Understanding Open-Ended Usages of Familiar Conceptual Metaphors: An Approach and Artificial Intelligence System

(Technical Report CSRP–01–05)

John A. Barnden and Mark G. Lee
School of Computer Science, The University of Birmingham
Birmingham, B15 2TT, United Kingdom
J.A.Barnden@cs.bham.ac.uk

Abstract

We present and evaluate an approach to the reasoning needed to handle a broad class of metaphorical utterances, and a computer program (ATT-Meta) partially implementing and further specifying that approach. The approach emanates from artificial intelligence but is offered also for consideration by cognitive scientists generally. The utterances of interest are ones that (a) rest on conceptual metaphors that are familiar to the understander but (b) transcend the mappings in the conceptual metaphors by using concepts not handled by the mappings. Our approach advocates possibly-extensive inferencing in the terms of the source (vehicle) domains of the conceptual metaphors, while avoiding as far as possible the extension of the mappings to deal with the concepts they do not handle. The general approach is similar in flavor to those of a small number of other metaphor researchers, but we provide a more extensive analysis, additional principles and a more thorough-going implementation. The approach contains a number of "view-neutral mapping adjuncts," which are default mapping principles that enable important source-domain aspects to be mapped to the target domain, independently of which specific metaphorical views are in play. Many discussions of metaphor appear to assume that such mapping actions occur, but rarely address them systematically and explicitly. In addition, in the approach, a conceptual metaphor can consist not only of a between-domain mapping but also of special, ancillary assumptions that serve to enrich the source domain with specific details needed by the metaphors. The implemented system supports ancillary assumptions but currently only has a preliminary handling of view-neutral mapping adjuncts.

1 Introduction

Lakoff (1994:p.212) gives the following example of a metaphorical utterance, attributed to a song lyric:

(1) We're driving in the fast lane on the freeway of love.

Our main interest is actually in mundane discourse such as ordinary conversation, popular magazine articles, newspaper reports, popular novels, etc., rather than songs, poems and other literary genres, but the example serves as a useful introduction to some issues. The example is a use of the familiar metaphorical view (familiar conceptual metaphor) of LOVE AS JOURNEY. The progress of the love relationship is viewed as a journey, the lovers are viewed as people undertaking the journey, and the love relationship as the vehicle in which they are traveling. To put it another way, the journey is stipulated to map to the relationship progress, the travelers to the lovers, and the vehicle to the relationship.

However, the view does not stipulate any mapping for the notions of driving, freeway or fast lane. Moreover, Lakoff (*ibid.*) does not mention any need for the understander of the utterance to *extend* the already-existing mapping (above) to cater for those notions. Lakoff states that the understander can reason, *within the terms of the source domain*, that the journey alluded in the utterance is exciting, using the source-domain knowledge that driving in the fast lane of a freeway is exciting. The excitement in the source domain is for the travelers. This maps, apparently, to excitement for the lovers, in the target domain. We discuss below the question of how this mapping act occurs, given that it is not an aspect of the LOVE AS JOURNEY view as described by Lakoff. The suggested overall process is a fairly simple application of a general approach that we advocate in this paper. The approach is characterized largely by its *de-emphasis of map extensions* and its *added, compensatory emphasis on within-source-domain reasoning.*

Lakoff does not present that approach in any explicit and general way, does not tease apart the fundamental principles and assumptions that are involved, and does not present arguments for it. The task of this paper is, partly, to correct these deficiencies. Also, strangely, Lakoff (1993), a work very similar to Lakoff (1994), does state (p.211) for the above freeway example that the "knowledge structures associated with" the words "freeway" and "fast lane" *are* mapped for the purposes of the particular utterance. Therefore, it is not clear, after all, that Lakoff's imagined and only vaguely described understanding process conforms to our approach!

Our approach is directed at metaphorical utterances that, like the freeway-of-love example, use metaphorical views that are stored in some form within the memory of understanders but where the utterances may be *map-transcending* in involving particular source-domain notions *not* directly mapped by those views, such as the notions of freeway and fast lane. Notice that map-transcendence does not of itself imply the the wording in the utterance is novel to the understander or particularly creative.

While we are on wording, we stress that we pay relatively little attention to "A is B" metaphors. See, e.g., Cameron (1999a) on the misleading stress traditionally placed on this particular syntactic clothing of metaphor.

As well as presenting a general approach, we will be presenting in this article an implemented AI system, ATT-Meta, that conforms to the approach. But the *general approach* itself is neither a psychological theory nor a specific computational model. Rather, it is a body of principles and stances that could be realized in many different specific psychological theories or computational models. We will suggest by example and general considerations that the approach is desirable and perhaps necessary, but by developing *a particular computational model* (ATT-Meta) we hope (a) to put the general approach more firmly to the test, (b) to make it plausible that the approach could indeed be realized as a testable psychological model or a viable computational model, and (c) to uncover problems and issues that might otherwise escape notice.[1] On the other hand, not only is the ATT-Meta system merely one possible specific realization of the approach, it is also only a partial realization of it. Indeed, we will be explicitly showing where ATT-Meta falls short of the full approach. And the general approach itself is not a complete answer to all problems concerning map-transcending metaphorical utterances.

The distinction between approach and system is not to be confused with Cameron's (1999a) theory-level/processing-level distinction. Rather, the approach is a somewhat unspecified processing-level account informed by some theory-level assumptions, and the system is a much more detailed, although therefore more limited, processing-level account.

Dealing with map-transcending utterances is difficult, because it seems inconceivable that, in general, an understander would possess a mapping for more than a minority of the notions in the source domain of any given metaphorical view, if the source domain is a commonsensical, familiar one,[2] and there seems to be no identifiable limit to metaphorically applying the language we use for talking about journeys to the love-relationship domain. For instance, *any* way of indirectly implying in a LOVE AS JOURNEY utterance that the journey is exciting would lead to the above conclusion that the lovers are excited. This productive use of language from one domain to talk about another is well known (Carbonell, 1982; Gibbs 1998; Lakoff 1993; Lakoff & Turner 1989; Lakoff & Johnson 1980; Martin 1990, 1994; Reddy 1979/1993; but see Croft 1998 and Gibbs 1998 on the difficulty of choosing between possibilities for accounting for the productivity). In the area of corpus studies of metaphor, Deignan (1999) has said that "speakers regularly exploit and extend existing metaphors as a way of creating new meanings" and "It seems that virtually any source domain collocate of a word that has an established metaphorical sense can be extended metaphorically into the same target domain."[3] None of this is to say that there is no curb at all on the productivity of metaphor, and limitations are often pointed out. However, Clausner & Croft (1997) and Grady (1997) argue compellingly that by suitably refining one's theoretical account of what metaphorical views are actually involved in utterances, one can go a long way towards eliminating apparent productivity gaps.

[1] ATT-Meta is not a complete natural language understanding system, but is an AI system in that some version of it, or some version of the algorithms in it, could eventually form part of a complete system. Of course, AI is far from achieving such understanding, for many reasons other than the problem of metaphor.

[2] Of course a minority of a large set can itself be a large set.

[3] We do not agree with the assumption of map-extension here, but only with the observation that the collocates are used metaphorically.

Map-transcending utterances form a very important subclass of all metaphorical utterances, as is suggested by the range of examples in this article and accompanying technical reports (Barnden 2001a,b; Lee & Barnden 2001b), and by the claims of Carbonell (1982), Lakoff & Turner (1989) and Martin (1990). But aside from some partial approaches in AI, metaphor research has not focussed in algorithmic detail on how map-transcendent metaphorical utterances are to be handled. Notably, map-transcendence is not addressed in writings on the dominant approaches in Psychology such as Structure Mapping Theory (Falkenhainer *et al.* 1989; Gentner 1983), ACME (Holyoak & Thagard 1989; Holyoak *et al.*, 1994), IAM (Keane 1988), feature-transfer approaches (e.g., Ortony, 1979); and the categorization approach of Glucksberg & Keysar (1990). All of these except the categorization approach are mostly to do with *creating metaphorical mappings (or similarly feature-transfers) from scratch*, without attention to any already-stored metaphorical views.

Map-transcendence overlaps notions that other authors have identified as important in the study of metaphor. For instance, it overlaps the notion of extension of metaphors (e.g.: Martin 1990; Deignan *loc. cit*), the notion of an utterance using "unused parts" of a source domain (Lakoff & Johnson 1980), and the notion of "creative" or "novel" use of metaphorical views (many authors, incl. Lakoff & Turner 1989), "modification" of metaphors (Goatly 1997), "exploitation" (Moon 1998), and use of "variants" of fixed expressions such as metaphorical idioms (Moon 1998).

In seeking to *avoid* the extension of known mappings, our approach is largely opposite in a crucial respect to psychological models such as SME (Falkenhainer *et al.* 1989) and ACME (Holyoak & Thagard, 1989) and to some salient AI models (including Fass 1997; Iverson & Helmreich 1992; Sun 1995; Veale & Keane 1997). These psychological and AI models are based largely on *creating an analogical mapping* between the source domain and the target domain, by means of a *search for partial isomorphism* between the domains. Partial-isomorphism search for the purposes of map-creation, and similarly map-extension, is highly complex and computation-intensive (see, e.g., Falkenhainer *et al.* 1989). We conjecture also that our avoidance tendency contributes to psychological plausibility, as we will discuss below. The tendency is broadly similar to the selective-projection stance of Fauconnier & Turner (1998). However, that work does not provide a precise method for selection of what to project from one space to another.

As for source-domain reasoning, most previous work on metaphor understanding has failed to appreciate its importance, possible extent, and detailed nature. Nevertheless, some authors have addressed it to some degree (notably: Hobbs 1990; Martin 1990; Narayanan 1997). Also, going by some brief remarks (e.g.: Carbonell 1982; Grady 1997: p.283; Lakoff 1994:p.212/3; Lakoff & Johnson on "unsused" portions of metaphors) some other authors are presumably adopting something like our map-extension minimization and reliance on source-domain reasoning. However, we develop the idea more extensively and make it a central tenet. Existing metaphor-processing accounts such as Structure Mapping Theory (SMT) seem compatible with the addition of source domain reasoning (and indeed one can only assume that some of the assumed structure in source analogues in reported applications of SMT would arise from source-domain reasoning rather than being stored in memory) but compatibility is a far cry from actual inclusion and from

coping with the ramifications of the inclusion.

The ATT-Meta system works on the basis, largely, of *user-supplied* knowledge rules that are about the domains involved in particular metaphorical views or that encapsulate mappings between source and target domains in particular metaphorical views. It is *not* our aim to propose any particular rules for any particular domain, or to commit firmly to any particular mapping rules (although we have views on the content of mapping rules). Rather, our main focus is on the high-level processing strategies involved, the amount of on-line mapping work needed in metaphor processing, and the amount of source-domain reasoning needed. Thus, any particular application of ATT-Meta should be interpreted as follows: *if* such-and-such rules are included *then* ATT-Meta can do so-and-so. Another proviso is that it is not part of our current research to determine how domain rules or mapping rules could be acquired by a metaphor-understanding agent during its life.

The plan of the paper is as follows. Section 2 presents most of the main principles of the advocated approach. Sections 3, 4 and 5 report applications of the approach to a wide variety of metaphor examples from the research literature and from real discourse. Section 6 resumes the description of the approach by presenting its emphasis on query-directed reasoning as a way of making metaphor interpretation responsive to discourse context. Section 7 is an introduction to the implementation of the approach in ATT-Meta. Section 8 illustrates ATT-Meta in action by describing its application to a complex real-discourse example from Section 5. Section 9 evaluates the system by comparing it to the approach and discussing its performance on examples. Section 10 evaluates the general approach in the light of the previous sections. It summarizes the coverage of metaphor phenomena by the approach and how well the approach copes with examples. It also shows how the approach deals with problems faced by other approaches and addresses the question of the role of literal meanings of metaphorical utterances in the computation of their metaphorical meanings. Our approach does give an important role to literal meanings for the type of metaphorical utterance in question, although the approach is not one that relies on prior rejection of a literal meaning. Section 11 presents further discussion and Section 12 concludes.

The general approach and/or ATT-Meta system have been described elsewhere (e.g.: Barnden 1992, 1998a,b; Barnden *et al.* 1994a,b, 1996; Barnden & Lee 1999; Lee & Barnden 2001a), but the present paper presents a more extensive, up-to-date account, with important new claims.

2 Main Notions and Principles of the Advocated Approach

2.1 Semantic Account Assumed

We take a roughly traditional stance to the basic way in which utterances have meanings. We assume that words, phrases, and other surface items can have "direct" senses (or, more precisely, internal sense representations) supplied by some lexicon internal to a language-using agent. We take no stance on whether such senses are therefore to be called "literal" or not. Rather, direct senses are simply senses that are

particularly directly and perhaps automatically accessible by the agent from the lexical items. Also, we make no appeal to the idea that some direct senses for a given surface item are more central or primary than others.

A well-formed linguistic expression has one or more *direct meanings* compositionally derivable from the individual direct meanings, direct senses, and referents of the expression's components. We do not comment in this paper on the difficult problem of how pronouns, demonstratives, etc. contribute to the direct meaning. A direct meaning could be called a *literal meaning* by a particular theorist if the individual senses/referents used to construct the meaning happen to be classed by that theorist as literal.

2.2 Metaphorical Views and Mental Storage Thereof

Metaphor is a hard thing to define and various theorists as well as ourselves take the view that it is not a precisely definable concept, being somewhat like the concept of "game" in a Wittgensteinian spirit (Cameron 1999b:p.106). Also, what a researcher takes metaphor to be can depend on research aims (Cameron 1999a).

For our purposes, using the word "agent" to mean someone producing or understanding discourse, a metaphorical view consists at least partly of a source-to-target mapping that can translate agent-internal (mental) representations of particular source-domain aspects—entities, properties, relations, states of affairs, etc.—into agent-internal representations of particular target-domain aspects, and perhaps also vice versa (see below). The mapping typically takes the form of a between-domain structured analogy of the sort described in, say, Falkenhainer *et al.* (1989). The domains are ones the agent perceives as being (at least partially) incongruous (following Cameron 1999b). More precisely, the domains involve, in part at least, qualitatively different individuals, properties, relationships, etc., as far as the agent is concerned.

We prefer the term "metaphorical view" to Lakoff's term "conceptual metaphor." We take no stance in this article on the idea that metaphorical views are an important part of the very *concepts* that people have for thinking the metaphors' target domains (see, e.g., Murphy 1996). Also, we do not assume Lakoff's Invariance Hypothesis (Lakoff 1990), and do not use the notion of an image-schema.

A given metaphorical view, or more precisely the mapping it involves, may be *stored* in the memory of a particular cognitive agent, albeit perhaps in some implicit, distributed fashion (cf. section 2.3). We also say that the agent *possesses* the view. Intuitively, when the view is stored, we can say the agent already knows the mapping, perhaps unconsciously. Unstored views/mappings used by an utterance may be novel to the agent, but equally the agent may have met the view before but not have stored it. We focus in this article entirely on metaphorical views that are stored in the understander of the utterances in question. But, importantly, we are concerned with cases where some or all of the use of source-domain concepts by the utterances transcend the mappings in the stored views.

The mapping in a view generally consists of several individual *mapping relationships* dealing with different aspects of the source and target domains. For instance, in LOVE AS JOURNEY one mapping relationship maps travelers to lovers, and another maps journeys to love relationships.

2.3 Potential Irreality of Domains and Metaphorical Views

Although we have mentioned "metaphorical views" and "domains," these are a discussive device and are not assumed to have individuated reality in our approach (although they *could* be real in a given instance of the approach). We only assume that what is stored in the understander are individual mapping relationships, not whole views. We do not assume that the user explicitly categorizes knowledge, lexical senses, etc. in terms of domains. Our talk of "domains" is just a loose way of alluding to two sets of properties, relationships, etc. related by a set of mapping relationships. These sets are not given as such in advance—it is a subjective and context-sensitive matter whether we wish, say, to talk about foxes and people (cf. "John is a fox") as being in separate domains or not. Cameron (1999a) makes a similar point.

Even if domains could be clearly and adequately individuated, the source domain and target domain in a metaphorical view would often overlap in major ways. In the freeway-of-love example, people, together with many of their emotions, beliefs, properties and relationships, are aspects both of the domain of journeys and the domain of love relationships. A similar point is clear also in the Small-Voice example to be discussed below, and Kittay (1989) makes similar observations.

2.4 Metaphoricity and its Agent-Relativity

We take a highly agent-relative stance to all things metaphorical (cf.: Cameron 1999b; Fass 1997:p.50; Goatly 1997:p.110; and Radman 1997). For instance, there is no objective fact as to whether the utterance is metaphorical or not, or what particular views are involved, or whether two domains are incongruous with each other and can therefore be the target and source domains in a metaphorical view, without reference to individual understanders. In the area of metaphorical-utterance understanding what is of primary importance is the metaphorical views that *that particular agent* possesses or is in a position to create from scratch. Of course, with too great a degree of difference between agents on this, communication could suffer, but this is not a feature of metaphor specifically, as agents may differ on many other things that could affect communication.

One type of relativity rests on what direct senses understanders happen to have in their lexicons. Consider an utterance of "*McEnroe killed Connors,*" taken by an understander U to claim that McEnroe caused Connors to lose in a tennis match. Suppose that U possesses a metaphorical view of SPORT AS PHYSICAL COMBAT, that maps dying to losing in a game, and that maps causation identically. Suppose that in U's lexicon the only direct sense for the word "kill" is the PHYSICAL COMBAT sense of cause-death. Then U can work out that McEnroe made Connors lose, using that source-domain sense and the metaphorical mapping. For U, the utterance is a use of the metaphorical view SPORTS AS PHYSICAL COMBAT.

But another understander V may also have the direct target-domain sense cause-to-lose-in-game for the word "kill" in her lexicon. The utterance is nevertheless, for V, still a use of the metaphorical view, if V possesses the view, because the above route via the source-domain sense and the metaphorical mapping is still *available.* We could say the metaphoricity for V is "sidelined" for V by V's direct target-domain sense.

(Other researchers have used closely related notions of "inactive," "dormant" or even "dead" metaphor but these terms are variously used, very vaguely defined, and potentially misleading — see also Radman 1997.) For simplicity, in our examples below we will normally assume that the metaphoricity is not sidelined.

For an understander W who did not possess, and does not now create, the SPORT AS PHYSICAL COMBAT view, the utterance is of course not a use of that view. It may not be metaphorical at all for W: W may have the cause-to-lose-in-game and cause-death direct senses for "kill," and possibly various other non-sport direct senses, but no potential route to a cause-to-lose-in-game conclusion from the non-sport direct senses. The metaphoricity of "kill" for *other* understanders is simply absent for W.

2.5 Informational Contributions and Metaphorical Meaning

In this article, an *[informational] contribution* of an utterance, whether metaphorical or not, that is about some domain D is simply some claim, query, instruction, etc. concerning D that is drawn in part by processing the utterance. Such processing can include the application of metaphorical mappings. So, in the McEnroe/Connors example the claim that McEnroe caused Connors to lose is a contribution in the target domain. The purpose of our approach is to account for the creation of internal representations, within a given understander, of target-domain informational contributions that metaphorical utterances may have for that understander.[4]

Conclusions drawn from a target-domain contribution by further inference in that domain could also count as contributions of the utterance. We do not need to draw any particular boundary around the notion of contribution. But in this article we will mostly be considering target-domain contributions that arise immediately. or almost so. from metaphorical mapping acts.

In discussions of metaphor, putative "metaphorical meanings" of metaphorical utterances are often mentioned. The metaphorical meaning of "John is a fox" in some context might be claimed to be the contribution that John is cunning. But in our approach we do not commit to any definition of which particular subcollection of the target-domain contributions an understander draws from a metaphorical utterance constitutes the "metaphorical meaning," if indeed any does. (This is much more neutral than the stance of Davidson 1979, and Cooper, 1986, who positively claim that metaphorical utterances have no meanings other than their literal meanings.)

2.6 Ancillary Assumptions

A metaphorical view can be (loose talk for) more than just a set of mapping relationships. There can be what we will call "ancillary assumptions" that constrain the nature of the imaginary situations in the source domain. For instance, Lakoff (*ibid.*, at least) seems to assume that the LOVE AS JOURNEY view involves the travelers as being in a vehicle, and as being, by default, alone in it. But of course this is

[4] Elsewhere we have referred to informational contributions as "connotations," but that term may have inappropriate connotations for some readers.

not a necessary feature of a journey. The view could be cumbersomely renamed as LOVE AS JOURNEY WHERE THE TRAVELERS ARE ALONE IN A VEHICLE, but that would not change the basic point that more is involved than a mapping between journey terms and love terms.

