system will always explore arguments against any seemingly successful hypothesis. And all of these arguments, for and against, are taken into account when deciding whether to accept the hypothesis. Thus we take a query or goal directed approach in which an initial query is posed by the discourse surrounding a particular utterance. Then this query and its converse are both investigated, by backward-chaining 'if-then' rules, each with a particular qualitative certainty level, until they are grounded in facts.

These arguments being investigated will involve information derived from metaphorical, literal information, and information derived from both, interacting in a complex manner. There will also be cases where the qualitative certainty levels of the respctive arguments may not decide between the two. In such case, our novel approach to specificity (see section 4), in which a complex examination of the complete argument structures of the competing hypotheses may decide.

And the winning arguments may consist of cases where 'Target' knowledge defeats inferences based on a metaphorical reading, but there will also be cases, which are not widely recognized within the metaphor literature, where inferences derived from a metaphorical

[1] School of Computer Science, University of Birmingham, Birmingham, B152TT, UK. email: J.A.Barnden@cs.bham.ac.uk

© 2004
CMNA IV, 4^{th} *Workshop on Computational Models of Natural Argument*
Edited by Floriana Grasso, Chris Reed and Giuseppe Carenini
ECAI 2004

reading may defeat default target assumptions (see [2]).

The ATT-Meta system is in a state of continuous improvement. The system of reasoning has recently been greatly improved and made much more general. Aspects of this new approach are described in this paper.

2 AN EXAMPLE OF INFERENCE IN METAPHOR

Our approach takes metaphor to be a way of seeing or describing something as though it were something else. Thus, we assume that the participants in a discourse involving a metaphor have implicitly made the following agreement: we know that TARGET is not SOURCE but for the current purposes we shall pretend that TARGET is a particular kind of SOURCE and consequently can license the same kind of inferential links that other kinds of SOURCE would license. We call this type of metaphorical reasoning based on a shared pretence "within-pretence reasoning" and the computational space in which this reasoning takes place the "pretence-space" or "pretence" as opposed to the "reality-space" where reasoning about the target takes place.

Consider the following example, which is slightly adapted from one in Cosmopolitan magazine:

"In the far reaches of her mind, Anne believes that Kyle is unfaithful".

Anne's mind is viewed within the pretence as a physical space of the kind that can have "far reaches". We assume that certain aspects of the source domain (here, that of physical space) are mapped onto corresponding aspects of the target domain (here, that of mental objects and processes). The mapping is, however, not necessarily complete - i.e., there may be aspects of the source domain that have no corresponding mapping to the target domain. For example, we assume that the understander may posses no direct mapping from "the far reaches" to any aspect of Anne's mind. Instead, using source domain knowledge, inferences will be made about possible relations between the "far reaches" of a space and more central areas and some of the conclusions reached may match mappings between the source and target. In brief, and omitting many details, we assume that it can be inferred that a distant object would be difficult to manipulate for an individual such as Anne's conscious self located in the centre of the mind-space. Crucially, to arrive at this mapping from the original statement, a process of inferencing was required and the inferences made were uncertain.

This inferencing took place in the pretence-space using source domain information but, in order to properly integrate the metaphorical utterance with the surrounding discourse, complex uncertain inferencing will also take place in the reality-space using target domain information. To complicate things further, reasoning within the pretence sometimes relies on target domain facts, -the existence of Anne as a real person would be such a case- and this may give rise to further conflict. This example concerning Anne and Kyle and the different types of knowledge involved is one that we have implemented. It involves a much more complex level of inferencing than has been sketched here, but it can be demonstrated.

3 ATT-META'S MODE OF REASONING

ATT-Meta itself has no knowledge of any specific metaphorical view and nor does it directly interpret natural language- it is simply a reasoning engine. However, it includes some built-in rules about pretence, beliefs, qualitative degrees, etc. The user supplies as data to ATT-Meta whatever target domain knowledge, source domain knowledge and metaphorical transfer relationships that s/he assumes will adequately express the meaning of the metaphorical utterance, the co-text, and the likely background knowledge. In particular the user supplies fact-rules that express the direct meaning of the metaphorical utterance. This information is expressed as 'if-then' rules, with fact-rules as a special case. Simplifying the rules somewhat for the purposes of this paper, a simple rule about birds would be expressed as follows[2]:

```
IF bird(X) AND alive(X) THEN {presumed} can-fly(X).
```

