Belief Modelling, Intentionality and Perlocution in Metaphor Comprehension

Tony Veale, M. T. Keane

Department of Computer Science, Trinity College, Dublin, Ireland.

1. Introduction

The view that metaphor plays a fundamental structural role in organizing our conceptual systems, rather than serving a deviant rhetorical effect, is now generally accepted. Indeed, this idea drives many recent computational models (Weiner , 1984; Martin, 1990; Way, 1991;Veale & Keane, 1992a, 1992b). However, these analyses are based on the de-contextualised "contents" of the metaphor and ignore the communicative context in which it is uttered. Additionally, Davidson (1979) has warned against confusing the content of a metaphor with the intended effect of a metaphor. In his view, metaphor is not a *carrier* of meaning, but a *recipe* from which meaning is constructed. The real meaning of a metaphor lies in the changes it makes to belief structures of the hearer/reader.

There exists a pragmatic imperative then, that in order to capture the richness of figurative utterances, metaphors need to be viewed as full speech acts. From this perspective a metaphor is uttered by a speaker with a specific communicative intent, in the context of a *speaker-specific* world model, and subsequently interpreted by a listener relative to a local world model which is similarly *listener-specific*. To arrive at a full interpretation of the utterance, a system must therefore characterise that component of meaning which is common to the world models of speaker and listener, before the *information content*, as opposed to the semantic content, of the metaphor can be represented. So, a proper analysis requires the system to model the belief structures of the speaker and listener relative to each other in the context of the tenor and vehicle concepts. Recognising the communicative intent of the speaker is also important if the speaker's beliefs are to be correctly modelled by the listener (e.g., is the speaker conveying a pejorative or ameliorative account of the tenor?).

Various researchers, notably Wilks, Barden & Wang (1991), have subsumed the metaphoric process within a larger framework of belief ascription. These authors treat metaphor as a structural/propositional transference between domains and show how some properties of metaphor naturally arise out of a model of belief-space *amalgamation*. The work of Wilks *et al.* informs much of what follows, but I am here concerned with various other properties of metaphor which do not seem to be captured by their framework: properties such as *salience imbalance* (see Ortony, 1979; Weiner, 1984), *domain incongruence* (see Tourangeau & Sternberg, 1981) and *domain interaction* (see Richards 1936, Black 1962).

1.1. Structure of this Paper

So, our emphasis here is more on the role of belief ascription and speech act analysis in metaphor interpretation, rather than on belief ascription in itself. To begin, section 2 considers those aspects of speech-act theory which are applicable to the metaphor comprehension process. Sections 3 then presents an extension to the Sapper model of memory to support the mechanics of belief ascription and thereby accommodate these pragmatic aspects of communication. Section 4 demonstrates how the activation dynamics of the Sapper connectionist network can be employed to derive a quantitative measure of both the emotive and communicative force of a metaphor, while sections 5 and 6 illustrate how various heuristic rules of belief revision capitalise upon this measure. This paper then finishes with an overall summary and some conclusions in section 7.

2. Metaphor as a Communicative Act

Viewed as a conceptual process, metaphor is a directional juxtaposition of two conceptual schemata (tenor and vehicle), which affects the structure of not only the tenor but also the vehicle. Consequently, metaphor comprehension is often viewed solely as an issue of knowledge representation and combination (see Weiner 1984). However, metaphor is also a communicative act which observes the pragmatic *contract* between speaker and listener in the transfer of information between both. This contract (comprising the rules of conversational implicature) defines the terms of mediation between the mental models of both parties. It follows that a metaphor cannot be analysed fully in the context of a single mental model (i.e., the system's knowledge-base - effectively semantic memory), but that assumptions regarding the knowledge possessed by both listener *and* speaker are also necessary elements of the analysis.

The term *metaphoric ground* is commonly used in reference to the set of common associations / attributes shared by tenor and vehicle, congruently or otherwise. However, to incorporate notions of conversational implicature into our treatment, it becomes necessary to elucidate these additional types of *ground*:

Speaker Ground: The set of associations shared by tenor and vehicle in the mental model of the speaker.