2.7 Direction of Mapping

Metaphorical mapping acts—that is, the use of mapping relationships to actually map items from one domain to another—are almost always talked of, in practice, as going from the source domain to the target domain (despite the point made by the interaction theory of metaphor that the source domain can be affected by metaphorical talk: Waggonner 1990). However, we feel this is misguided, as explained more fully in Barnden (in press, b).

Consider the metaphorical view of IDEAS AS PHYSICAL OBJECTS. First, if this view is being used in a discussion of Mary's mental states, then *any* idea of Mary's that starts being mentioned in the discussion can by default be assumed to be being cast as a physical object. This is a mapping act *from target domain to source domain*, not the other way round. Secondly, source-to-target mapping acts, from physical objects to ideas, are more constrained than the target-to-source ones. In "Mary buried the idea in a dark recess of her mind," the physical material in which the recess lies should presumably not itself be mapped to a set of ideas. So, there is no *generally applicable* mapping from the property of being a physical object to the property of being an idea, whereas there is a generally applicable map the other way round.

Thus, it is more appropriate to talk of mapping acts from ideas to physical objects than the other way round. Of course, there must ultimately be *some* source-to-target mapping act that creates information in the target domain. It is just that not all mapping acts performed in the understanding process need be in that direction. In our account below, metaphorical understanding can include not only source-to-target mapping acts but also target-to-source ones, although we will not be dwelling on the point.

2.8 Metaphorical Reasoning Subsumed under Pretence

Our approach encapsulates the proposed source-domain reasoning within a *pretence cocoon.* The cocoon is a special reasoning context, manipulation of which is tantamount to the reasoner pretending to take the source-domain scenario alluded to in the utterance as a true statement. The propositions inside the pretence cocoon can be thought of as forming a description of an imaginary, pretend world. Propositions outside are about reality as the understander sees it, and will in particular be about the target domain of the metaphorical utterance in question. Metaphorical mapping acts operate between the inside of the pretence cocoon and outside.

Thus, for the freeway-of-love example, the reasoner pretends, within the cocoon, that the the lovers *really are* driving along a freeway. Consequences of this, using knowledge about freeway driving, can be worked out in the pretence cocoon, and *some* consequences may (hopefully) then be able to be transformed, via the metaphorical mapping, to become propositions about the love relationship that are outside

the cocoon. For the reasoning within the cocoon we will use the terms "within-pretence reasoning" and "within-cocoon reasoning" interchangeably.

We stress that when a pretence cocoon is created, it is not tagged as being to do with any particular metaphorical view. It is only implicitly specialized towards some set of views by having the utterance's direct meaning placed within it. If within-cocoon reasoning concludes, for example, that someone's mind is a physical region (because, say, the mind has "dark recesses"), *then* it is becoming apparent that the particular metaphorical view MIND AS PHYSICAL SPACE is being used. The information in the cocoon that the agent's mind is a physical space may then allow particular mapping relationships to come into play. These mapping relationships can theoretically be cast as being part of that view, but they are not necessarily identified by the understander as being so. (Recall the irreality of views noted in section 2.3.)

Our use of pretence corresponds to the proposal by Levin (1993) in literary theory that metaphor sets up an imaginary world. Kennedy & Chiappe (1999) say that a metaphorical statement such as that something is a banana "allows us to pretend that [the thing] has all the properties of a banana." Our pretence approach can also be seen as a special case of the blending idea in cognitive linguistics (Fauconnier & Turner 1998), in that the pretence cocoon acts a blending space melding together to some extent the domain of love relationships with that of journeys. There is a weak connection to Hobbs also, in that in his treatment of conventional metaphor (Hobbs 1990) he applies, much as we do, source-domain properties and relationships directly to target-domain entities, although he does not make use of cocoons.

Within-pretence reasoning is similar to the *simulative reasoning* used or proposed by many researchers for reasoning about agents' beliefs and reasoning (see, e.g., Carruthers & Smith 1996, Creary 1979, Davies & Stone 1995). Our use of pretence is roughly tantamount to simulating the reasoning of a misguided agent who believes the metaphorical utterance to be literally true.

In general there can be more than one pretence cocoon (e.g., because of metaphor compounding, as explained in Lee & Barnden 2001a), but in this article we will assume just one.

2.9 Any Knowledge Usable within the Pretence

The approach allows *any* knowledge that the understander possesses to be used inside the pretence, not just source-domain knowledge, because the approach does not subscribe to any explicit or precise delineation of domains in the first place (recall section 2.3). However, in practice, most of the reasoning within the pretence cocoon is based on source-domain knowledge, with some reasoning steps sometimes being supplied by ancillary assumptions. This is so at least for the examples discussed in this paper. Thus the notions of within-pretence reasoning and within-source-domain reasoning are for the purposes of this paper approximately interchangeable.

Because any knowledge is usable in the cocoon, there is a danger that real-world knowledge that contradicts the information that is postulated or inferred in the pretence cocoon could interfere with the reasoning there. Moreover, even knowledge that could intuitively be regarded as pertaining to the source domain can

cause interference within a pretence. For example, in a usage of MIND AS PHYSICAL SPACE, the hypothesis that some particular person's mind is a physical space is likely to arise in the pretence cocoon. But, it is reasonable to suppose that the understander believes or can infer that minds are not, in reality, physical spaces. This information can be used in the pretence and will therefore cause a conflict. The approach to this taken in our approach is to include a special, but very generally framed, provision in the handling of reasoning conflicts. This provision says, intuitively, that the more that an argument within the pretence relies on particular information placed within the pretence, as opposed to information arising from knowledge of reality, the stronger it is.

The knowledge that minds are not physical spaces arguably pertains to the domain of physical objects and spaces as well as to the domain of mind. On this assumption, we have seen that even source domain knowledge may be defeated within a pretence. If the knowledge is not regarded as pertaining to the physical domain, then the example shows one way in which non-source-domain knowledge may be used in the pretence, albeit in this case defeated there.

2.10 Identities Rather Than (Some) Correspondences

We have already noted that the source and target domains in metaphor can overlap, in an particular can involve a set of shared individual entities, such as the people in the freeway-of-love example. This observation, coupled with our use of pretence, leads us to allow individual entities to be shared between the reasoning within the pretence and the reasoning outside concerning reality. Thus, some of the mapping between the pretence and reality actually consists of identity relationships between individual entities inside and individual entities outside. We consider identity to be a special case of mapping. A shared individual entity's properties and relationships within the pretence will of course usually conflict, in part, with those the entity has in reality.

The possibility of sharing does not prevent the pretence including individuals not appearing in reality and vice versa. In some cases, there may be non-identity mapping relationships linking within-pretence entities with within-reality entities.

2.11 Uniformity and Mixedness of Reasoning

In our approach, reasoning of the various different types—within-pretence, outside-pretence, mapping actions—can be freely interleaved or carried out in parallel. At one moment the understander can be pursing a reasoning line within the pretence, only to switch to doing some mapping and some outside-pretence reasoning, and then to switch back. The three types of reasoning are not packaged into different stages.

The three types are not to be thought of as intrinsically different in their detailed operation. They can all be handled, for instance, by the application of rules of a uniform shape, as is the case in the ATT-Meta system. Moreover, we stress that the source-domain reasoning done within a pretence is just like the

reasoning that really might be needed for reasoning about a situation that a discourse could describe in the source domain, such as a case of people driving on a real freeway.

2.12 View-Neutral Mapping Adjuncts

Carbonell (1982) proposed that certain aspects of sources tend to map over to targets, irrespective of what the particular metaphorical view at hand is. Roughly following his line, although departing significantly from the specifics of his suggestion, we assume at least the following *view-neutral mapping adjuncts* (VNMAs). They are defaults that can in principle be blocked by particular metaphorical views or by incompatibility with the understander's knowledge of the target. We call them mapping "adjuncts" and regard them as standard but implicit aspects of all view-specific metaphorical mappings. In the descriptions below we use the word "mappee" of a source-domain item to mean the target-domain item to which it maps. [5] Recalling section 2.10, the mappee may be the same as the mapped source-domain entity (though with somewhat different properties).

Causation/Ability VNMA: Causation, prevention, helping, ability and (dis)enablement relationships between events or other entities in the source map to causation, prevention, ... relationships between their mappees (if they have any).

Change VNMA: If there is a change event from one state of affairs to another in the source domain, where the states of affairs themselves have mappees in the target domain, then the change event has a mappee that is a change event between the latter states of affairs.

Time-order VNMA: The time-order of events in a source domain is the same as that of their mappee events, if any.

Duration VNMA: Qualitative length of time, in the context of the source domain, that is consumed by an event maps identically to qualitative length of time, in the context of the target domain, consumed by the mappee event, if any. For example, if something takes a long time in the context of the source domain then a mappee target event takes a long time in the context of the target. Also, qualitative duration comparisons map over.

Event-Shape VNMA: Aspectual features of events/situations/processes, such as whether they have a start or end, or are intermittent, map identically to mappee events/situations/processes.

Mental/Emotional States VNMA: If some agents in the source domain map to some agents in the target domain, then mental and emotional states of the agents map identically, except that their objects or propositional contents (if any) are modified suitably by any mapping relationships that apply, and provided that this modification can be done.

For example, if John and Mary arguing is metaphorically viewed as Harry and Sally (respectively) engaging in physical combat, then the source domain proposition that *Harry believes he is losing in the*

[5] In standard mathematical terminology it would be the "image" of the source item.

combat maps to the target domain proposition that *John believes he is losing the argument*, assuming that combat-losing maps to argument-losing.

The VNMA allows for object/content-free states. Thus, John's being happy in the source domain maps to John's being happy in the target domain.

Modality VNMA: Relative degree of necessity, possibility, obligation, tendency, etc. in the source domain, for actors to undertake actions or for a state of affairs to obtain maps identically to relative degree of necessity, possibility, etc. for mappee actors, if any, to undertake mappee actions, if any, or for mappee states of affairs to obtain, if any.

Value-Judgment VNMA: Levels of goodness, importance or other types of value assigned by the understander to states of affairs in the source domain map identically to levels of goodness, etc. of their mappee states of affairs, if any.

Uncertainty VNMA: The level of certainty with which situations hold in the source maps at least roughly to level of certainty with which their mappee situations, if any, hold.

Function VNMA: Functions (i.e., purposes served) of entities in the source domain map to functions of their mappees, if any.

Qualitative Degree VNMA: If the holding of a graded property or relationship in the source maps to the holding of a graded property or relationship in the target, or vice versa, then the qualitative absolute and relative degrees map over identically. For example, if presence of above-normal temperature maps to presence of anger, then a high temperature maps to intense anger, and the higher the temperature the more intense the anger.

Negation VNMA: If a property or relationship in the source maps to a property or relationship in the target, then non-possession of the source property/relationship maps to non-possession of the target property/relationship.

Set-hood VNMA: If entities of a certain type S in the source map to entities of type T in the target, then a set of entities of type S in the source maps to a set of entities of type T in the target.

Set-Size VNMA: Qualitative size (relative or absolute) of sets in the source maps identically to qualitative size of mappee sets in the target. E.g., if a set is large in the terms of the source domain then its mappee (if any) in the target domain is also large in the terms of that domain.

Physical-Size VNMA: Qualitative physical size (relative or absolute) of physical objects in the source maps identically to qualitative physical sizes of mappee objects (if physical) in the target.

The VNMAs used in examples in this paper are: Causation/Ability, Event-Shape, Mental/Emotional States, Value-Judgment, Uncertainty, Qualitative Degree and Negation.

It may in future be necessary to postulate other VNMAs, to generalize or combine some of them, or to arrange them in a precedence ordering as Carbonell does. In proposing the VNMAs we do not claim to be doing more than collecting and explicating some (generally tacit) assumptions of many metaphor

researchers. In some cases researchers explicitly set great store on relationships like the VNMAs above. Examples include the time-order/duration and event-shape VNMAs, which are strongly related to the mapping of event structure in Lakoff (1993), and the causation mapping in SME examples (Falkenhainer *et al.*, 1989), a primary special case of the general preservation of "higher-order" relationships in an attempt at systematicity. Also, since non-unary predicates map identically in SME as described in Falkenhainer *et al.* (1989), many (though not all) of the VNMAs would just fall out of that identical-mapping assumption. The reason we propose specific VNMAs, rather than proposing a general transfer principle of which many at least of the VNMAs would drop out as special applications, is that we wish to be conservative about what is proposed. It is a larger claim than *we* wish to make that higher-order properties or structure in general is mapped over.

Equally, it would be too bold without extensive further evidence to adopt a default that any source-domain property or relationship that makes sense without change in the target domain is to be mapped over identically. This issue is an important because in many cases the target and source domains in metaphor significantly overlap.

Going back to the freeway-of-love example, (1), Lakoff does not give (*loc. cit.*, nor as far as we know, elsewhere), an explanation of why the excitement maps over. In our approach it is mapped by the Mental/Emotions-States VNMA. Since the agents in the source domain (the occupants of the car) are indeed mapped (in fact, are identical to) agents in the target domain, namely the lovers, the emotion of excitement maps over identically. Suppose we take the excitement in the pretence to be about the whole freeway-driving situation encapsulated in the direct meaning of the utterance. This situation maps to some unanalyzed situation or other in the target domain. So the particular excitement in the pretence can be transferred to become the lovers' excitement about that target-domain situation.

2.13 Main Operations in Understanding

Figure 1 impressionistically depicts the main types of processing activity that we claim should take place in the understanding of map-transcending metaphorical utterances where the metaphoricity is not sidelined. The main types are:

- construction of the direct meaning of the utterance (recall section 2.1)
- placing of it in the pretence cocoon
- usually, performance of (predominantly) source-domain reasoning within the pretence cocoon, using that direct meaning together with general knowledge about the source domain
- possibly, application of ancillary assumptions to create further propositions within the pretence cocoon

- mapping acts, including at least one that goes from source to target; the acts can be applications of view-specific mapping relationships or VNMAs.

This listing does not imply any particular ordering of the operations mentioned, and indeed they are not thus ordered in ATT-Meta.

As exemplified by the freeway example, a crucial aspect of our approach is a default that goes in the diametrically opposite direction to map-creation:

> *Map-Extension Minimization:* By default, an extension of the mapping(s) in the stored metaphorical view(s) used in an utterance, to cover map-transcending aspects of the utterance, should *not* be sought. It should only be sought when there is a specific need for it posed by the discourse understanding task. (Also, we suspect that it may often be difficult or impossible to find a suitable, non-arbitrary, useful extension anyway.)

We stress that understanders are free to extend maps if they have reason to do so. It is just that we are trying to minimize the amount of extension our approach obliges them to do. We compensate for the reduced reliance on map-extension by

> *Reliance on On-The-Fly Source-Domain Reasoning,* which is reasoning conducted in the terms of the source domain, largely in an effort to inferentially *link* map-transcending source-domain notions used in the direct meaning of the utterance to source-domain notions for which there is already a mapping relationship, so that only an *existing* source-to-target mapping relationship is used.

The following is also involved:

> *Transfer-Warranting by Unmapped Structure:* in the case of map-transcending metaphorical utterances, the warrant for attempting to transfer a source-domain representation may include inferential linkage to [A] complex source-domain structures that arise from the utterances at hand and that are, typically, *not* themselves mapped to parallel structures in the target domain, and [B] general knowledge of the source domain that, again, is typically *not* mapped to the target.[6]

In the freeway-of-love example, aspect [A] applies because the propositions in the source domain that are transferred are warranted by inference from the freeway-driving situation, which is not mapped in any detailed, structured way to the target domain. Certainly, we can postulate that there is *something* the lovers

[6] By a transfer action we mean a mapping action that maps particular pieces of source-domain items that have not yet been subject to such an action in the ongoing metaphor understanding process.

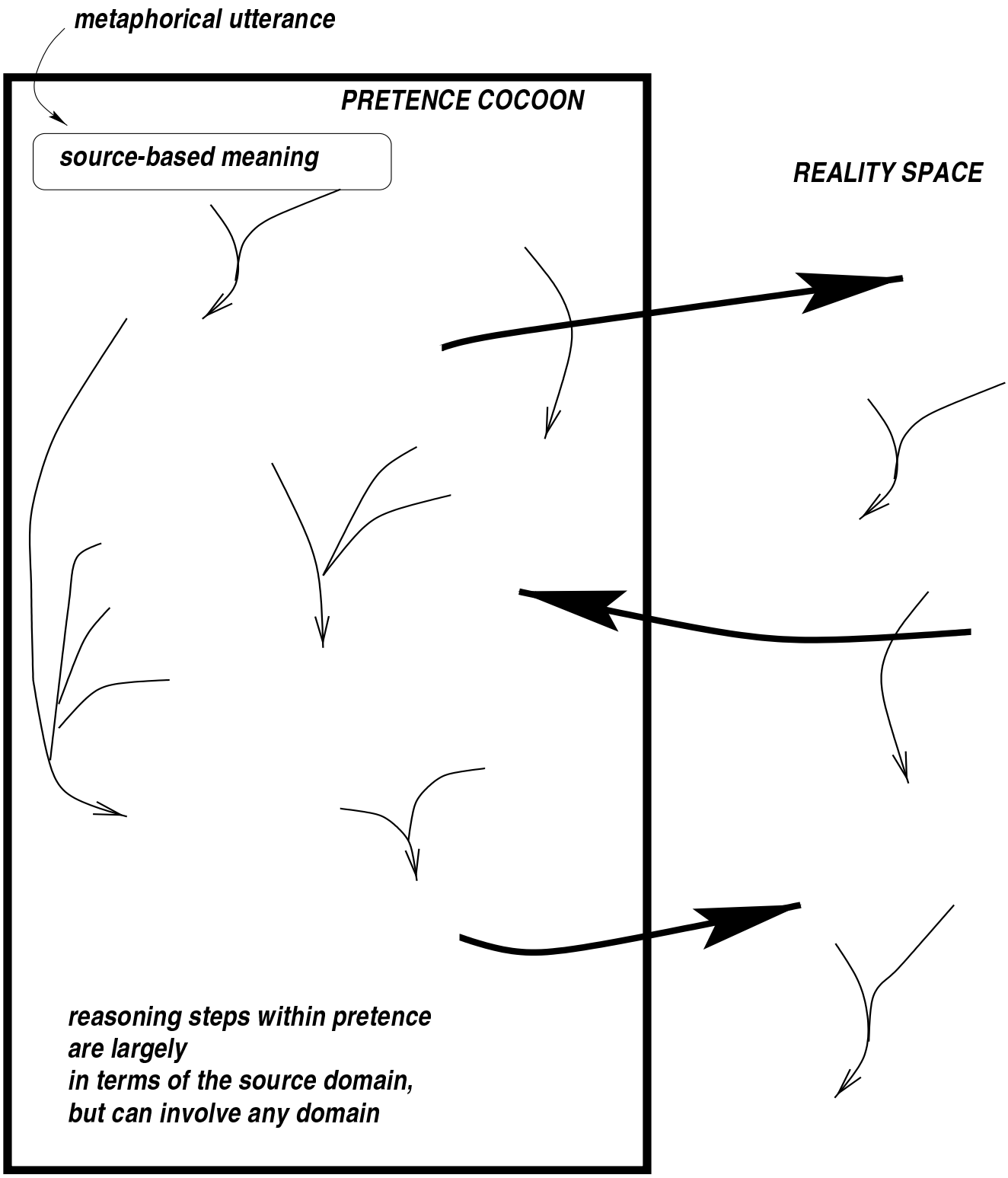

Figure 1: **Impressionistic depiction of reasoning in the general approach.** The heavily outlined box shows the pretence cocoon. The region outside is the top reasoning context in which the understander reasons about reality as he/she/it sees it. Bold arrows show the action of mapping relationships. Other arrows, apart from the one at the top left, show reasoning actions within reality or within the pretence. These steps can be the application of domain knowledge or (when inside the cocoon) of ancillary assumptions.

are doing that corresponds to driving in the fast lane. But that something is an unanalyzed entity. Aspect [B] applies because the source-domain reasoning rests in part on general knowledge about freeways and driving culture. By default (or so we must assume from Lakoff's comments), fast-lane driving is faster than driving in other freeway lanes, and freeway driving in relatively fast anyway. The speed leads to excitement. The general knowledge used here does not itself need to be mapped in any way at all but is crucial in Lakoff's interpretation of the utterance.

Our warranting principle contrasts with an important theme in much metaphor theory, including that based on analogy systems such as SME, ACME and IAM (Keane 1988), where the warrant for attempting to transfer source-domain information over to the target domain is its connectivity to other source-domain information that *does* have a detailed mapping to the target. For instance, in the water-flow/heat-flow analogy discussed in Falkenhainer *et al.* (1989), the warrant for transferring the proposition that the water flows (to become the target-domain proposition that the heat flows) is the detailed structural mapping already established between other aspects of the two situations. We are not against this type of warranting, but we draw attention to the fact that it is not enough by itself, and indeed may be of minor importance for map-transcending utterances based on stored metaphorical views.