Note that there are three main parts to a rule: the *if* part; the *then* part; and the *qualitative-certainty-level*. The symbol 'presumed' is a qualitative certainty qualifier and can also be read as "by default". Penguins, of course, cannot fly and would be an exception to this rule. If-then rules with a null IF part count as facts e.g.:

```
{certain} is-person(Anne).
```

ATT-Meta's reasoning is entirely query-directed. Query-directed reasoning, also called goal-directed reasoning, is a powerful technique, much used in AI (see, e.g., [13]). The process of reasoning starts with a 'top' query -i.e. a question as to whether some proposition holds or not. Queries are compared to known propositions for a possible match. They are also compared to the 'then' parts of rules. In the case of ATT-Meta, these may be either standard or conversion rules, and the facts may be stated as holding in the metaphor pretence or in reality.

ATT-Meta tries to find evidence for and against the top query, using the user-supplied knowledge and logical forms of the utterance(s). In cases of conflict, a conflict-resolution mechanism (see section 4) comes into play. If the top query contains variables ATT-Meta will, in addition, try to find values for the variables.

A further source of uncertainty is that at any time, any particular hypothesis H being entertained by ATT-Meta, including the top query, is tagged with a qualitative certainty level, one of **certain, presumed, suggested, possible** or **certainly-not**. **Presumed** means that H is a default, **suggested** means that there is evidence for H but not (yet) enough to enable H to be a working assumption; **possible** means that there is not yet certain evidence against H. When a hypothesis is created (as a query) it is immediately given a certainty value of **possible**. Reasoning may then upgrade or downgrade it as appropriate.

When a rule is applied, the certainty it contributes to the result is the minimum of its own certainty qualifier and the certainty levels assigned to the hypotheses picked up by the condition part. Multiple rule applications can support a given hypothesis - in this case the maximum of the certainty values contributed by the different applications is used. When there is evidence to level **presumed** for both a query and its negation, then the conflict-resolution mechanism discussed in the next section tries to adjudicate the relative evidence strength.

4 THE 'EVEN-THOUGH' OPERATOR AND CONFLICT RESOLUTION

Our approach to conflict resolution adopts the common approach that a specific chain of reasoning supporting a hypothesis should defeat

[2] In ATT-Meta, if and then conditions are expressed in an episode based logic (broadly similar in spirit to the logical scheme of Hobbs [6]).

a less specific chain of reasoning. For example, suppose that there is a defeasible rule stating that *students are untidy* and we know that *Ralph is a student*. We might conclude that *Ralph is untidy*. Now, suppose we know another fact, namely that *Ralph is middle-aged*. And suppose further that *middle-aged students are tidy* (IF middle-aged(r) AND student(r) THEN tidy(r)). We might then conclude that *Ralph is tidy*, since the fact that he is middle-aged and a student is more specific than the fact that he is a student.

This much is common. However, we also introduce an "even-though" exception handling operator. This does not appear to be paralleled by work elsewhere on uncertainty, and is the dual of the idea of expressing exceptions within general rules. Our approach has the exceptions indicating in a content-based way what general lines of reasoning they are exceptions to. Thus, we have a general reasoning rule about *birds* that they *can fly*, and we might state the following: *a penguin can't fly 'even-though' it is a bird*;. The effect of this "even-though" operator is to further specify the class of birds in an analogous manner to the further specification that being middle aged imposed on studenthood. Consequently, the fact that something is a bird and a penguin is more specific that the fact that it is a bird.

Note that under this approach, 'exceptionality' is located at the exceptions, not in modifications to general rules. In terms of system building this makes for a much more modular and convenient approach since the introduction of a new exception will not require the general rules to be modified.

In order to state the approach more formally, we employ the notion of sets of sets of paths, or pathset-sets, capturing the notion of specificity by requiring the less specific be a subset of the more specific. We shall first give a general definition of a path in order to introduce the notion and then discuss a modification which replaces equality between paths in our definition of subsets with subsumption.

Informally a path is a set of implicational links from a fact that grounds the competing hypotheses e.g. penguin(tweetie) upto, but not including, the hypothesis[3] e.g. fly(tweetie) or not-fly(tweetie). Here, we shall enclose paths between angled brackets. In the following discussion we shall ignore the **presumed** uncertainty-level. Also, we shall usually omit the predicate argument if it is clear. Thus, with the following rule and fact:

```
IF penguin(X) THEN NOT-fly(X); penguin(tweetie)
```

the path of links will be: <penguin>.