Listener Ground: The set of associations shared by tenor and vehicle in the mental model of the listener.

Conversational Ground: The common set of associations shared by tenor and vehicle in the mental models of both the speaker and listener (i.e., the set intersection of the Speaker Ground and the Listener Ground).

Imparted Ground: The set of associations shared by tenor and vehicle in the mental model of the speaker, but not in the mental model of the listener (i.e., the set difference of the Speaker Ground and the Listener Ground).

For a metaphor to be intelligible in a particular context, there must exist a Conversational Ground between speaker and listener. But if this were all that were required for comprehension, metaphor would indeed be a sterile form of comparison. However, metaphor is also a compact form of expression in which both common knowledge (the conversational ground) and new knowledge (the imparted ground) is transferred from speaker to listener. Essentially, the conversational ground provides the basis of understanding (the broad terms of agreement) against which the speaker can relate personal beliefs / feelings regarding the tenor. The choice of metaphoric vehicle thus reflects not only the speaker's knowledge of the tenor, but also the speaker's beliefs about the listener's knowledge of the tenor and vehicle.

Consider the metaphor "Bank managers are vampires", an emotively pejorative statement which reflects the speakers contempt for bankers (and banks themselves). The Sapper network description of this metaphor is illustrated in Figure 1.

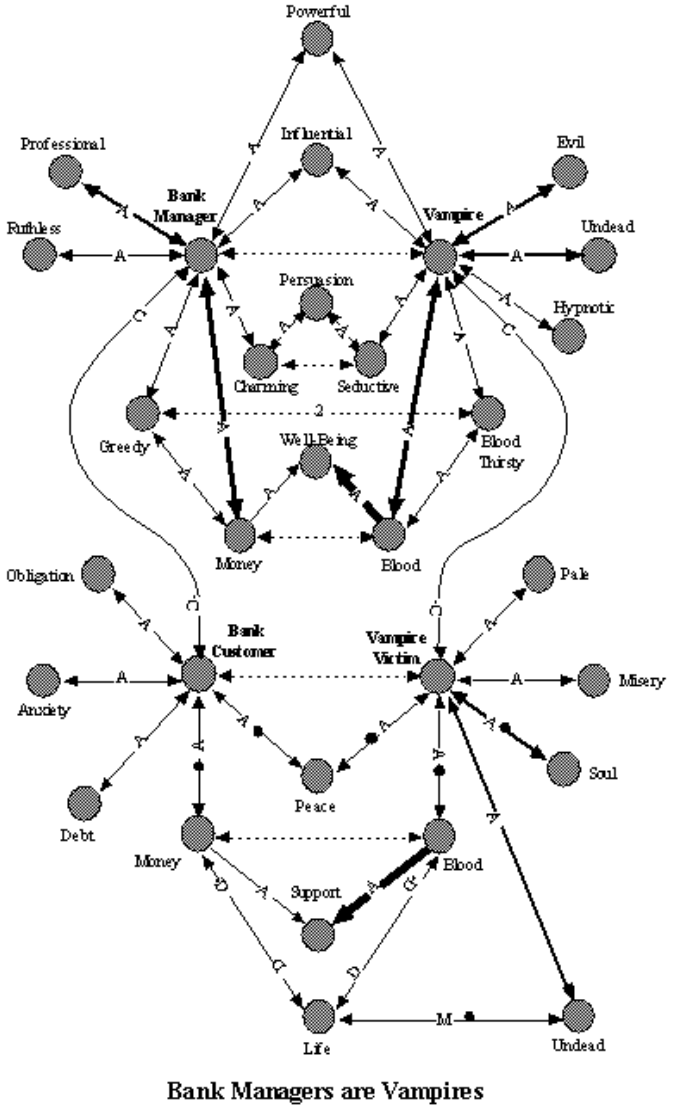

Figure 1: The Sapper description of the metaphor "Bank Managers are Vampires".