3 Application to Examples Discussed for Neighbouring Approaches

In this section and the next we will show how our approach applies to various examples that have been discussed in metaphor research. They are examples for which the researchers have indicated particular informational contributions (claims, etc. about the target domain), and we will show how those independently-proposed contributions, and others, can be arrived at by means of our approach.

In the present section and later ones, our statements about what metaphorical views are involved in examples, what views are stored in the assumed understander, and what direct senses of words are possessed by understanders are our best guesses but are not firm claims of this paper. Our focus is in the style of processing that can be based on stored metaphorical views and direct senses, whatever particular ones they happen to be in a given understander at a given time.

In the present section we address approaches that are "neighbours" of ours in that they place some stress (however sketchily) on source-domain reasoning and/or refraining from constructing maps for map-transcending notions. We omit Lakoff's approach as we have already considered the freeway-of-love example.

3.1 Grady's House of Many Mansions

Grady (1997) appears to recommend some degree of map-extension minimization and certain forms of source-domain reasoning (couched by him as metonymic acts). But we disagree with his claim (*ibid.*, p.280) that additional metaphorical views (i.e., additional mappings) are needed to deal with the map-transcending

aspects of the following example. It is part of an excerpt, cited by Grady (1997), from a *New York Times Book Review* article:

(2) The house of psychoanalysis has many mansions, but some of Freud's followers ... have not wanted to live in the main house and have built their own annexes and outbuildings.

We assume that Grady is correct in saying that the example from which the above excerpt is taken rests on metaphorical views of ORGANIZATION AS PHYSICAL STRUCTURE, the organizations in question being various forms of psychoanalytic thought-system (or theory), and ESPOUSED THOUGHT-SYSTEM AS HOME (our name for the view). Under the latter view, a body of ideas, an artistic style, etc. can be viewed as a person's home. Hence, having one of the dwellings in the example as home maps to espousing the corresponding thought-system.

The reasoning process we advocate for our excerpt from Grady's example is sketched in Figure 2. Presumably ORGANIZATION AS PHYSICAL STRUCTURES maps physical containment relationships to organizational containment relationships. In the source domain it can be inferred, from the direct meaning of the utterance, that psychoanalysis(-*qua*-house) physically contains the main house, the annexes and outbuildings. Therefore, by the containment mapping relationship, psychoanalysis is being claimed in reality to contain corresponding items. Because psychoanalysis is a thought-system, it is inferred (defeasibly) outside the cocoon that these items are also thought-systems.

The annexes and outbuildings have been built by various psychoanalysts. A source-domain inference from this is that these sub-buildings have been caused to exist by the psychoanalysts. Since these people are also in the target domain, the causation-to-exist is transferred by the Causation/Ability and Event-Shape VNMAs to become cause-to-exist propositions in the target domain. The mapping in ESPOUSED SYSTEM AS HOME combined with the Event-Shape VNMA can create the target-domain proposition that the people came to espouse the thought-systems corresponding to the sub-buildings.

Notice from the Figure that the ESPOUSED SYSTEM AS HOME mapping actions depend on it already being established that the relevant items, the ones viewed as sub-buildings within the pretence, are indeed thought-systems in reality. But that establishing relies on a use of the ORGANIZATION AS PHYSICAL STRUCTURES mapping. Thus the overall reasoning is a complex intertwining of mapping actions, within-pretence reasoning steps, and reasoning steps outside the pretence.

Notice also that significant, defeasible inference steps in source-domain terms are needed, such as going from stopping living in a building to stopping having that building as one's home. Much of the rich general knowledge used here is not paralleled in the target domain and therefore illustrates aspect [B] of *Transfer-Warranting by Unmapped Structure*. The treatment of the example clearly illustrates our main principles of Map-Extension Minimization and Reliance on On-The-Fly Source-Based Reasoning.

Tense in the English glosses in Figures is actually handled by the Time-Order VNMA (Barnden & Lee 2001), together with identity mapping of the present time point. Therefore, the mapping arrows marked

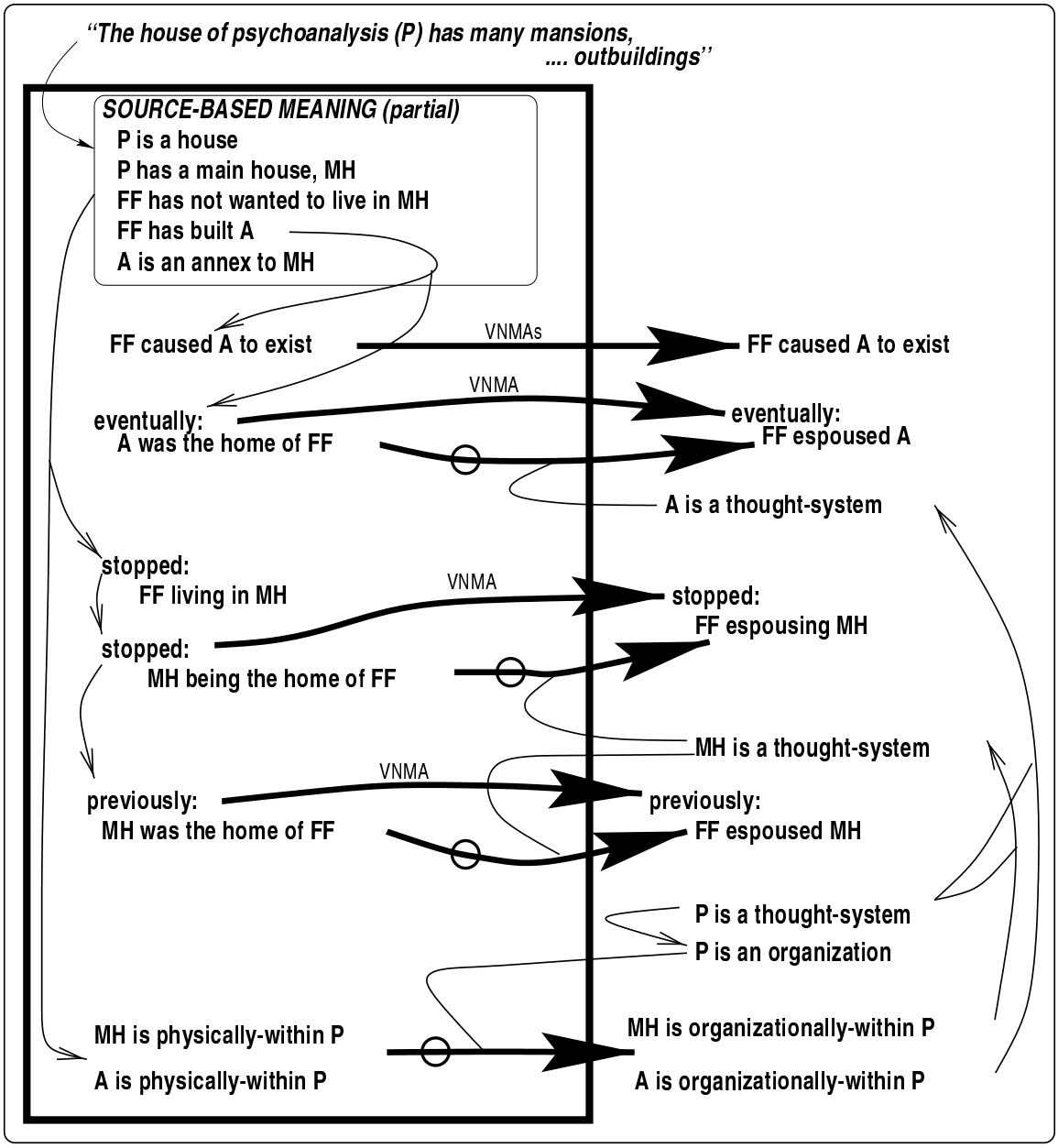

Figure 2: **Grady's House of Many Mansions.** Showing in simplified form how the approach could deal with the Grady excerpt (2). In this and later figures, a mapping arrow labelled VNMA or VNMAs shows the action of one or more VNMAs. A mapping arrow marked with a circle shows the action of a mapping relationship specific to a particular metaphorical view. The statements within the diagram are English glosses of expressions in some internal representation scheme used by the understander. Only a selection of the possible propositions and inferential links are shown. Statements with no apparent support are provided by the discourse or background knowledge. As suggested by the thin lines joining with the fat arrows, mapping applications often require certain side-conditions to be satisfied, e.g. that something is a thought-system or an organization. In the present figure, FF is one arbitrary follower of Freud. Theories are a special case of thought-systems.

by a circle as view-specific often represent in actuality the combined operation of a view-specific mapping relationship and the Time-Order VNMA.

3.2 MIDAS and the Cold

Martin (1990) applies his implemented metaphor understanding system, MIDAS (plus metaphor-acquisition subsystem MES), to examples such as

(3) John gave Mary a cold.

MIDAS can be equipped with knowledge of particular metaphorical views, such as one in which a state of *being infected* is cast as *possessing the infection.* Martin is specifically interested in how MIDAS can then deal with "gave a cold" even though MIDAS does not as yet have a mapping for the transfer-possession sense of "give" (nor a direct sense for "give" in terms of causing infections). Thus, in such examples Martin is focused as we are on map-transcending uses of stored metaphorical views.

MIDAS proceeds essentially as follows. From its built-in knowledge of giving and possessing of physical objects, it does a source-domain reasoning step to infer that John's giving Mary the cold caused Mary to possess the cold. It now applies its known mapping of possessing-TO-being-infected, together with an identity mapping of causation, to infer that Mary was caused to be in the state of being infected with the cold. It then searches for something in the target domain that can cause such a state and finds the action of infection, finally concluding that John infected Mary with the cold. Also, the system surmises that giving maps to infecting (when the given object is an infection), and adds this knowledge to the known mapping above so as to be able to handle future similar examples more directly.

All except these final steps of searching for the infection-act and creating a new mapping proceeds roughly according to our approach, although Martin does not couch his approach in terms of pretence. It follows that our approach can handle the example, except for not addressing those final steps. Note carefully however that our approach, because of the Causation/Ability VNMA, does provide for concluding that *something caused* Mary's being-infected-with-cold state, where that something is *whatever it is* in reality that is being described as John "giving" Mary the cold.

3.3 Variability in Hobbs

Hobbs (1990:p.60ff) discusses, among others, computer-science examples such as "The variable went from 0 to 100," on the assumption that the system knows that a variable being "at" a value V at some time t, using the spatial-position notion of "at," implies that the value of the variable is V at time t. Thus, the system possesses a mapping from spatial position to possessing values (at specific times). The "went from" is map-transcending, however, because there is no mapping of spatial *movement* of a variable to *change* of it value. But, using ordinary reasoning about spatial movement, the proposed system can infer that at

one time t1 the variable was spatially-at 0 and at a later time t2 the variable was spatially-at 100 (where 0 and 100 are implicitly being used as names for spatial positions). Application of the known mapping can then generate the conclusion that the variable's value was 0 at t1 and 100 at t2. Thus, our approach can handle examples such as Hobbs'. His proposal, like our approach, allows unlimited amounts of source-domain reasoning, even though the reasoning for the above example is very simple. Hobbs' proposal avoids creating a mapping for the physical movement (though his approach to novel metaphor, *ibid.* pp.64ff, does involve map-creation).

Narayanan (1997) is another author whose overall approach to map-transcending metaphor is similar to ours in not relying on map extension and relying instead on source-domain reasoning, so our approach can handle his examples as well in much the way we have indicated for Hobbs and Martin.

4 Application to Examples Discussed under Non-Neighbouring Approaches

4.1 The Fass Drinking Car

A representative example used by Fass (1997) to illustrate his meta5 system is

(4) My car drinks gasoline

taken from Wilks (1978).[7] The system can interpret this as meaning "My car uses gasoline" essentially by finding an analogical match, from scratch, between the knowledge that animals drink liquids and cars use gasoline. The meta5 system has no prior knowledge of any particular metaphorical views. However, we can arguably reanalyze the example more naturally as a map-transcending manifestation of a stored view.

Plausibly, ordinary English users possess a view of MACHINES AS CREATURES (as indeed Fass 1997:p.318, points out). Utterances such as "my radio is dead," "my car has life in it still," "he killed the engine," "a middle-aged toaster", and "an aggressive lawn-mower" seem to the present authors mundane and easily understandable. We assume that as part of the view, a machine's running corresponds to biological activity (internal or external). *This is the only view-specific mapping we need in order to be able to deal with (4).*

The reasoning process is sketched in Figure 3. From the direct meaning of the utterance and source-domain general knowledge, it can be (defeasibly) inferred in source-domain terms that gasoline helps the car to be alive (biologically). But, by default, being alive enables the creature to be active. Therefore, gasoline helps the car to be (biologically) active. The Causation/Ability VNMA combined with the above MACHINES AS CREATURES mapping now provide the target-domain contribution that gasoline helps the car to run. Further reasoning shown in the Figure could now produce the conclusion that the car *uses* gasoline.

[7] The example is varied to "The car drank gasoline" at some places in Fass (1997).

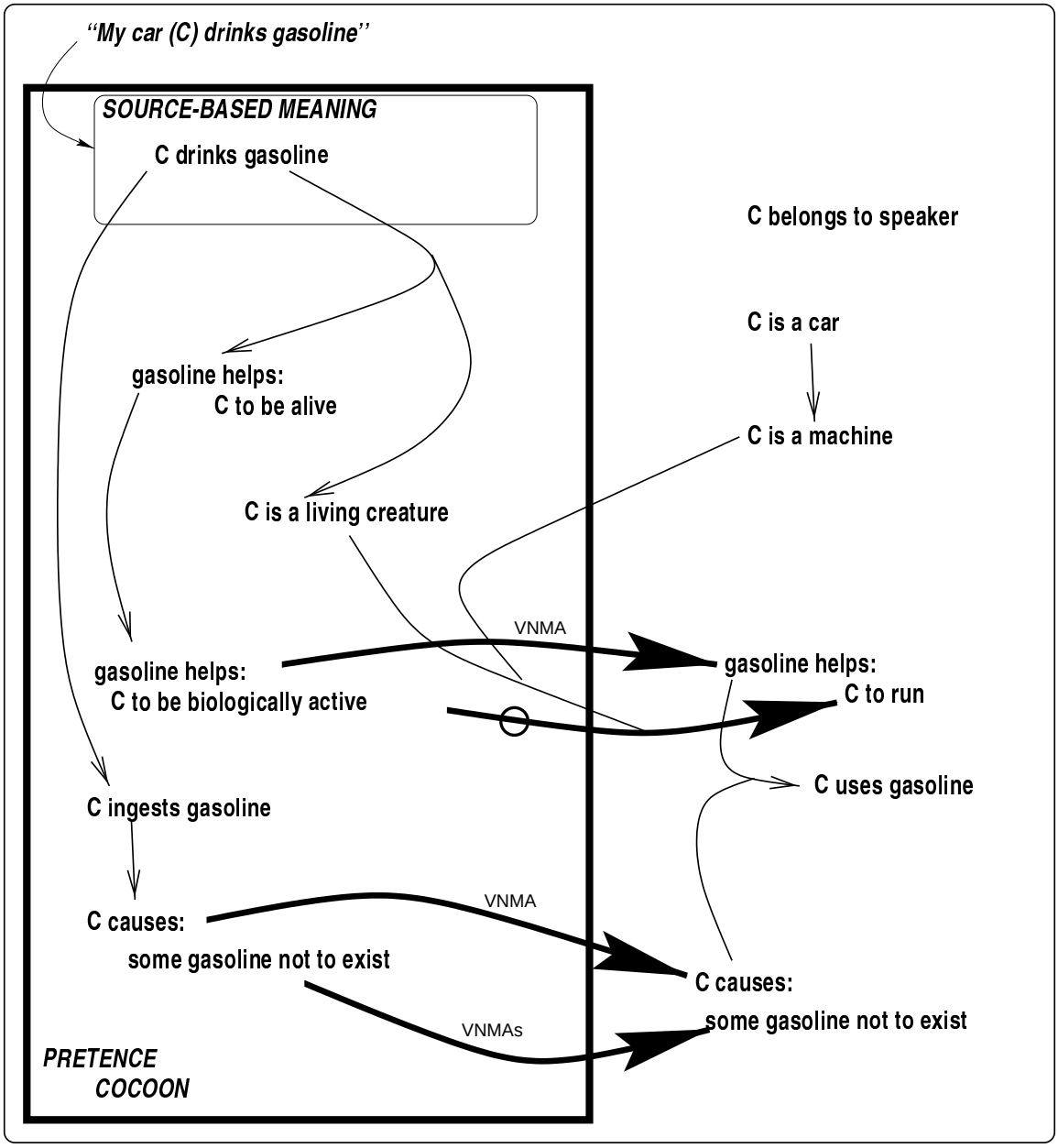

Figure 3: **Fass drinking-car example, (4).**

The production of these target-domain contributions does not require any mapping of drinking itself to be created. Of course, an understander *could* go on to do the extra work of mapping drinking itself to, say, the process of a car having gasoline put in it. Our point is that the informational contribution Fass proposes (that the car uses gasoline) does not require this extra work.

In our experience of discussions about metaphor, people actually prefer to interpret "My car drinks gasoline" as connoting that the car uses gasoline *quickly and copiously.* (Compare "the blotting paper drank up the ink." Also the entries in Webster's *Third New International Dictionary* suggest at least moderate rapidity of ingestion.) Fass (1997:p.192) himself makes a similar observation, but dismisses it with only a brief comment. But, because an act of drinking is normally moderately fast, a use of the time-duration VNMA would allow our approach to conclude that the car's use of gasoline is moderately fast. In order to for meta5 to come up with this contribution, it seems that it would need to already have in the target domain a representation of cars using gasoline moderately fast, because otherwise there would be nothing in the target domain to be found to be analogous to the source-domain situation. This seems to us an undesirable restriction on the system's abilities.

4.2 Jail Job

A central example used by Glucksberg & Keysar (1990) in their category-inclusion approach to metaphor is

(5) My job is a jail.

According to those authors, the utterance uses the jail as an exemplar of a (possibly ad hoc) category that also covers the speaker's job. This category could, for instance, be that of things that confine one against one's will, are unpleasant and are difficult to escape from. We actually change the example to "Tom's job is a jail [for Tom]" for minor expository reasons. We note that the job could be confining for Tom in the sense that he cannot quit it *and/or* in the sense that he cannot partake (or is strongly inhibited from partaking) in activities that are not dictated by his job, such as leisure activities, even outside working hours, *and/or* that he cannot leave his job-site suring working hours. There are further alternative informational contributions.

In applying our approach to the example we note that, surely, one thing that underlies one's understanding of the sentence is that Tom is, metaphorically speaking, *physically in* the jail. Evidently, the metaphorical view that actually motivates the example is simply a familiar metaphorical view of LIFE-ROLES AS LOCATIONS, a special version of STATES AS LOCATIONS. One commonly talks of, for instance, being "put in", "leaving" or being "thrown out of" a job or other sort of position, The language of physical location and movement is fairly systematically used for discussion of jobs and the like. Thus, our analysis of the example is radically different from that of the category-inclusion approach.

We assume that the direct meaning of the utterance, namely *TJ is a jail for Tom,* where TJ is Tom's job, is put into the pretence cocoon. Let us suppose that a possible source-based default inference from TJ being a jail for Tom this is that *Tom must continue to be physically-in TJ.* And suppose that another such

inference is that *Tom cannot partake in activities physically-outside TJ.* Given that mappings involved in LIFE-ROLE AS LOCATION assert a correspondence of (suitable cases of) being physically in a role, X, and having X as one of one's roles, and a correspondence between activities taking place physically-within X and activities forming part of X *qua* role, the inferences displayed above straightforwardly transfer to give the target-domain contributions that *Tom must continue to have role (job) TJ* and that *Tom cannot partake in activities that are not part of TJ.* We have also appealed here to the Event-Shape, Ability and Negation VNMAs.

5 Application to Our Databank

The pair of examples in the present section come from our metaphor-of-mind databank, cited above. We suggest informational contributions that, we believe, are plausible ones in the actual contexts of the examples. We treat the examples in some detail as they richly illustrate the principles of our approach.