Often an argument for a particular hypothesis will depend on a conjunction of paths. Take the untidy student example. We cannot argue that Ralph is untidy if he is a student since he is a mature student and so tidy. The pathset supporting tidiness would thus be:

(<student>, <mature>).

In other words, a set of paths, or a pathset, is required when a conjunct is used in a rule. The pathset supporting the untidy hypothesis is the singleton set

(<student>).

Tidiness would win if we assume the following:

```
Pathset 'A' is more specific than pathset
'B' iff 'B' is a subset of 'A' and the
reverse is not true.
```

However, it would often be the case that there is one argument, and thus more than one pathset, supporting a hypothesis. Thus our algorithm needs to be broadened to include pathset-sets:

```
Pathset-set 'A' is more specific than
pathset-set 'B' iff for every member pathset
of 'B', the pathset is a subset of some
member pathset of 'A' and the reverse is not
true.
```

Let us now return to Tweetie. We can assume an inference from the fact that Tweetie is a penguin to the hypothesis that he cannot fly, via an inference rule that penguins cannot fly. However, another set of implication links might be that if something is a penguin, then it is a bird, and if something is a bird, then usually it can fly.

```
IF bird(X) THEN fly(X)
IF penguin(X) THEN bird(X)
penguin(tweetie).
```

We have an argument for 'NOT-fly' consisting of the pathset-set: ((< penguin >)), and an argument for 'fly' consisting of the pathset-set: ((<penguin bird>)).

Neither pathset is a subset of the other and so neither wins, which is intuitively incorrect.

It is to solve this problem, that we propose the new "Even-Though" operator, which for current purposes we shall assume is identical to the conjunct AND. Thus, we can add to the NOT-fly pathset-set which previous contained only the pathset that made use of the rule about penguins not flying, a pathset making use of the following NOT-fly rule:

```
IF bird(X) AND penguin(X) THEN NOT-fly(X)
```

(ie if X is a penguin then X will not fly even-though X is a bird.).

With this new rule adding the (<penguin bird>,<penguin>) pathset to the old ((<penguin>)) pathset-set, the pathset-set for NOT-fly is as follows:

((<penguin bird>,<penguin>), (<penguin>)).

Since the (<penguin bird>) pathset supporting the 'fly' hypothesis is a subset of the (<penguin bird>, <penguin>) pathset supporting the 'NOT-fly' hypothesis, but the reverse is not true, the 'NOT-fly' hypothesis wins.

Consider now a more complex case involving a subspecies of penguins that can fly: 'flenguins'.

```
IF flenguin(X) THEN penguin(X)
IF penguin(X) AND flenguin THEN fly(X)
IF penguin(X) THEN bird(X)
IF bird(X) THEN fly(X)
IF bird(X) AND penguin(X) THEN NOT-fly(X)
flenguin(tweetie)
```

The pathset-sets for fly and NOT-fly are now as follows:

Fly ((<fl p b>), (<fl p>, <fl>))
NOT-fly ((<fl p b>,<fl p>))

Let us ask whether fly is more specific than NOT-fly? The 2 element pathset (<fl p b>,<fl p>) from NOT-fly cannot be a

[3] Or the point at which paths to two plus hypotheses diverge.

subset of the first, single element (<fl p b>), pathset from the fly pathset-set. And nor is it identical to the second, 2 element pathset (<fl p>, <fl>) from the fly pathset-set. Consequently fly cannot win. This seems to be the incorrect result.

Is NOT-fly, then, more specific than fly? The first pathset of fly: (<fl p b>) is a subset of (<fl p b>,<fl p>). However, the second pathset (<fl p>, <fl>) is not a subset of (<fl p b>,<fl p>). Hence, NOT-fly does not win either.

Suppose we take a step back and consider why a winning pathset wins. It wins if it is more specific than the alternative i.e if the loser subsumes the winner. Consider now the two pathsets (<fl p>, <fl>) and (<fl p b>,<fl p>) taken from fly and NOT-fly respectively. In the <fl p b> path, birds are less specific than penguins, which are in turn less specific than flenguins. Consequently, the path <fl p b> subsumes the path <fl p>.

Now reconsider the pathset (<fl p b>, <fl p>) from NOT-fly that failed to match (<fl p>,<fl>) from fly when matching required identity. Suppose we say that if a path subsumes another path, then it matches it for the purposes of determining subsets. Under the new definition it is now a subset of (<fl p>,<fl>), since <fl p b> subsumes <fl p>, and <fl p> subsumes <fl>.

Note that the reverse cannot also be true ie (<fl p b>, <fl p>) is not more specific than (<fl p>,<fl>), so NOT-fly is not more specific than fly.

Space precludes further examples such as 'ill-flenguins' that cannot fly. However, it can be demonstrated that the new algorithm makes the correct predictions.

5 CONCLUSION

The approach described in the previous section is not particular to metaphorical reasoning, and was designed solely to deal with conflict resolution between conflicting hypotheses. However, it has turned out to have repercussions beyond what it was designed for and is highly appropriate to our view of metaphor reasoning. Metaphors might be viewed as specific ways of viewing something. ATT-Meta takes source domain facts to be more specific than target domain facts, and makes implicit use of the 'Even-Though' operator when introducing source domain facts. In other words, a source domain fact holds Even-Though there are target domain facts that might be thought of as contradictory. And this is important, since we assume that all rules and facts can in principle be used in any reasoning space. Without some means of protecting, for example, the source-domain facts that Anne's mind is a physical space, inferences about it are likely to be defeated by the knowledge that Anne's mind is not actually a physical space.

More generally, we would argue that specificity is a powerful tool for comparing arguments. Our path-based approach approach involves the complex examination of complete argument structures supporting competing hypotheses.

Finally, with respect to metaphor theory, our approach makes source/target conflict adjudication a complex argument-based matter rather than on some simple principles such as the 'Invariance Principle' [8] that target inferences should always defeat inferences based on the source interpretation.

ACKNOWLEDGEMENTS

We would like to thank the referees for their comments which helped improve this paper.

REFERENCES

[1] J.A. Barnden, S.R. Glasbey, M.G. Lee, and A.M Wallington, 'Reasoning in Metaphor Understanding: The ATT-Meta Approach and System', *Proceedings of the 19th International Conference on Computational Linguistics (COLING)*, (2002).

[2] J.A. Barnden, S.R. Glasbey, M.G. Lee, and A.M Wallington, *Application of the ATT-Meta metaphor-understanding system to examples of the metaphorical view of TEACHERS AS MIDWIVES*, Technical Report CSRP-02-10, School of Computer Science, The University of Birmingham, U.K, 2002.

[3] B. Falkenhainer, K.D. Forbus, and D. Gentner, 'The Structure-Mapping Engine: algorithm and examples', *Artificial Intelligence*, **41 (1)**, 1–63, (1989).

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

[5] S. Glucksberg, *Understanding Figurative Language: From Metaphors to Idioms*, Oxford University Press, Oxford, 2001.

[6] J.R. Hobbs *Literature and cognition* CSLI Lecture Notes, **21**, Center for the Study of Language and Information, Stanford University, Stanford, 1990.

[7] Z. Kovecses, *Metaphor: A Practical Introduction*, Oxford University Press, Oxford, 2002.

[8] G. Lakoff, 'The Invariance Hypothesis: is abstract reason based on image-schemas?', *Cognitive Linguistics* **1-1**, 3–74, 1990.

[9] G. Lakoff and M. Johnson, *Metaphors we live by*, University of Chicago Press, Chicago, 2nd edn., 2003.

[10] G. Lakoff and M. Turner, *More than cool reason: a field guide to poetic metaphor*, University of Chicago Press, Chicago, 1989.

[11] S. Parsons, *Qualitative Methods for reasoning under Uncertainty*, MIT Press, Cambridge, Mass, 2001.

[12] J.L. Pollock, 'Defeasible reasoning', *Cognitive Science* **11**, 481–51, 1987.

[13] S. Russell and P. Norvig, *Artificial Intelligence: a Modern Approach*, Prentice Hall, New Jersey, 2002

[14] T. Veale, and M. Keane, 'The competence of sub-optimal theories of structure mapping on hard analogies', *Proceedings of the 15th International Joint Conference on Artificial Intelligence*, Morgan Kaufmann, Los Altos, (1997).