5.1 Anne and Kyle

(6) Then one day ... Anne found a recent ticket stub in her husband's [Kyle's] pocket; he'd never mentioned taking in a show. "If any of my girlfriends had told me a similar story, I would've asked if they were sure their mate wasn't seeing someone else," says Anne. "Instead, I just kept saying, 'This is the thin part of the *through-thick-and-thin* section of our wedding vows.' "

In the far reaches of her mind, Anne knew Kyle was having an affair, but "to acknowledge the betrayal would mean I'd have to take a stand. I'd never be able to go back to what I was familiar with," she says. Not until eight months had passed and she finally checked the phone bill did Anne confront the reality of her husband's deception.[8]

We concentrate on the clause shown in bold font. We assume that this clause manifests two metaphorical views, MIND AS PHYSICAL SPACE and IDEAS AS PHYSICAL OBJECTS, for an understander U. The use of the second is more implicit than that of the first, as we will see. We assume that U does not have a mapping of the particular notion of "far reaches" of a mind-space into the mental domain, so that the utterance is map-transcending for U in this respect. Moreover, we assume it is map-transcending in that U has no mapping for a *physically-located* episode of knowing something. For the example, the only mapping relationship that we assume U to possess as part of MIND AS PHYSICAL SPACE is:

(7) When a person's mind is being viewed as a physical space, an idea's being physically located in the space corresponds to the person's being able to operate mentally on the idea, to a very low degree at least.

[8] In Linden Gross, "Facing up to the Dreadful Dangers of Denial," *Cosmopolitan, 216*(3), USA ed., March 1994. Bold font added, italics in original.

We also assume that U holds that:

(8) When a person's mind is being viewed as a physical space, the person's conscious self is viewed as a person physically located in that space. Moreover, if the space has a main subregion, the conscious self is in that subregion.[9]

(9) When an agent X is viewed as having a mental state or process physically within some subregion L of his/her mind, then the ideas involved in that state or process are normally also physically located in L.

(8) and (9) are *ancillary assumptions* forming part of the view. They enrich the nature of the alleged physical space rather than being source/target mapping relationships. The second ancillary assumption reflects the fact that in uses of MIND AS PHYSICAL SPACE in real discourse, mental *states and processes* are often portrayed as having physical locations in the mind.

As for IDEAS AS PHYSICAL OBJECTS, the only mapping U needs to possess is as follows:

(10) When an idea entertained by a person is being viewed as physical object, and the person's conscious self is viewed as a person, then the ability of the conscious self to *physically* operate on the idea corresponds to the real person's ability to operate in a *conscious mental way* on the idea.

In fact, this is a special case of a more general correspondence, where the single idea is replaced by a set of ideas, and physical or mental operation on a set of ideas is assumed to mean conjoint operation on them. However, the more general version is not needed for the current example.

Stipulations (8-10) help to make sense of a large array of metaphorical utterances based on MIND AS PHYSICAL SPACE and/or IDEAS AS PHYSICAL OBJECTS, not just the present example. In fact, for many examples other the Anne/Kyle one, no mapping relationships other than the ones in this subsection (supplemented by VNMAs) appear to be needed.

We suggest that a likely informational contribution of the bold clause in (6) to the understanding of the discourse as a whole is the proposition that

(11) The idea that Kyle was having an affair was something of which Anne had only a very low degree of conscious awareness (i.e. on which Anne had only a very low ability to mentally operate in a conscious way).

The contribution can be produced as follows. (Figure 5 below shows the operations done by ATT-Meta on this example, to be explained in Section 8.) The direct meaning of the sentence is to the effect that

[9]We have also developed a variant account in which the person herself, rather than her conscious self, is in the physical mind-space. Note also that the very notion of a conscious self may be metaphorical.

physically-in the far reaches of Anne's mind, Anne knew that Kyle was having an affair. From this and (9), U can infer within the pretence cocoon that *the idea K that Kyle was having an affair was in the far reaches of Anne's mind.* Then, if we assume U is equipped with suitable commonsense knowledge about people, physical objects and regions, U can infer within the pretence cocoon that *idea K was physically very inaccessible to Anne's conscious self,* because by (8) Anne's conscious self is a person and is in the main part of Anne's mind, distant from the far reaches. A further step of source-domain reasoning can conclude that *the conscious self had only a very low level of ability to operate physically on the idea.* Then the physical-operability/conscious-mental-operability correspondence (10) can map this source-domain conclusion to become the target-domain contribution (11). This step appeals to the Qualitative Degree VNMA as well.

Map-Extension Minimization comes in because we claim that understanding of the clause does *not* involve an allegation that Anne's mind *really has some aspect that corresponds to the "far reaches" of the physical space.* Finding such a correspondent would probably be impossible for an ordinary understander U anyway, because there is no generally held, *non-metaphorical,* common-sense or scientific model of mind that would allow the identification of an aspect that would count as "far reaches." U could trivially just stipulate that Anne's mind has something that corresponds to "far reaches," but it is not clear what purpose this would serve. Not only is there no need to map far-reaches, but also there is no need to map relationships of physical distance within the mind-space to mental relationships within the mind. Rather, it is a *consequence* of physical distance, namely a very low degree of physical operability for the idea in question, that ends up being mapped to the target domain.

Aspect [A] of Transfer-Warranting by Unmapped Structure is illustrated by the fact that the direct meaning is not mapped, and aspect [B] by the fact that the utilized general knowledge about physical objects and space—for instance that distance implies inaccessibility by people—has no mapping into the target domain. Finally, the within-pretence reasoning steps that are not produced by ancillary assumptions just rest on rules about *ordinary physical entities,* and have nothing to do with metaphor or the target domain (Anne's mental states).

5.2 One Small Voice Amongst Others

The following are typical examples of one common metaphorical way of talking about mental states in ordinary text and conversation: *"Part of me was thinking, do I really want to have a serious relationship with someone I have to hide part of my life from?"*[10]*; "Did part of you think, 'Yes, ... I'm flattered ...'?"*[11]*;* and *"Part of you wants to talk [about the personal problem] but part of you doesn't."*[12] We will concentrate on the following more subtle example:

(12) Suddenly I was having second thoughts. About us, I mean. Did I really want to get married and spend the rest of my life with Mick? *Of course you do,* one small voice insisted. *Are you quite*

[10] *Glamour,* January 1996, p.184.
[11] Said by presenter Sue Lawley to interviewee in Desert Island Discs program, Radio 4, England, 12 May 1995.
[12] Said by counsellor Ann Dale in lecture at Wycombe Abbey School, England, 10 May 1995.

sure about that? another nudged. So much was going on in my head, I couldn't sleep.[13]

We assume these examples to manifest a stored metaphorical view of MIND PARTS AS PERSONS, closely connected to metaphorical views of self discussed by Lakoff (1996) (see Barnden 1997, for further discussion). An agent's inner persons under this metaphorical view can have have their own mental states, emotions, etc. and can communicate in natural language with each other and with the whole agent (as in (12), and notice that "one small voice" could felicitously be replaced by "one small part of me"). MIND PARTS AS PERSONS is a good example of source and target domains overlapping (recall section 2.3). Mental states and natural language communication are important in both domains.

Our application of Map-Extension Minimization to MIND PARTS AS PERSONS examples lies mainly in our contention *there is no need to try to find some real aspects of the agent's mind that correspond to mentioned or implied "parts,"* Rather, we only need to assume that the understander, U, has the following mapping relationship:

(13) the proposition that a sub-person P of an agent X believes, desires, intends, mentally-questions ...
T maps to the proposition that X *has some reason to* believe, desire, intend, mentally-questions ... T.

We also include the following ancillary assumptions in the metaphorical view, although they are not needed for the present example:

(14) Any utterances by a sub-person P of an agent X are part of a private conversation involving (only) X and his/her sub-persons.

(15) This conversation is physically inside X.

Because of this second assumption, MIND PARTS AS PERSONS examples that involve utterances by the mind-parts are also manifestations of IDEAS AS INTERNAL UTTERANCES, a metaphorical view we will not address extensively in this article but that is discussed in Barnden (1997).

Example (12) further illustrates the Map-Extension Minimization default, as there is no need either to map the notion of "insisting" or of a "small voice" to the mental domain or to map the particular small voice mentioned. Notice that the "you" in that example arguably refers to the whole agent, and that this is consistent with assumption (14).

The overall rough pattern of reasoning suggested for the example is sketched in Figure 4. We assume that the understander has already presumed that IDEAS AS INTERNAL UTTERANCES is in play, and therefore that the small voice is an *internal* voice, not a real one. Then, using knowledge about the source domain of verbal interaction, the understander can infer within the pretence cocoon that there is a person

[13] *My Story* magazine, May 1995, Editions Press Limited, Gibraltar, p.6/7. Italics in original. Paragraph breaks deleted.

P whose voice it is, and since the voice is inside the writer (the "I" in the discourse) then presumably that person is too. Similarly, the owner of the second mentioned voice is presumably a sub-person of Mary. Now, the fact that the first sub-person *insists* something implies that it asserts it, which in turn implies by default that it believes it. The fact that the second sub-person verbally questions the first's assertion implies by default that the second has a mental questioning attitude towards what was asserted by the first. These inferences are again performed using ordinary commonsense knowledge about speech and thought. Thus, by means now of the mapping relationship (13) the understander can infer (by default) the contributions that (a) the agent has some reason to believe that she wants to marry Mick and that (b) the agent has some reason to question whether she wants to marry Mick.

MIND PARTS AS PERSONS examples illustrate Transfer-Warranting by Unmapped Structure particularly vividly, as the sub-persons are not mapped to real aspects or components of the agent. Also, examples like the small-voice one where natural language communication comes in require considerable commonsense knowledge about verbal interaction. This knowledge does not need to be mapped.

6 Approach Contd.—Discourse-Query-Directed Guidance of Reasoning

We have postponed the issue of the present section until now because the approach's abilities discussed so far were able to be be presented independently of it, and because it raises large, outstanding discourse-understanding problems that have relevance far beyond metaphor. The section shows that our approach is an attempt to make metaphor-processing research more attentive to discourse context and goals arising from discourse, a need strongly emphasized by Cameron (1999a).

In previous sections, we have purported to reason forward from the direct meaning of utterances (in the pretence cocoon) and then applied metaphorical mappings to arrive at informational contributions in the target domains. We ignored other possible inferences that could be made, not specifying any principles whereby the reasoning could be guided towards the particular contributions discussed. But we now observe that powerful guidance is often, and perhaps typically, at hand. We propose two guidance principles, both appealing to the notion of *query-directed reasoning* (more commonly but less clearly called goal-directed reasoning). Query-directed reasoning is common in AI because of its tendency to guide reasoning usefully. The queries in question are hypothesized to arise from (partial) processing of the co-text and other context.

In query-directed reasoning in general, a query is simply some proposition posed within a reasoner as a question that is of interest: it is not necessarily something that the reasoner *wants* or *expects* to be true. The reasoner proceeds backwards from queries rather than forwards from provided facts. For a particular query, the reasoner performs reasoning acts to find support for and/or against it. The reasoner investigates the query by seeing whether it matches known facts or could be the output of reasoning rules or other mechanisms. The latter path typically causes further queries to arise, which we will call subqueries. For instance, given the query "John resents Mary?" one subquery could be "John is often rude to Mary?" This might arise because there is a default rule saying that if X is often rude to person Y then X resents Y.

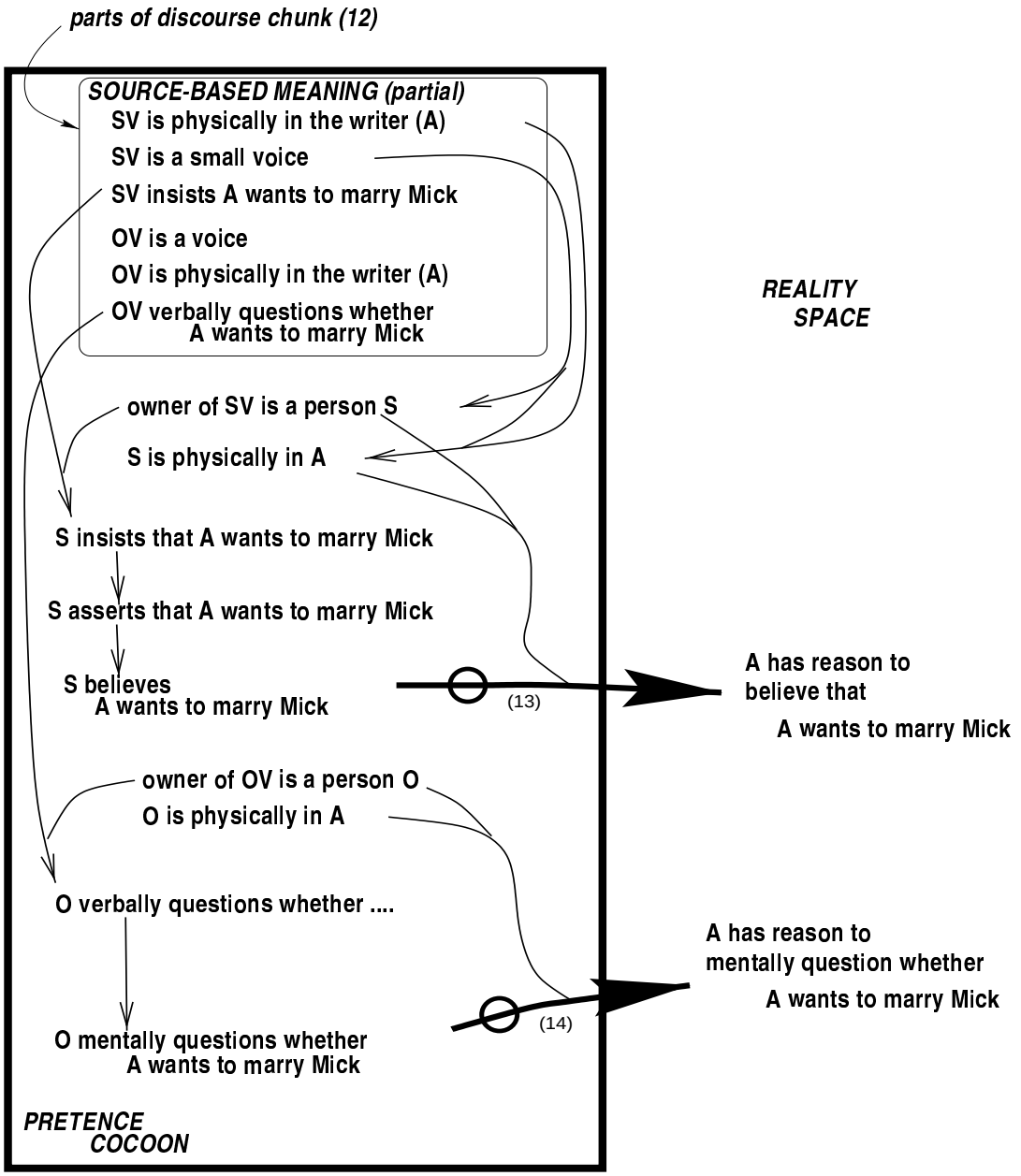

Figure 4: **Small-Voice example, (12).** The example uses mapping relationship (13), as depicted by the upper and lower fat arrows, but not the ancillary assumptions (14) and (15).

A query can have some unspecified parts, and one effect of gathering evidence for the query could be to specify those parts. For example, a query could be "Sam likes X?" where X is a variable, and then evidence may be found that Sam likes piano music, Sam likes peanut butter, and Sam likes himself.

The result of investigating a query can be to refute it. When a query has been supported to less than full certainty it is beneficial to investigate its negation, in case the evidence for the negation is stronger. So, to continue the example, the system might set up "Sam does not like peanut butter?" as a new query, and this may turn out to be more strongly supported than "Sam likes peanut butter."

In our approach, the two guidance principles based on query-directed reasoning that were alluded to above are:

A. In attempting to draw contributions from a metaphorical utterance, the understander can exploit its surrounding context by guiding the reasoning towards addressing any queries set up by processing that context.

B. The Map-Minimization Default is another powerful reasoning-guidance tool, by guiding reasoning towards linking up with existing mapping relationships. The conditions of application of those relationships are tantamount to queries.

The informational contributions addressed in examples above can *in principle* be derived without this help—the question is the practical one of how to guide reasoning towards them so as to reduce the effort spent on fruitless or irrelevant inferencing. We now consider the two principles in reverse order.

In MIND PARTS AS PERSONS of the metaphorical mapping relationship (13) maps mental/emotional states of subpersons. By this very fact, within-pretence reasoning can be guided (in part) towards concluding that subpeople have mental/emotional states. Thus, in the Small-Voice example, no mental/emotional state or subperson is explicitly mentioned in the utterance, but principle (B) provides guidance towards finding such subpeople and their mental/emotional states if the understander has any reason to suspect that the MIND PARTS AS PERSONS view is relevant.

In principle (A) we assume a partially discourse-query-directed account of understanding, for example as detailed in Lee (1998). Hobbs (1985) and Mann & Thompson (1987) have discussed how the need to find coherence between utterances causes reasoning queries to arise. And it is well recognized in the metaphor literature that the interpretation of metaphorical utterances can be affected by context. In particular, context-derived reasoning queries has played an important role Hobbs' approach to metaphor and other matters. They have done so also in the study of analogy, including the research surrounding ACME (Holyoak & Thagard 1989), the SME version mentioned by Markman (1997), and IAM (Keane 1988). Some psychological experiments on bridging inferences in discourse understanding—inferences aimed at tying different utterances together—suggest that much of the information extracted from utterances is only extracted in order to help explain other utterances. For instance, if there is mention in one utterance of coffee being

stirred, then there is no inference of a particular type of stirrer unless that information is helpful for understanding another utterance (e.g. "The spoon was dirty": processing of this sentence can then set up a query roughly like "Is there a spoon?") However, the psychological evidence appears mixed (see the various studies in Graesser & Bower 1990) and to be clouded by methodological dispute (Keenan *et al.*, 1990). Finally, Cameron (1999b:p.127) stresses that metaphorical utterances are often accompanied by explanatory paraphrases close-by.

Thus, consideration of contextually-derived queries is not unique to our approach, but it is unusual for the matter to be addressed in an algorithmically detailed way in metaphor research. Of course, utterances must be able to yield informational contributions to some extent without the help of queries arising from context. Therefore, we expect that the reasoning that occurs during understanding is in general some mixture of reasoning that moves forward from the direct meaning of the utterance and reasoning that moves backward from the queries arising from context. We do not as yet have a detailed account of the management of this mixture, as we have paid attention mostly to the query-directed aspects. Nor do we have a complete theory of how queries are mutually posed by utterances in a discourse in the first place. However, in the following we will sketch plausible processes whereby queries could arise in particular examples.

We do not assume that utterances in a discourse are *fully* understood in order of presentation, although no doubt they are partially understood in that order. Thus, queries suggested by an utterance can guide fuller understanding of another, only partially understood, utterance, which may be either earlier or later in the discourse. In particular, a reasonable strategy for many types of troublesome utterance, including metaphorical ones detected as being such, is to be prepared to suspend the understanding process until succeeding context can be examined.

In the Anne/Kyle example we addressed the clause "In the far reaches of her mind, Anne knew Kyle was having an affair." We showed how the contribution that *the idea that Kyle was having an affair was something of which Anne had only a very low degree of conscious awareness* could be reached. We now show how reasoning could be guided in a rather direct way towards this contribution by a context-derived query, illustrating the action of principle (A). The clause in question is immediately followed in (6) by *"but 'to acknowledge the betrayal would mean I'd have to take a stand. I'd never be able to go back to what I was familiar with,' she says."* We assume that one sense the understander possesses for "to acknowledge [that]" is to have conscious awareness that something is the case.[14] The "but" joining the two clauses suggests that the issues raised by the second clause are important for the first clause, and that there may be a contrast.[15] Therefore, it is reasonable to suppose that preliminary semantic/pragmatic processing of the second clause can give rise to the following query posed with respect to the first clause:

to what degree is Anne consciously aware, within the situation reported in the first clause?

[14]cf. "to recognize to one's self" in Webster's *Third New International Dictionary*. Note that even if a non-mental sense of "acknowledge" were adopted in the current passage, it would defeasibly imply conscious mental awareness in any case.

[15]The sense of contrast is heightened by the fact that the verb "acknowledge" suggests an opposition of views explicit or implicit in the overall discourse or discourse situation, as evidenced by Hunston's (1995) corpus study of the use of this verb.

Since preliminary processing of the first clause can yield that that situation is to do with knowing that Kyle was having an affair, the query can readily be refined to become

(16) *to what degree is Anne consciously aware that Kyle was having an affair?*

This is similar to the informational contribution, (11), in section 5.1. Therefore, the reasoning presented there can arise *in response to the query* rather than be produced in a forwards direction from the clause's direct meaning. We propose that the presence of the last query above, which is outside the pretence cocoon, causes a within-pretence-cocoon query to be created, by virtue of the IDEAS AS PHYSICAL OBJECTS correspondence between conscious awareness and physical operability that was mentioned in section 5.1. This within-cocoon query is

to what degree can Anne's conscious self physically operate on the idea that Kyle was having an affair?

This subquery creation is an example of the operation of query-directed reasoning. The new query can then lead to a straightforward query-directed reasoning process in the pretence cocoon, eventually going back to the provided information that Anne's knowing that Kyle was having an affair was in the far reaches of her mind. The process envisaged here sets up the reasoning steps in broadly the reverse order to that in which we described them in section 5.1.

Notice that the query *to what degree is Anne consciously aware that Kyle was having an affair?* is not the contribution itself. In a different example, the answer to this query might be that Anne had a *high* degree of conscious awareness. Thus, we are *not* claiming that context of an utterance explicitly suggests the very contributions that the utterance may have, so that they need only be *checked* by processing that utterance, but rather that context supplies *queries* that usefully guide reasoning. The reasoning may be towards contributions that are filled-out versions of the queries, or that are negative answers to the queries.

In a different context, the discussed contribution (11) would require more imagination on the part of the understander. For instance, if the "In the far reaches ... Anne knew ..." clause were uttered as the reply to the question "Did Anne know that Kyle was having an affair?" and no further information were given, then it would be more up to the understander whether the work to derive the above contribution is done. In the new context, the query posed is merely whether Anne knows about the affair. This query would merely guide the understander to produce another, simpler, contribution, namely that Anne does have some awareness of the affair.

7 Implementation of the Approach in ATT-Meta: Introduction

Here we outline some aspects of ATT-Meta that are important for this paper, and in the next section we give a detailed account of the application of the system to the Anne/Kyle example. We do not go into the fine technical detail of the representation scheme and reasoning algorithms, for reasons of space.

Information is expressed in a situation-based or episode-based first-order logic (broadly similar in spirit to the logical scheme of Hobbs 1990). But this paper will use a different, easier-to-read syntax. The exact formulations are not important for the purposes of this paper.

7.1 ATT-Meta as Opposed to the General Approach

The ATT-Meta system is merely a partial, limited implementation of our approach, and the capabilities of the approach and those of the system must be carefully distinguished. The aspects of the approach that ATT-Meta does capture are:

- Use of pretence cocoons.
- Finding target-domain contributions by a combination of within-pretence reasoning (incl. application of ancillary assumptions) and application of already-known metaphorical mapping relationships, while avoiding the extension of mappings.
- Ability to perform mapping actions in either direction between source and target domains.
- The irreality of domains and metaphorical views. ATT-Meta has no explicit knowledge of domains or views. In a given application of the system, it is given an undifferentiated set of rules covering various domains, and some individual mapping relationships and ancillary assumptions that are not explicitly marked as pertaining to particular metaphorical views.

  Rules from any domain can be considered for application inside the pretence cocoon. In practice, the applications are mostly of rules from the source domain.
- Guidance of reasoning by reasoning queries.
- The need to allow for uncertainty in reasoning, whether the reasoning is within a pretence cocoon, outside, or between the cocoon and outside. (This aspect is beyond the scope of the present paper. See Barnden 1998b and Barnden & Lee 1999.)

ATT-Meta also captures various other aspects of the approach that are omitted from this paper, such as some principles concerning the handling of reasoning conflicts (Barnden, in press, a), and principles concerning the handling of metaphor compounds (Lee & Barnden, 2001a). ATT-Meta also includes general-purpose facilities for reasoning about the beliefs and reasoning of other agents (Barnden 1998b; Barnden & Lee 1999); the facilities can operate either independently of or in alliance with metaphorical reasoning. The aspects of the approach that ATT-Meta does *not* capture are

- View-neutral mapping adjuncts (VNMAs), except for the fact that there is a preliminary though general approach to the Uncertainty VNMA, and the fact that view-specific rules can be written to include the effects of other VNMAs.

- The process whereby reasoning queries are generated by processing of surrounding discourse (section 6).
- Mixing of query-directed reasoning and data-directed reasoning (reasoning that goes forward from provided information)—all reasoning in ATT-Meta currently is query-directed.
- The creation of direct meanings of utterances.

As regards the last point, ATT-Meta is currently only a reasoning system: the user supplies to it hand-constructed logical forms for direct meanings of imagined natural language inputs. The construction of a front-end for dealing with text directly is an important long-term aim of our project, but our focus so far has been on reasoning. This reasoning is a difficult matter especially because of the needed uncertainty and the connotative richness and variability of metaphor.

Also, we have not developed representational syntax and reasoning rules adequate for dealing properly with such matters as modal conditions, aspectual conditions, time, changes of state, etc. However, there is no bar to the addition of the necessary representational and rule machinery. In particular, the system's representations use a logic that reifies events (as does the logic of Hobbs 1990) and is thus well-placed to handle many of the neglected matters.

Neither the general approach nor the implemented system yet addresses the question of how to detect that an utterance is metaphorical. However, preliminary considerations on this are mentioned in section 12.

ATT-Meta has so far largely been applied to metaphorical views of mental states and processes, such as IDEAS AS PHYSICAL OBJECTS and MIND PARTS AS PERSONS. However, neither the general approach nor its instantiation in ATT-Meta are in any way restricted to mental-state metaphor.

7.2 What ATT-Meta Does Overall

ATT-Meta itself has no knowledge of any specific metaphorical view or any particular domain. It is merely a reasoning engine. It does, however, include some built-in rules about pretence, beliefs, qualitative degrees, and some other matters. The user supplies, as data for ATT-Meta, whatever target-domain knowledge, source-domain knowledge, metaphorical mappings and ancillary assumptions the user wants to try out for handling examples. These pieces of information are expressed as rules (with facts as a special case). The rules are applied by ATT-Meta.

The user supplies facts that are intended to couch the logical forms of the direct meanings of a small set of utterances that form a (real or imaginary) chunk of discourse. Utterances may or may not be metaphorical. In the metaphorical case, the user explicitly signals the metaphoricity in the facts, as exemplified below. The user also supplies needed additional facts about the entities referred to in the imaginary discourse, such as that Anne and Kyle are people.

The user supplies a top reasoning query that ultimately drives all the reasoning. For example, the query could be *"Anne knows that Kyle is having an affair?"* or *"Anne is consciously aware to degree at least D of the idea that Kyle is having an affair?"* When the top query does not contain a variable (such as the D just above), ATT-Meta tries to find evidence for and against the top query, using the user-supplied knowledge and the user-supplied logical forms of utterances. In cases of conflict, a conflict-resolution mechanism attempts to favour one side or the other. When the top query contains variables, ATT-Meta tries to find values for the variables such that the instantiated version of the query are supported. The top query and its negation and instantiations if investigated are ultimately given qualitative certainty levels.

The main output of the system are those certainly levels. The system also maintains a record of all queries created, certainty levels assigned to them, and rule applications performed. Various statistics of the system's operation are output.

All rules (including facts) are on a par with each other and are applied equally and with arbitrary mixing in the sense that ATT-Meta has no knowledge of which rules are intended by the user to be rules private to some domain, or to be ancillary assumptions in metaphorical views, or to be rules couching metaphorical mapping relationships.

7.3 Hypotheses and Rules

The information manipulated by ATT-Meta consists of *hypotheses* and *if-then* rules. If-then rules are mostly supplied by the user, though there are some built into ATT-Meta. The set of rules is constant through a given application of ATT-Meta, but the set of hypotheses is incrementally grown during processing. User-supplied rules couch the domain knowledge, knowledge of metaphor and the logical forms for the discourse chunk as mentioned above. In the illustrative syntax used in this article, a simple rule about birds could be

```
(17)     IF is-bird(X) AND NOT(dead(X)) THEN {presumed} can-fly(X).
```

The "presumed" is a qualitative certainty qualifier, and can also be read as "by default." A rule's IF-part is usually a conjunction of atomic conditions, or negations of atomic conditions, where atomic means that there are no internal connectives like AND and OR, or use of logical quantification, and the THEN-part is usually an atomic condition or a negation of one. If-then rules with a special null IF part count as fact rules. We will write a fact rule without any IF or THEN part, as in

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

Facts can have any of the available certainty levels (see below), but in our current practice they are generally "certain."

Hypotheses are reasoning queries or propositions created by rule applications. A query is either the top query supplied by the user or a subquery created by ATT-Meta in the course of reasoning in relation to the

top query. Once a query has had some evidence found for it, it can be viewed as a (partly and preliminarily) established proposition. As a query-hypothesis is reasoned about and accumulates support, it becomes more and more proposition-like until it is eventually "finalized" and treated as a proposition established to some final level of certainty. One simple way a query can be supported is if it matches a user-supplied fact rule. In this case the query is deemed to have become a *fact*.

Hypotheses can contain conjunction, disjunction and negation. Currently, no explicit quantification is allowed in hypotheses. However, queries can contain variables, and these are regarded implicitly as existentially quantified. Thus, intuitively, the query `can-fly(X)` is asking whether there is at least one entity that can fly. The system tries to find all possible values for X for which `can-fly(X)` is supported.

At any time, any particular hypothesis H is tagged with a qualitative certainty level, one of `certain`, `presumed`, `suggested`, `possible` or `certainly-not`. The last one just means that the negation of H is `certain`. `Possible` just means that the negation of H is not certain but no evidence has yet been found for H itself. `Presumed` means that H is a *default*: i.e., it is taken as a working assumption, pending further evidence. `Suggested` means that there is evidence for H, but the evidence is not (yet) strong enough to enable H to be a working assumption. When a hypothesis is created (as a query), it is immediately given a certainty value of `possible`. Of course, reasoning may sooner or later downgrade it to `certainly-not` (i.e., impossible).

Queries cause the appearance of other queries through a standard backwards use of rules. For instance, consider the above rule about birds flying. If the query is `can-fly(Bertie)` then the subquery `is-bird(Bertie)` is created, and if this receives adequate support the subquery that `NOT(dead(Bertie))` is created. Suppose that both of these hypotheses are actually `certain` facts because of fact-rules given by the user. Then the rule provides support for the hypothesis that Bertie can fly, so that this hypothesis is now a proposition that the system to some extent holds to be true. This hypothesis is, however, still query-like in that it can continue to collect support from other rules.

When a rule is applied, the certainty it contributes to its result is the minimum of its own certainty qualifier and the certainty levels assigned to the hypotheses picked up by the condition part. So the above bird rule can at most give a level of `presumed` to its conclusion, and will give less than that if the one of the hypotheses picked up by the IF-part is less than `presumed`.

Multiple rule applications can support a given hypothesis. In this case, the maximum of the certainty values contributed by the different applications is used. On the other hand, when a query is investigated, and does not contain variables, the negation of the query is also investigated. This is done in case the evidence against the query (i.e., the evidence for the negation) is at least as strong as that for it. When there is evidence to at least level `presumed` for both a query and its negation, then a heuristic conflict-resolution mechanism is called in which tries to adjudicate the relative evidence strength. This can result in one hypothesis winning, in which case the other is downgraded in certainty value. If no strength difference can be decided upon, then both hypotheses are downgraded.

Hypotheses are annotated as to which "reasoning contexts" they exist in. In this article, we need consider only two contexts: a single pretence cocoon and the reasoning environment outside the pretence cocoon, which we call the "top" context. The top context is the system's own reasoning context, in which the hypotheses are about reality as the system sees it. A hypothesis can be replicated in different contexts, but if so the different copies are handled entirely separately, as if they were unrelated hypotheses. This is because even if a hypothesis is supported in two different contexts, the lines of reasoning supporting it can differ radically, and the level of certainty can differ. Hypotheses within pretence cocoons have certainty levels, just as hypotheses in the top context do.

Except in the case of special types of rule to be mentioned below, when a rule is applied to a query in a specific reasoning context, the queries emanating from its condition part are set up in the same reasoning context. Thus if the above rule about birds flying is applied to the query `can-fly(John)` in the pretence cocoon, the subqueries `is-bird(John)` and `NOT(dead(John))` arise in the pretence cocoon, not in the top context.

Rules can be considered to be of three types: ordinary rules, like the birds-flying one, that can intuitively be thought of as representing knowledge within a single domain; *conversion rules*, that encapsulate metaphorical mapping relationships; and *ancillary-assumption rules*, that encapsulate the ancillary assumptions mentioned in section 2.6. The various types of rules will be exemplified below. The different types of rules are not explicitly typed within the system—they are all treated equally.

7.4 Use of Knowledge in Pretence

As stated earlier, any non-fact rules pertaining to any domain (including even the target domain) can be used inside the cocoon. However, if they have strength `certain` they are downgraded to strength `presumed` to allow greater fairness in conflict resolution inside the cocoon. Somewhat analogously, any fact in the top reasoning context can be "imported" into the cocoon, in effect by a special view-neutral conversion rule that goes from outside the cocoon into it. `Certain` facts are downgraded to `presumed`. Importation is exemplified below.

Taking up a point in section 2.9, there is a special provision in the conflict resolution mechanism whereby the more that the support for a within-pretence hypothesis relies on within-pretence facts and within-pretence reasoning steps (as opposed to imported facts, other results of outside-to-inside conversion rule applications, and reasoning steps outside the cocoon), the stronger it is taken to be.

8 Application of ATT-Meta to Anne/Kyle Example

Figure 5 shows the reasoning done by an application we performed of ATT-Meta to the Anne/Kyle example. Therefore, below, we were the "user" of the system. We have slightly modified the example by replacing "Anne knew" by "Anne believed," to get

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

This is because it makes no difference to the metaphorical aspects of the example, and because ATT-Meta merely handles knowledge as true belief.

8.1 User-Supplied Information about Specific Situation

The information we supplied about the specific situation to be reasoned about was as follows. We imagine the information to arise by preliminary processing of (18) and its context.

```
{certain} is-person(anne).
{certain} is-person(kyle).
{certain} married(anne, kyle).

IF  currently-within-metaphor-pretence
AND IN SURROUNDING CONTEXT: true
THEN {certain} has-far-reaches(mind-of(anne)).

IF  currently-within-metaphor-pretence
AND IN SURROUNDING CONTEXT: true
THEN {certain} physically-in(believes(anne, having-affair(kyle), certain),
                             the-far-reaches-of(mind-of(anne)) ).
```

Notice how the direct meaning of the metaphorical utterance is couched as the last two rules. They are designed for application only to queries within the pretence cocoon, as shown by the special condition `currently-within-metaphor-pretence`. In general, the conditions in such a rule establish subqueries within the pretence cocoon, except for conditions prefixed by "IN SURROUNDING CONTEXT: " which establish subqueries in the next reasoning context out, in other words the top context. In the present case those subqueries will be just the vacuous query `true`, which is always supported to full certainty. Overall, the effect of the two rules is to ensure that, within the pretence cocoon, the queries

```
has-far-reaches(mind-of(anne))

physically-in(believes(anne, having-affair(kyle), certain),
              the-far-reaches-of(mind-of(anne)) )
```

will be supported to level `certain`.

The function `the-far-reaches-of` is intended to be a *physical* domain function, and only properly applies to physical regions. The facts about Kule being a person and married to Anne are not actually used by the reasoning.

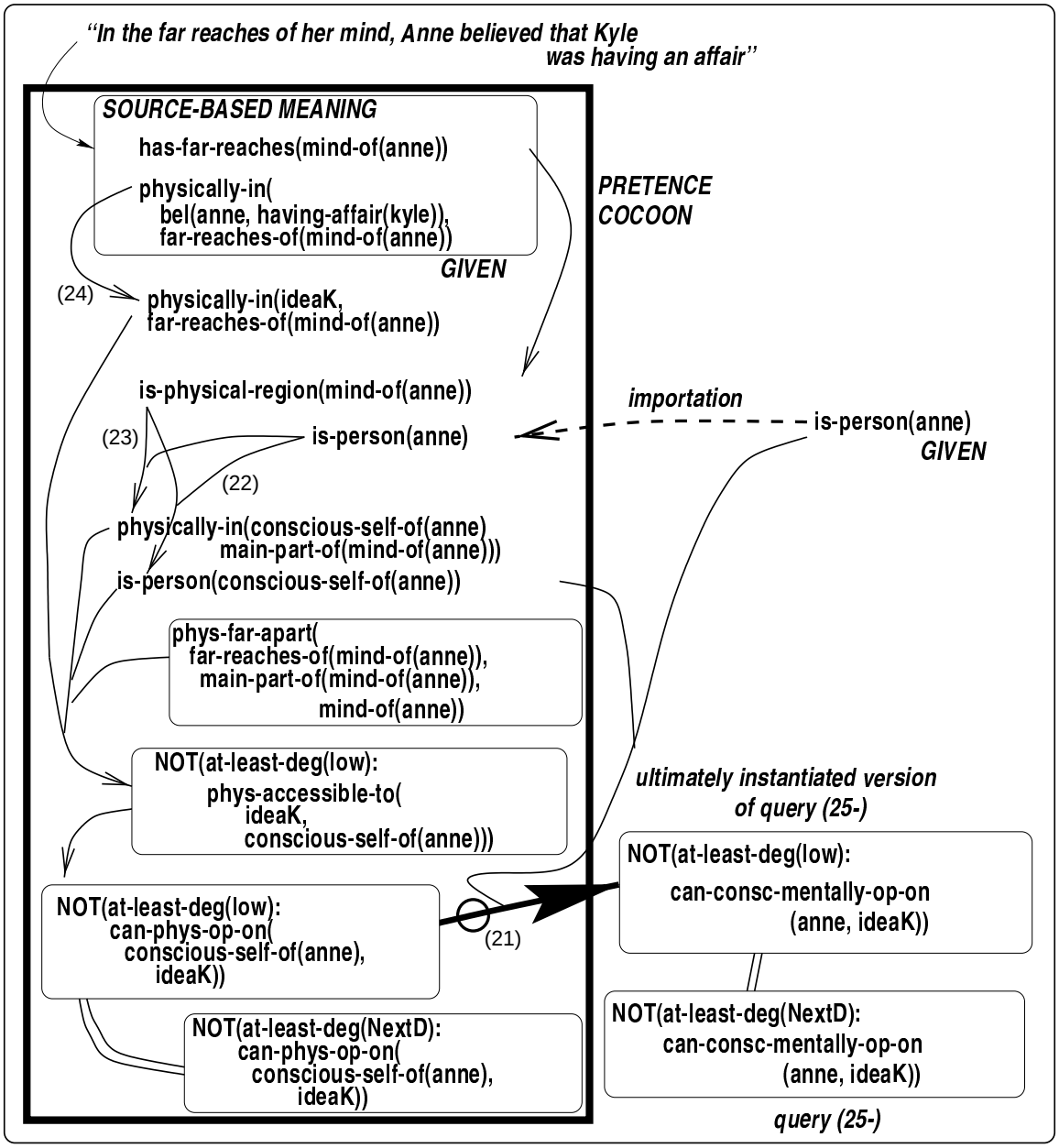

Figure 5: **Modified Anne/Kyle example, (18):** main features of the implemented ATT-Meta system's metaphor-based reasoning for the example. The logical notation is an abbreviated form of that used in the text. The symbol `ideaK` is an abbreviation for `the-idea-that(having-affair(kyle))`. Facts are marked as *GIVEN*. The whole of the source-domain meaning of the utterance is a collection of facts. All facts are `certain`. All other hypotheses shown attain a certainty qualifier of `presumed`. The fat arrow depicts the application of conversion rule (21). Applications of ancillary-assumption rules are numbered (22-24), following their numbering in the text (and see also their informal expression as (8) and (9).) Double lines connect general queries with their instantiations.

8.2 User-Supplied Query

We supplied the following top query to the system:

```
(19)   to-degree-exactly(D): can-consciously-mentally-operate-on(
                                anne, the-idea-that(having-affair(kyle))).
```

where `D` is a variable. We were effectively asking to what particular degree is the affair idea consciously entertainable by Anne. Recall from section 6 that in real discourse processing this query could arise in a demand-driven way from the processing of the discourse as a whole.

ATT-Meta allows properties and relationships to hold to various different positive qualitative levels: `very-low, low, medium, high, very-high, absolute.` There are various operators such as `to-degree-exactly` and `to-at-least-degree`, and various simple rules for handling them in ATT-Meta.

8.3 User-Supplied Rules about Mental States

This and following subsections present the main knowledge rules that we supplied, starting with two rules about mental states, then going on to some common-sense rules about physical objects and space, and finishing with metaphor-related rules. We only show rules that are *strongly used* in the reasoning (i.e. deliver a certainty contribution of at least `presumed` to some hypothesis), together with a few non-strongly-used ones because they round the picture out a little more. We do not show all the rules that were strongly used.

```
{certain} is-mind(mind-of(P)).
{certain} is-idea(the-idea-that(P, X)).
```

These fact rules characterize the nature of the outputs of the functions `mind-of` and `the-idea-that`. They apply to all values of variables P and X. The next three rules relate conscious mental operation to mental operation in general, on an idea J by a thinker P. The second is a quasi-*contrapositive* of the first. The third is a quasi-converse of the first, but is merely a default rule. The second is not strongly-used.

```
IF  is-person(P) AND is-idea(J)
AND to-degree-at-least(Degree): can-consciously-mentally-operate-on(P, J)
THEN {certain}
     to-degree-at-least(Degree): can-mentally-operate-on(P, J).

IF  is-person(P) AND is-idea(J)
AND NOT(to-degree-at-least(Degree): can-mentally-operate-on(P, J))
THEN {certain}
     NOT(to-degree-at-least(Degree):
                            can-consciously-mentally-operate-on(P, J)).
```

```
IF  is-person(P) AND is-idea(J)
AND to-degree-at-least(Degree): can-mentally-operate-on(P, J)
THEN {presumed}
     to-degree-at-least(Degree): can-consciously-mentally-operate-on(P, J).
```

8.4 User-Supplied Rules about Physical Objects and Space

The following physical-domain rules are certainly cruder and less numerous than the rules that would be used in a fully realistic system, but serve to illustrate the ATT-Meta style of reasoning. We stress that they are *designed for ordinary, commonsense reasoning about physical matters, and would be of the sort needed independently of the need to do any metaphorical processing.* They are in no way tailored to the task of performing metaphorical reasoning. Analogous points apply to the mental state rules in the previous subsection. We have not listed rules that have nothing to do with the type of physical situation operating in the example (such as rules saying that buildings have rooms). However, these rules were present in the experiments performed.

```
IF physically-in(O, R)       THEN {presumed} is-physical-object(O).
IF physically-within(R1, R2) THEN {presumed} is-physical-region(R1).
IF physically-within(R1, R2) THEN {presumed} is-physical-region(R2).

IF  physically-in(O,R) AND physically-within(R,S)
THEN {presumed} physically-in(O,S).
```

The fourth rule here is a simplification of what we actually provided. It collapses two actual rules, specialized as to whether the expression bound to S contains variables. This division is needed for minor technical reasons.

```
IF  is-physical-object(O) AND is-person(P)
AND to-degree-at-least(Degree): can-physically-operate-on(P, O)
THEN {presumed} to-degree-at-least(Degree): physically-accessible-to(O, P).

IF  is-physical-object(O) AND is-person(P)
AND NOT(to-degree-at-least(Degree): physically-accessible-to(O, P))
THEN {presumed} NOT(to-degree-at-least(Degree): can-physically-operate-on(P, O)).

IF  is-physical-object(O) AND is-person(P)
AND physically-in(O, R1) AND physically-in(P, R2)
AND are-far-apart-within({R1, R2}, S)
THEN {presumed} NOT(to-degree-at-least(low): physically-accessible-to(O, P)).

IF has-far-reaches(S)
THEN {certain} physically-within(the-far-reaches-of(S), S).

IF has-main-part(S)
THEN {certain} physically-within(the-main-part-of(S), S).
```

```
IF has-far-reaches(R) THEN {presumed} has-main-part(R).

IF  is-physical-region(S)
AND has-far-reaches(S) AND has-main-part(S)
THEN {certain} are-far-apart-within(
                 {the-far-reaches-of(S), the-main-part-of(S)}, S),
```

Note how the notion of being-far-apart is relative to a surrounding region S. What counts as far-apart depends on the physical context at hand.

8.5 User-Supplied Conversion Rules for IDEAS AS PHYSICAL OBJECTS

The metaphorical correspondence (10) in section 5.1 was captured in four conversion rules that we supplied:

```
(20)
   IF  is-idea(J) AND is-person(P)
   AND WITHIN PRETENCE: is-physical-object(J)
   AND WITHIN PRETENCE: is-person(conscious-self-of(P))
   AND WITHIN PRETENCE: to-degree-at-least(Degree):
                         can-physically-operate-on(conscious-self-of(P), J)

   THEN {presumed} to-degree-at-least(Degree):
                         can-consciously-mentally-operate-on(P,J).

(21)
   IF  is-idea(J) AND is-person(P)
   AND WITHIN PRETENCE: is-physical-object(J)
   AND WITHIN PRETENCE: is-person(conscious-self-of(P))
   AND WITHIN PRETENCE: NOT(to-degree-at-least(Degree):
                            can-physically-operate-on(conscious-self-of(P), J))

   THEN {presumed} NOT(to-degree-at-least(Degree):
                             can-consciously-mentally-operate-on(P,J)).
```

These are accompanied by converse rules for each, mapping mental operability in reality, or lack of it, to physical operability or lack of it inside the pretence. The two rules displayed are intended for application in the top reasoning context, i.e. outside the pretence cocoon. The special form "WITHIN PRETENCE: ϕ" queries ϕ within the pretence cocoon as opposed to the top context. The resulting certainty level for ϕ is what is used in the computation of the rule's contribution to the certainty of the query it is applied to. Recall that this contribution is the minimum of the rule's own level and the levels acquired from the conditions, so that the contributed certain is no more than that obtained for ϕ. We therefore achieve a crude implementation of the Uncertainty VNMA in section 2.12.

The explicit inclusion of degrees in conversion rules constitutes the current preliminary way of incorporating the desired effect of the Qualitative Degree VNMA.

8.6 User-Supplied Conversion Rules for MIND AS PHYSICAL SPACE

The correspondence (7) in section 5.1 was again captured in four conversion rules, the most important for our example being:

```
IF  is-idea(J) AND is-person(P)
AND WITHIN PRETENCE: is-physical-region(mind-of(P))
AND WITHIN PRETENCE: physically-in(J, mind-of(P))

THEN {presumed} to-degree-at-least(very-low): can-mentally-operate-on(P, J).
```

The other three rules are a quasi-contrapositive version of the one above and then quasi-converses for these two rules.

8.7 Metaphor Ancillary-Assumption Rules

We supplied the following rules, encapsulating ancillary assumptions (8, 9) for MIND AS PHYSICAL SPACE (recall section 5.1). They are designed for application only to queries within the pretence cocoon. Thus, the conditions establish subqueries within the pretence cocoon, except for conditions prefixed by "IN SURROUNDING CONTEXT: " which establish subqueries in the next reasoning context out, in other words the top context.

```
(22)   IF  currently-within-metaphor-pretence
        AND is-person(P)
        AND is-physical-region(mind-of(P))
        THEN {presumed} is-person(conscious-self-of(P)).

(23)   IF  currently-within-metaphor-pretence
        AND is-person(P)
        AND is-physical-region(mind-of(P))
        AND has-main-part(mind-of(P))
        THEN {presumed} physically-in(
                          conscious-self-of(P), the-main-part-of(mind-of(P))).

(24)   IF  currently-within-metaphor-pretence
        AND is-person(P)
        AND physically-in(believes(P, X, certain), L)
        AND physically-within(L, the-mind-of(P))
        THEN {presumed} physically-in(the-idea-that(P,X), L).
```

We have omitted showing a rule that is like (23) but that merely locates the conscious self in the mind of P as a whole, not assuming it has a main part.

In reading these rules, recall that `currently-within-metaphor-pretence` restricts the rule to being applicable only in the pretence cocoon.

8.8 Main Sequence of Reasoning Events

Because of built-in rules about qualitative degrees, ATT-Meta sets up the following two subqueries as a result of investigating the top query, (19), that was given in section 8.2:

```
(25+)   to-degree-at-least(D): can-consciously-mentally-operate-on(
                                    anne, the-idea-that(having-affair(kyle)))

(25-)   NOT(to-degree-at-least(NextD):
                                can-consciously-mentally-operate-on(
                                    anne, the-idea-that(having-affair(kyle))))
```

with the additional constraint that NextD should be the degree just above D. ATT-Meta will, in fact, eventually find the value `very-low` for `D` and `low` for `NextD`. Hence, the instantiation of (19) with D set to `very-low` is eventually produced with certainly level `presumed`. We will concentrate on (25-) as the more interesting subquery. (25+) becomes satisfied by virtue of the conversion rule for MIND AS PHYSICAL SPACE that was displayed in section 8.6, combined with the rule in section 8.3 that infers conscious mental operability from general mental operability by default.

Query (25-) matches the THEN part of conversion rule (21), binding rule variables P to `anne`, `J` to `the-idea-that(having-affair(kyle))` and `NextD` to `Degree` in the rule. Therefore, the following queries arise within the cocoon:

```
is-physical-object(the-idea-that(having-affair(kyle)))

is-person(conscious-self-of(anne))

NOT(to-degree-at-least(NextD):
       can-physically-operate-on(conscious-self-of(anne),
                                    the-idea-that(having-affair(kyle)))).
```

These are supported by backchaining through the within-cocoon reasoning chains shown in Figure 5. This reasoning involves both source-domain rules and ancillary assumption rules (22-24). When the support for the third query just above has been found, the variable NextD is instantiated to `low`. We therefore get within the pretence the instantiated version of that query as shown in the figure. This leads to the "instantiated version of query 25-" in the figure, by the completion of the application of the conversion rule (21). This conclusion is merely a default conclusion (i.e., it has certainty level `presumed`), for two main reasons: the conclusion within the pretence that the idea is very physically inaccessible is merely a default conclusion; and the conversion rule used is a default rule. Thus, information in the target domain could potentially defeat the conclusion.

The ancillary assumption rules need Anne to be a person within the pretence as well as outside. This explains the importation link shown in Figure 5.

Notice that the within-cocoon hypotheses that Anne's mind is a physical space, that Anne's conscious self is a person, and that this person is in the main part of the mind-space gain support *during* this backchaining. None of these, not even the first, are given at the start. In the same vein of on-demand processing, *the need for the cocoon itself arises dynamically through query-directed reasoning.* The cocoon is created when a rule is applied whose conditions contain a "`WITHIN PRETENCE:` ϕ" item. The query ϕ is then set up inside the cocoon (assuming conditions before it in the rule achieve some support).[16]

9 Evaluation of ATT-Meta System: I

Here we evaluate the ATT-Meta system, rather than our overall approach, partly by comparing it to the approach and partly by reporting on how well it performs.

9.1 Coverage of Approach and Examples

We commented in section 7.1 on various qualitative respects in which ATT-Meta falls short of completely implementing our overall approach as well as on respects in which it does capture the approach. We have elsewhere reported on the successful application of the ATT-Meta system to examples other than the Anne/Kyle and Small-Voice examples. (See section 10.3 for further information.) In section 8 above we showed that the system produces desirable contributions in the case of a quite complex example adapted from real discourse (6). We have done the same for the Small-Voice example (12) in Barnden & Lee (2001). On the basis of these two examples it is reasonable to make some specific assertions:

- ATT-Meta has the reasoning and representational power needed to cope with a wide array of MIND AS PHYSICAL SPACE examples. Specifically, the Anne/Kyle example suggests that any qualification that replaced "In the far reaches of her mind" but that still implied distance of the idea from the main part of her mind could be used by ATT-Meta to come up with the contribution that Anne had only a very low degree of conscious awareness of the idea, provided ATT-Meta was given the appropriate physical-domain rules by the user.
- Similarly, ATT-Meta has the reasoning and representational power needed to cope with a wide array of MIND PARTS AS PERSONS examples. The Small-Voice example suggests that any MIND PARTS AS PERSONS utterance that implied that two subpersons had contrary ideas could be used by ATT-Meta to infer that the real agent had two contrary tendencies. (Also, only one subperson, or voice, etc., need be mentioned in the utterance: ATT-Meta itself can infer the existence of at least one other. Normally, when someone insists something in a conversation, someone else has at least questioned it, and may have denied it. This point is backed up by Hunston's (1995) corpus study of the usage of the verb "insist".)

[16] Actually, in the data structures in the implementation, the cocoon only has a distributed, implicit reality, in the form of tags attached to hypotheses. A tag is a simple description of the context the hypothesis is in. The cocoon is not itself a data structure that has hypotheses as parts or contents in any explicit sense.

- The rules about physical space and conversations used in the two examples are of the sort that would be needed for common-sense reasoning about real physical matters and real verbal interactions. They are not specific to the needs of MIND AS PHYSICAL SPACE and MIND PARTS AS PERSONS.

- The various examples and considerations mentioned in this list suggest that a small number of metaphorical mapping relationships (each implemented as a small set of conversion rules) and a small number of ancillary-assumption rules provide great power in the handling of a wide array of complex examples.

A further word about the last point. We do not claim that the mapping relationships actually used in this article for the particular views considered are the only ones needed. However, only a small number of others appear to be needed. (This is in line with Grady's 1997, account of metaphor.)

9.2 Some Computational Statistics

We ran various experiments on the Small-Voice and Anne/Kyle and examples, (12) and (18) respectively. The former case is reported in Barnden & Lee (2001). We used the same set of user rules for both examples, aside from the fact rules depicting the particular situations at hand as described by the two different discourse chunks. The rest of this subsection discusses only the Anne/Kyle example.

The set of user rules included metaphorical mapping and ancillary-assumption rules for IDEAS AS PHYSICAL OBJECTS, MIND AS PHYSICAL SPACE and MIND PARTS AS PERSONS, and mini knowledge bases for the domains of verbal interaction, physical space/objects (section 8.4), and mental states (section 8.3). As a result, there were 62 user rules not counting the specific-situation fact rules, of which there were 6. In addition, at the time of experimentation ATT-Meta had 27 built-in rules, dealing with qualitative degrees, metaphorical pretence in general, belief states in general, and miscellaneous general inferential matters.

In the Anne/Kyle example, the system created 193 hypotheses in the course of reasoning. Of these, 59 attained a certainty level of at least `presumed.` Of the total 193, there were 164 in the pretence cocoon (46 of them at at least `presumed`).

Partly in order to check the overhead resulting from the inclusion of extraneous rules, we also ran the system after removing all except the 31 *strongly-used* user rules: a rule is strongly used if at some point it delivers a certainty contribution of at least `presumed` to some hypothesis. (But all built-in rules were kept even if they were not strongly used.) As a result, the number of hypotheses created dropped to 98. Of these, 53 attained a certainty level at least `presumed.` Of the total of 98, there were 75 in the pretence cocoon (40 of them at at least `presumed`).

Comparing the full run with the slimmed-down run, the 37 non-strongly-used user rules added 95 hypotheses, i.e. an average of about 2.6 hypotheses per rule. Without any user rules at all, 2 hypotheses

are generated, so that the 31 strongly-used user rules (including specific-situation fact rules) added 96 hypotheses, giving an average of 3.1 hypotheses per rule. Thus, the more irrelevant rules add a somewhat smaller number of hypotheses per rule (2.6 above) compared to the more relevant rules (3.1 above). This reduction is in the right direction but not as good as we would like. (The reduction is much greater in the Small Voice experiment reported in Barnden & Lee 2001.) However, the rate of 2.6 hypotheses per rule obscures some important detail. When only the three MIND PARTS AS PERSONS rules, none of which were strongly used, were removed from the full original set of rules (which led to 193 hypotheses), the number of hypotheses dropped to 146 (a reduction of 15.7 hypotheses per rule excluded). Thus the other 34 non-strongly-used user rules were only responsible for adding 146−95=51 hypotheses, representing a rate of only 1.5 hypotheses per rule.

As a further test of the effects of including more rules, we added the additional rules used in the various implemented examples in Lee & Barnden (2001b). The domain rules amongst these rules concerned possession of objects, fashion, resource usage, life and death, vision, physical constructions and personal relationships. The additional metaphorical-view rules (i.e., conversion rules and ancillary assumptions) concerned views of beliefs as possessions, mental operation as resource usage, idea as animate beings, cognizing as seeing, ideas as physical constructions and beliefs as locations. None of the 127 additional rules were strongly used. The effect of the addition was to increase the number of hypotheses generated from 193 to 405. The increase was therefore at a rate of 1.67 hypotheses per rule, consonant with the rate of 1.5 noted above. (The increase was much lower in the case of the Small-Voice experiment.)

These observations are encouraging, in that many of the additional rules were on subject matter strongly related to that of the original rule-set, and the added metaphorical mapping rules were related to the original ones in also being about mental states. However, we do not yet known how the effect will scale up with larger numbers of domains, rules per domain and metaphorical views.

When the three MIND PARTS AS PERSONS rules are removed from the implementation augmented with the additional rules from Lee & Barnden (2001b), the hypotheses dropped hugely from 405 to 237 (a reduction of 56.0 hypotheses per rule excluded). Comparing the figure of 237 with the figure of 146 resulting from excluding those three rules from the original experiment, we see that the 127 additional rules from Lee & Barnden (2001b) are adding hypotheses at the rate of merely 0.72 of a hypothesis per rule when the MIND PARTS AS PERSONS rules are excluded.

Clearly there are significant interactions between different subsets of rules, and we should consider reformulating rules such as the three MIND PARTS AS PERSONS that cause many added hypotheses even when not strongly used.

9.3 Some Current and Future Computational Short-Cuts

Much remains to be done to make ATT-Meta more economical in the way it operates. For example, apart from a current short-cut to be discussed, ATT-Meta tries to to apply all rules to all queries. Most attempts

will fail because the rule's THEN side will fail to match the query. Usually, the individual match operations are not computationally expensive because the mismatch is usually large, but it would nevertheless be advantageous to economize on rule-application attempts by, for example, pre-analyzing the rules to see which might conceivably be applicable in the situation at hand, and then hiding the others.

The short-cut that is included is that the only rules tried for a query are those whose certainty qualifier is at least as high as one of the certainty levels listed as being "of interest" for the query. As explained further in (Barnden 1998c), each query has a set of of-interest certainty levels that act as useful constraints on evidence gathering.[17] As a simple example, if the user is only interested in the top query if it can be established with absolute certainty, then only completely-certain rules are tried for it. Usefully-constraining desired levels also arise because there is a scheme for backpropagating of-interest certainty levels across rules to the subqueries they generate, and for setting the of-interest levels of negations of queries as high as possible.

A further short-cut is that, usually, if a proposition has been established with complete certainty, or is not established to at least some of-interest level, its negation is not investigated, whereas otherwise it may be.

A major source of inefficiency in ATT-Meta is the fact that it tries to find *all* arguments supporting a hypothesis that is being investigated, rather than curtailing the investigation when the highest of-interest level is reached. Relatedly, it seeks *all* supported variable bindings when the hypothesis has variables. The reason for these measures is that, in cases of conflict, it is dangerous to exclude possible arguments. Having said that, we plan to modify the system that it first investigates a hypothesis and its negation to the extent necessary to see whether there could be a conflict, and only does a full investigation when conflict is possible, otherwise it curtails as suggested. This change, though conceptually quite simple, is a complex one in programming terms.

9.4 Dispelling Real-World Interference with Pretence

Here we take up an aspect of the approach discussed in section 2.9. In the Anne/Kyle case, it is inferred within the pretence that Anne's mind is a physical space. But in our experiments on the example a rule that minds are not physical spaces was included (though not displayed in section 8). This rule, along with any other, is a candidate for being applied within the pretence. We therefore have a reasoning conflict. But the hypothesis that Anne's mind is a physical space relies on the pretence postulate that Anne believes something in the far reaches of her mind, whereas the hypothesis that Anne's mind is not a physical space does not rely on any pretence postulate. Therefore, by ATT-Meta's special conflict-resolution provision mentioned in section 7.4, the conflict is successfully resolved in favour of the former hypothesis. The conflict-resolution provision is described in more detail in Barnden (in press, a).

[17] Actually, in that article they are somewhat misleadingly called "desired" certain levels.

The inclusion of the knowledge that minds are not physical spaces, and the resulting reasoning operations, do not add much to the overall amount of computation for the example. Removal of the rule reduced the number of hypotheses in the Anne/Kyle experiment (with otherwise non-reduced rule set) mentioned in section 9.2 from 193 to 189.

10 Evaluation of Overall Approach

Here we summarize some evaluative aspects of the article so far, and present additional considerations. Our focus is on the overall approach, not the ATT-Meta system. A general comment is that although the number of specific examples discussed in the article is necessarily limited, the examples are quite diverse and should therefore be suggestive of the power of the approach.

10.1 Performance of Approach: Negative Points

Our approach is, obviously, only a partial approach to metaphor in general, in that it does not address the question of how entirely new metaphorical mappings are constructed by an understander, or how known mappings are extended when there needs to be an exception to our Map-Extension Minimization default. We conjecture that techniques of mapping-construction used in well-developed systems such as ACME, SME, IAM and MIDAS/MES could be added in a principled way to our approach, but that is something we have yet to do.

The processing in the approach (as opposed to the ATT-Meta system), though not absolutely relying on guidance from context as to what contributions to produce, is facilitated if it is available. We have no detailed account of how such guidance is produced from context. Nevertheless, this point is perhaps not particularly against our approach, because the selection of appropriate contributions is a notoriously difficult problem in metaphor research as a whole, and very little work of an algorithmically detailed nature has been done on how the understanding of a metaphorical utterance is to relate to the processing of context.

We have no complete or settled account of how an understander using our approach is to know, when faced with an utterance that may or may not be metaphorical, that it is appropriate to use a metaphorical pretence cocoon. Of course, detection of metaphoricity is generally recognized as a difficult problem in general (see, e.g., Kittay 1989). Considerable help with detection should be available from known heuristics such as detection of selection-restriction violations (Fass 1997) and detection of phraseological, lexical and morphological cues for metaphoricity (Goatly 1997), though it is well known that such heuristics are not a complete solution. We make a further comment on metaphoricity-detection in the Conclusion section.

We have no detailed account of how the processing advocated by the approach is to be combined with other semantic/pragmatic processes, and in particular with processing needed to cope with other figurative aspects of utterances such as irony, metonymy, exaggeration, etc. However, few researchers have attempted this. Notable exceptions are Hobbs and Fass in AI. Accounts such as Relevance Theory make important

contributions by discussing the phenomena in a general framework but do not provide a detailed, unified processing account.

10.2 Performance of Approach: Abilities I

In section 3 we showed how we could at least match the capabilities of other approaches that "neighbour" ours in holding back on map-extension, to some significant degree at least, or emphasizing on-the-fly source-domain reasoning. In section 4 we showed how representative examples that have been presented by authors of two *non*-neighbouring approaches (Fass meta5, and category-inclusion) could be handled in our approach. In addition, in both cases the authors presented the example as a case of novel mapping (novel, that is, to the understander) but we showed them to be better analyzed as cases of map-transcendence with respect to stored metaphorical views. In both section 3 and section 4 we addressed the informational contributions claimed by the relevant authors, irrespective of whether we ourselves agree that the contributions are appropriate, whether the authors provides any justification for the contributions, or whether the authors found or invented the examples. The important point is that the contributions were put forward independently from our own work.

In section 5 we showed how the approach could work on two examples involving altogether four important metaphorical views (IDEAS AS PHYSICAL OBJECTS, MIND AS PHYSICAL SPACE, MIND PARTS AS PERSONS, IDEAS AS INTERNAL UTTERANCES) featuring in our own databank of views of mental states. The two examples, Anne/Kyle and Small-Voice, are quite rich in terms of the reasoning needed, and provide useful demonstrations of the principles of our approach. Barnden (2001b) shows how the approach works on further representative examples from the databank.

The databank consists entirely of examples from real discourse we have encountered. That meant we had to make our own subjective evaluation of what contributions are appropriate and important, but we buttressed this in the Anne/Kyle example by an examination of context in section 6 (the Small-Voice case is dealt with in Barnden & Lee 2001). We conjecture that the set of views in the databank accounts for most cases of metaphor for mental states, though we are not in a position to substantiate this by statistical analyses of corpora, given that corpus-based study of metaphor is in its infancy.

10.3 Performance of Approach: Abilities II

In Lee & Barnden (2001b), we applied the ATT-Meta system itself to the majority of the metaphorical views of mental states/processes in Lakoff's Master Metaphor List, including all the ones which explicitly mention "mind" or "idea" in their names (such as IDEAS ARE CHILDREN). We took one example of each view considered, by doing our own corpus search (on the COBUILD Bank of English). Therefore, we had to suggest informational contributions for ourselves (and anyway the List is sparse on these). The successful application of the approach to producing the contributions gives us some confidence that it is on the right track.

Also, as recounted in Barnden (2001a), we have recently applied the approach (though not yet ATT-Meta itself) to a suite of examples from Goatly (1997). These examples are quite complex and account for many of the clearly map-transcending examples that Goatly gives.

In the following subsections we point out that the approach is able to cope with difficulties that would crop up in other approaches.

10.4 Complexity of Informational Contributions

Here we make a point that bears against many approaches in the literature, though the advantage is not so much to our particular approach but rather to any approach that allows target-domain contributions to be as complex as ours does. Such approaches include Hobbs', SME, ACME, and Lakoff's.

First we take feature-transfer approaches (e.g., Ortony 1979). The complexity of most of the contributions in Section 5 is considerably beyond what is envisaged in these approaches. The usual sort of features discussed are ones like being cunning, (e.g., from "John is a fox"). These are nowhere nearly as complex as contributions such as the contribution that the agent tends to believe that she wants to marry Mick (in the small-voice example).

Analogously for the categorization approach (Glucksberg & Keysar 1990). That approach has obscured the fact that the confiningness attributed to the job is really a relationship, not a unary property. A job or jail is confining *for someone,* and one could say something like "Tom's job is a jail for people less assertive than he". So "the category of confining things" is too simple a tool in the first place. One needs a good theory of categories of multi-entity *situations,* not just categories of single entities. Moreover, the particular possible contributions we came up with for the job/jail example (e.g., Tom cannot partake in activities that are not part of his job) do not sit naturally with a category approach.

10.5 Unaccounted-For Transfers

We mentioned earlier that Lakoff provides no account of why the excitement for the travelers in the freeway example maps to excitement for the lovers. Our approach includes a VNMA to handle such things. Similarly, in the psychoanalyst mansions example, Grady (1997) does not comment at all on the mention in the discourse of some Freud followers *wanting not* to live in the main house, and therefore wanting not to have the main house as their home. Yet, the mental state of wanting is not mapped by any of the metaphorical views mentioned by Grady. In our approach, the wanting-not can be transferred to the target domain by the Mental/Emotional States and Negation VNMAs, since the having-as-home is already mapped to espousing. Thus we get the target-domain contribution that the followers wanted not to espouse the thought-system that is the "main house." Grady also does not comment on significance of the the adjective "main." In our approach, main-ness maps over by virtue of the Value-Judgment VNMA.

Again, the Hobbs and Martin approaches provide no mechanism for automatically providing the types of transfers discussed so far in this subsection. Indeed, all approaches that lack either some equivalent of our VNMAs or something like the ability of ACME (as expanded by a transfer facility as in Holyoak *et al.* 1994) and SME to transfer unmapped structure in general would have difficulty with the above types of transfer unless suitable mappings of mental states, etc., were included in particular metaphorical views.

We should also note that some of the VNMAs are about types of transfer that have been given very little explicit attention in the literature despite their evident importance. These include the VNMAs for mapping Modality, Uncertainty, Qualitative Degree, Negation, Set-hood and Set-Size.

10.6 Richness and Extent of Non-Parallelism between Source and Target

MIND PARTS AS PERSONS examples, and notably the Small Voice example, vividly illustrate the problems approaches other than ours would have. This is because of the richness of the structure posited in the source domain and the substantial lack of parallelism of that structure with any target-domain structure that could reasonably be posited to feature in commonsense knowledge about the mind. The rich source-domain structure consists of the various subpersons together with their individual mental/emotional states and their intercommunications. Little of this structure is plausibly to be mapped to conjectured structure in the target—it is merely that subperson mental/emotional states map to whole-agent tendencies towards those states. The division into subpersons does not map to real components (physical or abstract) of the agent; and the internal utterances, as *relationships* between subpersons, do not map to *relationships* concerning the whole agent. We can say that an internal utterance, to the effect that X, that is made by one subperson to another maps to the agent's occurrent thought that X, but this preserves only part of the relational structure being mapped.

Therefore, MIND PARTS AS PERSONS examples are a problem for approaches such as Fass' (1997), Veale & Keane's (1997), SME and ACME that are centered on finding or positing parallelism of structure between source and target. We now make some specific comments about SME and ACME.

Let us assume for the moment that SME has already established some mapping from the source to target domain. Recall Figure 4 and consider the task of transferring the proposition that one subperson believes that the whole agent wants to marry Mick. For one thing, the "believe" relationship would have to map to "has some reason to believe," which is a mapping connection different from any apparent in writings on SME (especially as writings on SME generally require predicates to map identically). But also there is the question of what the subpeople would map to. They need to map to something, and indeed to entities capable of having mental states, because of the mapping-consistency constraint in SME. One would either have to posit that the whole agent really does have subagents, so that different source-domain subpeople would map to different target-domain subagents, or one would have to make the different subpeople all map to the whole agent, in which case the 1-1-ness of mappings that is required in SME would be contravened.

Also, notice that the mapping actions in Figure 4 are warranted in part by the direct meaning of the

metaphorical utterance. That meaning maps to an unanalyzed event outside the pretence cocoon (such events is not shown in the Figures). But such mappings of a complex to an atomic entity are not countenanced in SME. Indeed, SME would presumably try to find some complex structure in the target-domain to act as a parallel for the source-domain proposition. The system would therefore have to commit to some mappee for the small voice itself, whereas our approach does not.

Published examples of SME require transferred source-domain substructures to be linked to source-domain structures that already have a detailed mapping to the target. Again, in Figure 4 the only linking is to the direct meaning of the utterance, which does not have a detailed mapping to the target.

The problems for ACME are lesser because it does not firmly impose 1-1-ness on mappings (it merely prefers it), can tolerate failing to find target-domain mappees that would ideally be needed for source-domain items, and is open to non-identity mapping of predicates. But the very fact that ACME, and SME, *always attempt* to find parallels for source-domain structure, quite apart from the question of whether they can tolerate failure to find it, is a deficiency of those systems as applied to map-transcending uses of stored metaphorical views. Such attempts will often result in much wasted computational work. Notice that this work is over and above the work our approach requires, not just a substitute for some of the work in our approach. A system based on ACME or SME would *still* need to draw the within-source-domain connections that we illustrate in Figure 4 if it was to come up with the contributions shown.

The Anne/Kyle far-reaches example, resting on MIND AS PHYSICAL SPACE, would also lead other approaches to waste effort. Systems such as Fass's, Veale's, SME and ACME would inappropriately try to to find a mappee for the far reaches in the target domain. (In the SME and ACME cases we assume for the sake of argument that a source-to-target mapping is already known, but that it does not map the notion of far reaches.) SME would also have the above problem that the transferred source-domain propositions are not linked to source-domain structure that themselves have a detailed mapping to the target.

However, this subsection does not substantially bear against the approaches of Hobbs, Lakoff, Narayanan, or perhaps Martin. The doubt in Martin's case is that it would wastefully try to find target-domain structure parallel to source-domain structure in certain cases (e.g., when causation relationships are at issue).

10.7 Richness of Inference

We place no restrictions on the richness and extent of the reasoning that can take place within the pretence cocoon. To our knowledge, only Hobbs's approach is comparable to ours in this respect. When we note the types of reasoning required in the examples in sections 3, 4 and 5, buttressed by the detailed Anne/Kyle experiment (section 8 and Figure 5) we see that it it would be dangerous to apply any particular constraint on the reasoning. We do not see how other approaches could handle the examples without the ability to create within-source-domain linkages of much the sort that we present above. Of course, some constraint on total amount of reasoning may be cognitively necessary for all utterances, whether metaphorical or not.

After Hobbs, the approach next closest to ours is Narayanan's (1997). Although to our knowledge he

does not explicitly restrict the extent of source-domain reasoning, his approach is partially connectionist and is (at least currently) constrained by the qualitatively limited types of reasoning that are readily achievable in connectionist frameworks. This is so even in localist connectionist frameworks such as the one he uses, or more complex localist frameworks such as that of Shastri & Ajjanagadde (1993).[18]

Martin's scheme, while close in spirit to ours in several ways, appears to place very tight restrictions on the within-source-domain connections drawn. These are fine for the examples he addresses but would be too restrictive for the Anne/Kyle example, the freeway-of-love example and the Small-Voice example, amongst others.

As we noted in the case of ACME and SME, systems such as meta5 and Veale's appear to make no provision for performing within-source-domain reasoning of any complexity. A provision could no doubt be added, but would have a major effect on their overall handling of metaphor.

10.8 Pretence Cocoons, Nested Pretence and Metaphor Compounding

The difficulties listed so far in this section do not bear against Hobbs' approach, and bear only in a limited way against Narayanan's or Martin's (or Lakoff's, to the extent that Lakoff can be said to have specified a processual approach at all). But in this subsection we do present considerations that bear against these systems.

Hobbs has nothing that corresponds to our pretence cocoons, and does not mark pretended pieces of information as being such. Essentially, he mixes what we would put into the pretence-cocoon in with the target-domain information. This, not surprisingly, can lead to inferential difficulty. For example, suppose the system knows that (in reality) minds are certainly not physical spaces. Then any inference in the Anne/Kyle example that Anne's mind is a physical space would be defeated by that certain knowledge. This would then prevent the application of mappings that apply only when a mind is viewed as a physical space, and source-domain reasoning that relies on the mind being a physical space. What one needs is a way of protecting the pretended source-domain information from interference from reality-based information. This effect is achieved by the pretence cocoons in our approach.

Also, multiple cocoons may be needed. For example, metaphor might be used in quoted speech within a newspaper article whose text is independently metaphorical; or there might be utterances close by in discourse that discuss different target domains. In Lee & Barnden (2001a) we apply our approach to some types of metaphor compound by having different cocoons for the different metaphorical views involved.

Although the approaches of Martin and Narayanan do separate the source-domain information from the target-domain information, and can therefore be said to have something akin to pretence cocoons, they have not addressed the question of how to modify their system to allow for (their equivalent of) multiple cocoons, possibly nested. Also, Narayanan's system, because of its connectionist basis and separation of

[18] But Narayanan's scheme could be made inferentially more powerful by using the our own implemented connectionist reasoning framework (Barnden 1995).

the source and target domain reasoning into two different subnetworks, would have difficulty accounting for the dynamic creation and removal of cocoons in response to the needs of individual utterances.

To summarize subsections 10.4 to 10.8, our approach can deal with examples that present major difficulties to other approaches. Each other approach considered in this article was found wanting in some respect or another. The approaches that came out most unscathed were, not surprisingly, those of Hobbs, Narayanan and Martin, in decreasing order of unscathedness. However, even these approaches have some disadvantages compared to ours, and the research surrounding them has not teased out the general principles that we have isolated. We also reiterate that our approach is more liberal on the types of source-domain reason allowed than Narayanan's or Martin's. As for Hobbs' approach, it has not been implemented in a working system, whereas ours (in common with those of Martin and Narayanan) has been.

10.9 Potential for Unwarranted Informational Contributions

The MIDAS example (3) of John "giving" a cold to Mary raises the question of whether John still has the cold afterwards. Normally we would assume he does, despite the fact that, in the source domain of possession-of-objects, we might think that giving implies non-possession afterwards by default. Such a non-possession hypothesis would be mapped by the Negation VNMA in our approach to become the hypothesis that John is not infected after causing Mary to be; and this hypothesis would be natural under the MIDAS approach as well. But it would be undesirable for it to stand as a conclusion.

ATT-Meta is equipped to apply conflict resolution between the hypothesis and target-domain knowledge that causing-infection implies that the causer still has the infection. However, as ATT-Meta's conflict resolution mechanism stands at present, the John-not-infected-afterwards hypothesis could defeat the target-domain knowledge if the latter is also uncertain. However, we make the following observations.

If there were a `certain` rule that states that people still have infections even after causing them in other people, then in ATT-Meta this rule would indeed defeat the John-not-infected-afterwards hypothesis. It is only when the target-domain knowledge is uncertain that the problem will arise. The question of direction of defeat (override) when target-domain information is itself uncertain has hardly been discussed in the metaphor literature.

And it would be wrong in general to have a simple principle that says that target-domain knowledge should win conflicts with informational contributions arising from metaphorical mapping actions. This point requires much more discussion than is possible in this paper, but metaphorical utterances should often defeat target-domain defaults. Consider, for instance, the possible variant *"My car doesn't drink gasoline, it drinks methane"* of (4). We wouldn't want the informational contribution of this sentence, that the car uses methane instead of gasoline, to be defeated by a default target-domain rule that cars use gasoline.

It may be that the word "give" or the concept of giving is somehow special compared to possession-transfer in general. Consider *"John transferred / passed on his cold to Mary"* and *"John made a gift of his cold to Mary."* The John-not-infected-afterwards hypothesis *is*, we would claim, appropriate as the winner

in such variants. If the concept of *giving* can be considered exceptional, then it would be possible in ATT-Meta to have rules that handle this exception specially and override the normal John-not-infected-afterwards conclusion.

Notice, in fact, that *ordinary giving OFTEN does NOT imply subsequent non-possession.* One sense of "X gave Z to Y" where Z is something copiable (or multiply realizable), such as an academic paper or wallpaper pattern, is that X gave Y a copy of Z. So, if the underlying metaphorical view is one of DISEASE AS *COPIABLE* POSSESSED PHYSICAL OBJECT, rather than just DISEASE AS POSSESSED PHYSICAL OBJECT, the non-possession inference would not arise in the pretence cocoon in the first place, or would arise as merely a suggestion there rather than a default. In either case, ATT-Meta would end up with the default conclusion that John was still infected, because of target-domain knowledge. Note also that ATT-Meta could be given an ancillary assumption that states that when a disease is viewed as a possessed physical object it is copiable, and the no-possession-afterwards rule could apply only to non-copiable things.

We now turn to a related potential problem for our approach. It may seem to be in danger of overgenerating conclusions about mental and emotional states because of the VNMA for such states. Giving is normally deliberate, so that our approach would give the default conclusion that John deliberately caused Mary to be infected with the cold. Target-knowledge that infection-causing is usually unintentional would conflict with the inferred deliberateness, but then we again have the problem of ensuring the conflict is resolved in the desired way. It would be resolved as desired by ATT-Meta in some but not all circumstances. Again, giving may be an exceptional form of possession transfer, because the variant sentences mentioned above using "transferred" and "made a gift of" do seem to us to connote deliberateness.

MIDAS itself does not create a deliberateness conclusion, but that is merely because the system does not have any general approach at all to handling mental states that are inferred in the source-domain.

10.10 Concerning Psychological Evidence

We have not sought to generate detailed psychological predictions from our approach, and we do not believe that experiments reported in the literature so far work strongly for or against the approach. However, since we are interested in the possibility of the approach leading eventually to detailed psychological processing models, we have considered the connection to psychological experiments. Most importantly, we have discussed elsewhere (e.g. in Lee & Barnden 2001a) how our approach connects to the debate about the extent to which metaphorical understanding involves the construction of a "literal" meaning of the utterance. For review and discussion of relevant experimental results, see Gerrig (1989), Gibbs (1989, 1990, 1998), Gineste & Scart-Lhomme (1998), Lytinen *et al.* (1992), Onishi & Murphy (1993), and Récanati (1993). Although some experiments have been taken to show that literal meanings are not constructed, some recent work (Brisard *et al.* 2001, Giora 1997, Giora & Fein 1998, Honeck *et al.* 1998) throws serious doubt on the claim that literal meanings are avoided.

Assuming for the sake of argument that literal meaning and our direct meaning are the same, our ap-

proach does require the construction of a literal meaning, placed in the pretence cocoon (although this construction can be at the level of a clause rather than a whole sentence). But, as we explain in Lee & Barnden (2001a), this does not run foul of the arguments put forward by some authors to the effect that the computation of literal meanings is denied by experimental results. For one thing, our approach does not assume any *rejection* of the literal meaning before metaphorical meaning can be extracted, and for another we argue that much of the processing we advocate *replaces*, rather than adding to, processing that would be done if the literal meaning were the correct one.

11 Further Discussion

11.1 Metaphorical Views Are Not Explicit Entities

Metaphorical views are not identified as such by the approach or system, and metaphor-related rules (conversion rules and ancillary-assumption rules) are not allied to particular views except by their contents. Rather, views are implicit and distributed across the rules. For instance, several rules in sections 8.6 and 8.7 include the condition `WITHIN PRETENCE: is-physical-region(mind-of(P))`. As a result, these rules can only apply if it can be shown within the pretence that P's mind is a physical space. To that extent the rule implicitly forms part of the MIND AS PHYSICAL SPACE view. Moreover, a rule can in this way form part of several different views.

Moreover, such conditions can become satisfied in the middle of metaphorical reasoning, rather than at the beginning. Indeed, in the Anne/Kyle example, the hypothesis that Anne's mind is a physical space arises only indirectly, from the postulation that her mind has far reaches. And the hypothesis never achieves full certainty. Thus, it is not even the case that the the cocoon needs to rely on definite decisions about the metaphorical nature of the relevant entities.

11.2 The Prevalence, Utility and Dangers of Map-Extension

Our Map-Extension Minimization should not be construed as a denial that map-extension is often useful. But one point is that the usefulness can be for ends *beyond* those of merely understanding the utterance at hand. For instance, consider the learning of a new giving-to-infecting mapping relationship by the MIDAS/MES system (section 3.2). This is arguably a useful thing to do *from the point of view of understanding later utterances.* But it has no beneficial effect on understanding the utterance currently at hand. We believe it is useful to distinguish processing needed for current understanding from processing that could be useful for future understanding, simply because if the latter processing is computationally expensive (as indeed the discovery of new mapping relationships can be) it is undesirable to *assume* it is needed in current understanding.

In the MIDAS example the precursor of establishing the giving-to-infecting mapping relationship is finding that infecting is something that causes the state of being infected. The question is whether our

approach is wrong not to *oblige* the understander even to effect this precursor. Note carefully that our approach *allows* the understander to effect it. Indeed, ATT-Meta is fully equipped to be able to use a target-domain rule that said that if someone X causes someone Y to be infected with Z then X infects Y with Z. This rule can be either a certain rule or a default rule (i.e. one with strength `presumed`). If there was a need to know that John *infected* Mary as opposed to knowing that he *caused Mary to be infected* then the system could indeed produce that conclusion. The point is that we do not want to *oblige* the understander to do more work than necessary. Also, there could be many inferences one could make in the target domain from John causing Mary to be infected. Which of these inferences should be done? Which should count as part of the understanding of the utterance?

In addition, it would be appropriate for the rule just imagined to be merely a default rule if "infecting" means something more special than just causing infection, such as a *normal* way of causing infection, e.g. by touching them. But it is then incumbent on the designer of a system such as MIDAS to do provide a way of handling exceptions. For instance, perhaps John made Mary touch some other infected person, and it is this action of John's that caused Mary to be infected. But would we want to say that "John infected Mary"? Martin does not consider issues such as this.

It might be objected that the exceptional situation just mentioned would not normally be described as John "giving" a cold to Mary, so that in fact MIDAS would be correct in uncritically producing the conclusion that John infected Mary in a normal way. But it is just luck that the system conforms to the idiomatic use of "giving" here. The system would have operated in much the same way on the sentences "John transferred / passed on his cold to Mary", "John made a gift of his cold to Mary," on the assumption that the system views transferring and making-a-gift-of as causing possession, *unwarrantedly* concluding that John infected Mary in the normal way. Also, MIDAS would learn a mapping between transferring or making-a-gift-of and infecting-in-a-nomal-way; this mapping could mislead, or at least cause pointless interference with, later understanding.

The issue here is that "giving [a disease]" may in fact have as a direct, idiomatic sense the notion of (probably unintentionally) infecting-in-a-normal-way. Our whole discussion of the example has been on the basis of the understander not having such as direct sense, in order to conform to Martin's discussion. Naturally, if the understander *does* possess that special direct sense, then under our approach the conclusion that John infected Mary in the normal way would of course arise, over and above any more general inference that might be made by the lengthier route already discussed.

We return also to the default deliberateness conclusion mentioned in section 10.9. The above idiomatic direct sense could, under our approach, lead to the blocking of the deliberateness conclusion, given that that direct sense includes a *non*-deliberateness default. The deliberateness conclusion in the "made a gift" and "transferred" variants of the example *does* seem plausible to us, indicating to us that the mental-state VNMA is indeed an appropriate *default.*

In the case both of the drinking-car example and of the job/jail example (sections 4.1 and 4.2), our analysis was on the basis of a stored metaphorical view, whereas the authors concerned treated the examples

as being based on views not stored in the understander. We suspect that many other examples so treated in the literature are actually better handled as manifestations of stored views. A case in point is the Socrates-as-midwife example (from Plato's *Theaetetus*) discussed by Holyoak & Thagard (1989). Holyoak and Thagard present this as an elaborate application of the ACME system for map-discovery, even though Plato's text explicitly spells out most of the assumed mapping relationships in detail (which is just as well because, as Plato himself makes plain, some distort the reality of the midwife domain). The remaining mapping relationships, as well as some not discussed by Holyoak, arise from the VNMAs. It is reasonable to wonder whether any real text that was not setting out to be obscure or cute would merely state that Socrates is a midwife without considerable explicit elaboration or without rich context that implies particular mapping relationships more directly than the statement "Socrates is a midwife" does.

An example of a map-transcending utterance that would probably require map-extension for proper understanding is "Sally was unhappy with the basement of John's theory." This invented example rests upon the view of THEORIES AS BUILDINGS (Lakoff & Johnson 1980; Grady 1997). An understander would probably need to search for a way of mapping the mentioned basement, and perhaps the basement notion in general, in order to know what it is that Sally was unhappy with. The question of the circumstances under which map-extension is in fact needed is a matter for future research. In this article we present an approach that has no provision at all, currently, for extending mappings but yet can deal with broad, general types of map-transcendence.

12 Conclusion

We have provided an approach to the handling of map-transcending metaphorical utterances, and an implemented system capable of handling the core of the reasoning involved in dealing with such utterances, as long as no map-extension is needed. We have argued, and demonstrated by example, that rich metaphorical understanding often does not require map-extension, and that avoiding map-extension is a reasonable default. We have shown in this report and accomapanying ones (Barnden 2001a,b, Barnden & Lee 2001, Lee & Barnden 2001a) how the approach applies to a variety of existing, salient examples in the literature, to examples from an authoritative, independently produced databank (the Master Metaphor List), and to two complex examples from our own databank of real discourse. We have demonstrated (here and in related technical reports) that the implemented ATT-Meta system works on those two examples and others. We have also shown how our approach deals with difficulties that would face other approaches,

In place of map-extension, we have emphasized reliance on within-pretence (and mostly within-source-domain) reasoning that can be extensive. But it would be wrong to infer from this that we are actually adding work that would not otherwise be necessary. The within-pretence reasoning is in the service of drawing linkages between the direct meaning of the metaphorical utterance and the already available metaphorical mapping relationships. But an approach based on map-extension would itself need, in general, to uncover the linkages, in order at least to ensure that the new mapping relationships cohere with the existing ones, and

perhaps also as a way of guiding the search for the extension. Another point about the amount of reasoning in our approach is that it does not preclude stock metaphorical phraseology having direct meanings in the target domain, so that the pretence-based reasoning processes we have described are sidelined. For example, although "in the far reaches of her mind" might receive the type of processing we have proposed and have no direct meaning in the mental domain, "in the back of her mind" could have a direct meaning such as "in a mentally subsidiary way."

The approach contrasts with analogy systems such as SME and ACME that have been applied to metaphor. One main point of contrast is that in their work a piece of source information is transferable only if strongly linked (typically by causal or inferential relationships) to source structure that has been put into detailed item-by-item correspondence with target structure. In our approach, a piece of source information is transferable merely if it matches conversion rule conditions and is inferentially linked within the pretence to the direct source-domain meaning S of the metaphorical input. We pointed out that the linkage may use general knowledge, mainly about the source domain, that itself has no parallel with anything in the target domain. Also, S can be taken to correspond to *something unknown* in the target domain, but that unknown something is not analyzed into parts that are put into item-by-item correspondence with parts of S.

Our approach diverges from many others by not reifying metaphorical views or domains. Instead, on the question of views, metaphor-related rules contain conditions of application that make them suitable in an opportunistic way for particular views that we could theoretically identify. On the question of domains, the approach takes on board the observations that individuation of domains is a highly subjective and context-sensitive matter and that source and target domains in metaphor can overlap (so that it is misleading to say baldly that metaphor involves mappings between distinct domains).

As a particular consequence of the stance on domains, the pretence cocoon can mix together the reasoning from any collection of domains, and in particular from the target and source domains. Therefore, the approach and ATT-Meta system stand to provide explications of aspects of the theory of conceptual blending from cognitive linguistics (Fauconnier & Turner 1998), whether or not the blending is in the service of metaphor. Indeed, Barnden (1998d) reported an application of ATT-Meta to the two-ships example, an important illustration in existing accounts of blending. We have, relatedly, applied the system (without change) to a style of counterfactual reasoning (Lee & Barnden 2000), and wish in future to use ATT-Meta to study cases of mixtures of metaphor and counterfactuality.

Major gaps exist before the approach (let alone the implemented system) could be said to be a complete account of metaphor. More work needs to be done in meshing the approach with the rest of the overall process of discourse-understanding, Also, work needs to be done on the question of preliminary processing of a metaphorical utterance coming up with the decision that the utterance is indeed metaphorical, so that the direct meaning is placed within a pretence. In fact, we aim to study a proposal whereby that decision is *not* made until it becomes demanded by the backwards process of query-directed reasoning. Until that point the direct meaning would not be assigned either to the pretence or to the top (i.e., reality) space). Of course, in many cases, heuristic tools such as detection of selection-restriction violations (Fass 1997), or

of the presence of lexical/phraseological metaphoricity cues (Goatly 1997), could suggest metaphoricity. Particularly difficult cases for metaphoricity detection arise in the case of "implicit" metaphor (Steen 1999; see also Kittay 1989) where the semantic or pragmatic tension is between sentences rather than within a sentence.

Acknowledgment

The research was supported in part by grant number IRI-9101354 from the National Science Foundation (USA). It is being supported at the time of writing by grant GR/M64208 from the Engineering and Physical Sciences Research Council (UK).

References

Barnden, J.A. (1995). High-level reasoning, computational challenges for connectionism, and the Conposit solution. *Applied Intelligence, 5* (2), pp.103–135.

Barnden, J.A. (1997). Consciousness and common-sense metaphors of mind. In S. O'Nuallain, P. McKevitt & E. Mac Aogain (Eds), *Two Sciences of Mind: Readings in Cognitive Science And Consciousness*, pp.311–340. Amsterdam/Philadelphia: John Benjamins. Amsterdam/Philadelphia: John Benjamins.

Barnden, J.A. (1998a). An AI system for metaphorical reasoning about mental states in discourse. In Koenig, J-P. (Ed.), *Discourse and Cognition: Bridging the Gap* (pp.167–188). Stanford, CA: CSLI.

Barnden, J.A. (1998b). Combining uncertain belief reasoning and uncertain metaphor-based reasoning. In *Procs. Twentieth Annual Meeting of the Cognitive Science Society,* pp.114–119. Mahwah, N.J.: Lawrence Erlbaum Associates.

Barnden, J.A. (1998c). Uncertain reasoning about agents' beliefs and reasoning. Technical Report CSRP-98-11, School of Computer Science, The University of Birmingham, U.K. Also in *Artificial Intelligence and Law,* in press.

Barnden, J.A. (1998d). Conceptual blending and an implemented system for metaphor-based reasoning. Poster presentation at *Fourth Conference on Conceptual Structure, Discourse, and Language*, Emory University, Atlanta, Georgia, October 1998.

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

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

Barnden, J.A. (in press, a). Uncertainty and conflict handling in the ATT-Meta context-based system for metaphorical reasoning. To appear in the proceedings of *Third International Conference on Modeling and Using Context* (Dundee, U.K., 2001), Lecture Notes in Artificial Intelligence, Springer-Verlag.

Barnden, J.A. (in press, b). The utility of reversed transfers in metaphor. To appear in *Procs. Twenty-Third Annual Meeting of the Cognitive Science Society* (Edinburgh, August 2001). Lawrence Erlbaum.

Barnden, J.A. (n.d.). ATT-Meta Project Databank.
`http://www.cs.bham.ac.uk/~jab/ATT-Meta/Databank`.

Barnden, J.A., Helmreich, S., Iverson, E. & Stein, G.C. (1994a). 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. (1994b). Combining simulative and metaphor-based reasoning about beliefs. In *Procs. 16th Annual Conference of the Cognitive Science Society* (Atlanta, Georgia, August 1994), pp.21–26. Hillsdale, N.J.: Lawrence Erlbaum.

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), pp.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* (CONTEXT'99, Trento, Sep 9–11, 1999), *Lecture Notes in Artificial Intelligence, 1688,* pp.28–41, Springer.

Barnden, J.A. & Lee, M.G. (2001). Application of the ATT-Meta metaphor-understanding system to an example of the metaphorical view of MIND PARTS AS PERSONS. Technical Report CSRP–01–09, School of Computer Science, The University of Birmingham, U.K.

Brisard, F., Frisson, S. & Sandra, S. (2001). Processing unfamiliar metaphors in a self-paced reading task. *Metaphor and Symbol, 16*(1&2), pp.87–108.

Cameron, L. (1999a). Operationalising 'metaphor' for applied linguistic research. In L. Cameron & G. Low (Eds), *Researching and Applying Metaphor,* pp.1–28. Cambridge, U.K.: Cambridge University Press.

Cameron, L. (1999b). Identifying and describing metaphor in speken discourse data. In L. Cameron & G. Low (Eds), *Researching and Applying Metaphor,* pp.105–132. Cambridge, U.K.: Cambridge University Press.

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, N.J.: Lawrence Erlbaum.

Carruthers, P. & Smith, P.K. (Eds). (1996). *Theories of theories of mind.* Cambridge, UK: Cambridge University Press.

Clausner, T.C. & Croft, W. (1997). Productivity and schematicity in metaphors. *Cognitive Science, 21*(3), pp.247–282.

Cooper, D.E. (1986). *Metaphor.* Oxford, U.K.: Blackwell.

Creary, L. G. (1979). Propositional attitudes: Fregean representation and simulative reasoning. *Procs. 6th. Int. Joint Conf. on Artificial Intelligence* (Tokyo), pp.176–181. Los Altos, CA: Morgan Kaufmann.

Croft, W. (1998). Linguistic evidence and mental representations. *Cognitive Linguistics, 9*(2), pp.151–173.

Davidson, D. (1979). What metaphors mean. In S. Sacks (Ed.), *On Metaphor.* University of Chicago Press.

Davies, M & Stone, T. (Eds) (1995). *Mental simulation: evaluations and applications.* Oxford, U.K.: Blackwell.

Deignan, A. (1999). Linguistic metaphors and collocation in nonliterary corpus data. *Metaphor and Symbol, 14*(1), pp.19–36.

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, Connecticut: Ablex.

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

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), pp.235–241.

Gibbs, R.W., Jr. (1989). Understanding and literal meaning. *Cognitive Science, 13*, pp.243–251.

Gibbs, R.W., Jr. (1990). Comprehending figurative referential descriptions. *J. Experimental Psychology: Learning, Memory and Cognition, 16* (1),pp.56–66.

Gibbs, R.W., Jr. (1998). The fight over metaphor in thought and language. In A.N. Katz, C. Cacciari, R.W. Gibbs, Jr., and M. Turner, *Figurative Language and Thought*, pp.88–118. New York, Oxford: Oxford University Press.

Gineste, M-D. & Scart-Lhomme, V. (1998). Comment comprenons-nous les métaphores? *L'Année Psychologique,* 1998.

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

Giora, R. & Fein, O. (1998). Irony: graded slaience and indirect negation. *Metaphor and Symbol, 13*(2), pp.83–101.

Glucksberg, S. & Keysar, B. (1990). Understanding metaphorical comparisons: beyond similarity. *Psychological Review, 97* (1), pp.3–18.

Goatly, A. (1997). *The language of metaphors.* London and New York: Routledge.

Grady, J.E. (1997). THEORIES ARE BUILDINGS revisited. *Cognitive Linguistics, 8*(4), pp.267–290.

Graesser, A.C. & Bower, G.H. (Eds). (1990). *Inferences and text comprehension.* San Diego, CA: Academic Press.

Hobbs, J.R. (1985). On the coherence and structure of discourse. Report No. CSLI–85–37, Center for the Study of Language and Information, Stanford University, Calif., USA.

Hobbs, J.R. (1990). *Literature and cognition.* CSLI Lecture Notes, No. 21, Center for the Study of Language and Information, Stanford University.

Holyoak, K.J., Novick, L.R. & Melz, E.R. (1994). Component processes in analogical transfer: mapping, pattern completion, and adaptation. In K.J. Holyoak & J.A. Barnden (Eds), *Advances in Connectionist and Neural Computation Theory, Vol. 2.* Norwood, N.J.: Ablex Publishing Corp.

Holyoak, K.J. & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13* (3), 295–355.

Honeck, R.P., Welge, J. & Temple, J.G. (1998). The symmetry control in tests of the standard pragmatic model: The case of proverb comprehension. *Metaphor and Symbol, 13*(4), pp.257–273.

Hunston, S. (1995). A corpus study of some English verbs of attribution. *Functions of Language, 2E(2): pp.133–158.*

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

Keane, M. (1988). Incremental analogising: Theory and model. HCRL Tech. Rept. No. 38, Human Cognition Research Laboratory, Milton Keynes, U.K. Also in K. Gilhooly, M. Keane, R. Logie & G. Erdos (Eds), *Lines of Thought: Reflections on the Psychology of Thinking,* Chichester, U.K.: Wiley.

Keenan, J.M., Golding, J.M., Potts, G.R., Jennings, T.M. & Aman, C.J. (1990). Methodological issues in evaluating the occurrence of inferences. In A.C. Graesser & G.H. Bower (Eds), *Inferences and Text Comprehension,* pp.295–312. San Diego, CA: Academic Press.

Kennedy, J.M. & Chiappe, D.L. (1999). What makes a metaphor stronger than a simile? *Metaphor and Symbol, 14*(1), pp.63–69.

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

Lakoff, G. (1990). The Invariance Hypothesis: is abstract reason based on image-schemas? *Cognitive Linguistics, 1* (1), pp.39–74.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and Thought,* 2nd edition. Cambridge University Press.

Lakoff, G. (1994). What is metaphor? In J.A. Barnden & K.J. Holyoak (Eds.), *Advances in Connectionist and Neural Computation Theory, Vol. 3: Analogy, Metaphor and Reminding,* pp.203–258. Norwood, N.J.: Ablex Publishing Corp.

Lakoff, G. (1996). Sorry, I'm Not Myself Today: The metaphor system for conceptualizing the self. In G. Fauconnier & E. Sweetser (Eds), *Spaces, Worlds, and Grammar*, pp.91–123. Chicago, IL: University of Chicago 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.* U. Chicago Press.

Lee, M.G. (1998). Belief, rationality and inference: A general theory of computational pragmatics. Ph.D. Thesis, Department of Computer Science, The University of Sheffield, U.K.

Lee, M.G. & Barnden, J.A. (2000). Metaphor, pretence and counterfactuals. Oral presentation at *Fifth Conference on Conceptual Structure, Discourse, and Language*, May 2000, University of California, Santa Barbara, California, U.S.A.

Lee, M.G. & Barnden, J.A. (2001a). Reasoning about mixed metaphors with an implemented AI system. *Metaphor and Symbol, 16* (1&2), pp.29–42. Also as Technical Report CSRP–00–11, School of Computer Science, The University of Birmingham, U.K.

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, School of Computer Science, The University of Birmingham, U.K.

Levin, S.R. (1993). Language, concepts, and worlds: three domains of metaphor. In A. Ortony (Ed.), *Metaphor and Thought,* 2nd edition, pp.112–123. New York and Cambridge, U.K.: Cambridge 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), pp.416–432.

Mann, W.C. & Thompson, S.A. (1987). Rhetorical theory of text organization. *Text, 8* (3), pp.167–182.

Markman, A.B. (1997). Constraints on analogical inference. *Cognitive Science, 21*(4), pp.373–418.

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

Martin, J. (1994). Metabank: A knowledge-base of metaphoric language conventions. *Computational Intelligence, 10* (2), pp.134–149.

Moon, R. (1998). *Fixed idioms and expressions in English.* Clarendon Press: Oxofrd, U.K.

Murphy, G.L. (1996). On metaphoric representation. *Cognition, 60*, pp.173-204.

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

Onishi, K.H. & Murphy, G.L. (1993). Metaphoric reference: when metaphors are not understood as easily as literal expressions. *Memory and Cognition, 21* (6), pp.763–772.

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

Radman, Z. (1997). Difficulties with diagnosing the death of a metaphor. *Metaphor and Symbol, 12*(2), pp.149–157.

Récanati, F. (1993). The alleged priority of literal interpretation. *Cognitive Science, 19*(2), pp.207–232.

Reddy, M.J. (1979/1993). The conduit metaphor—a case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and Thought*, Cambridge, UK: Cambridge University Press. Reprinted in second edition, 1993, pp.164–201.

Shastri, L. & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: a connectionist representation of rules, variables, and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences, 16* (3), pp.417–494.

Steen, G. (1999). Metaphor and discourse: Towards a linguistic checklist for metaphor analysis. In L. Cameron & G. Low (Eds), *Researching and Applying Metaphor*, pp.81–104. Cambridge, U.K.: Cambridge University Press.

Sun, R. (1995). A microfeature based approach towards metaphor interpretation. In *Procs. Fourteenth Int. Joint Conference on Artificial Intelligence* (IJCAI-95), Montreal, Canada, pp.424-429

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

Veale, T. & Keane, M.T. (1997). The competence of sub-optimal structure mapping on 'hard' analogies. In *Procs. Int. Joint Conf. On Artificial Intelligence* (Nagoya, Japan), August 1997.

Waggoner, J.E. (1990). Interaction theories of metaphor: psychological perspectives. *Metaphor and Symbolic Activity, 5* (2), pp.91–108.

Wilks, Y. (1978). Making preferences more active. *Artificial Intelligence, 10*, pp.75–97.