Interpretation of this concept juxtaposition relies on the following grounds:

Speaker Ground *(unhappy with his bank manager)*: Bankers are powerful, influential and blood-thirsty people, which act in an evil, cruel fashion, seducing customers with loans and then sucking the life (blood) from them. Bankers, like vampires, often bring misery and death (financial) to their customers.

Listener Ground *(reasonably happy with his bank manager)*: Bankers, like vampires, are powerful, influential and charming people, which often grant loans but expect more money in return. This expectation is often considered greedy and ruthless.

Conversational Ground: Bankers, like vampires, are powerful and influential. They are blood-thirsty inasmuch as they can be greedy, suck blood inasmuch as they take money (i.e., financial life-blood), seductive inasmuch as they can be charming, and cruel inasmuch as they can be ruthless.

Imparted Ground: Bankers are cruel, evil, seductive blood-suckers and bringers of misery and ruin.

Notice that it is the Imparted Ground which lends the metaphor colour and life. The Conversational Ground, however, provides the bridge that allows the listener to link the concepts of Banker and Vampire; once this link is established, interaction of both conceptual schemata is possible, in the fashion of Black (1962), enabling the attributes of Vampire to mix with those of Banker, and vice versa (e.g., vampires may be seen as *respectable* and *business-like*).

3. Belief Ascription in Sapper

A belief maintenance architecture is superimposed onto the Sapper network model to accommodate this speech-act perspective on metaphor. Following the terminology and graphic conventions of Wilks, Barnden & Wang 1991, as employed within their *ViewGen* model, the belief sets of sentient agents (e.g., people, computers) are modelled as *viewpoint environments* which contain different *topic environments*, each relating to a different concept as viewed by a particular agent . This situation is illustrated in Figure 2.

Again following the graphical conventions established by Wilks, Barnden & Wang (1991), Figure 2 illustrates a belief-fragment from the speaker's mental model regarding the earlier metaphor of Figure 1. This speaker, who clearly believes bank managers to be seductively evil, has restructured his mental representation of said bureaucrats around the concept of vampirism, via a process known as *viewpoint amalgamation*. This amalgamation process, which is functionally isomorphic to Sapper's establishment of cross-domain analogues, allows predications believed of one concept to be applied to another. Now, it is the purpose of the metaphor "Bank Managers are Vampires" to persuasively convey this state of affairs to the hearer, whose own a priori representation of bank managers might look like Figure 3.

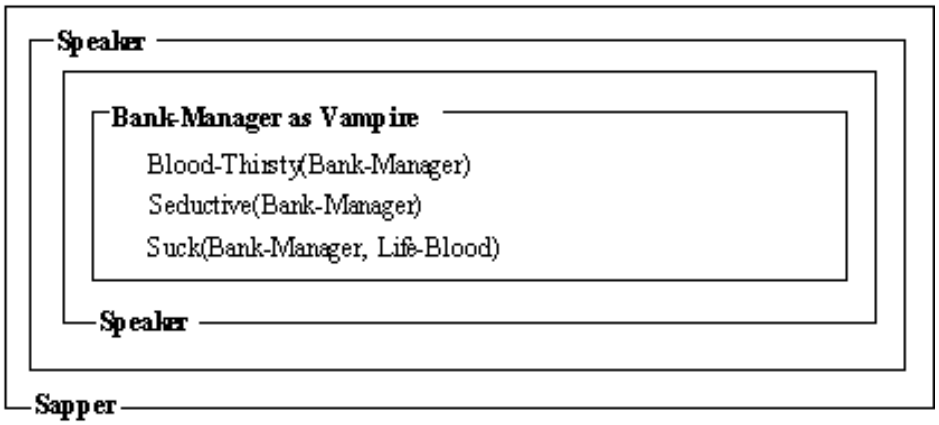

Figure 2: The belief-model representation of a speaker's beliefs regarding bank managers. The inner topic represents an amalgamated belief environment. Notation: Boxes with lower labels represent viewpoint or believer environments, while boxes with upper labels represent topic environments.

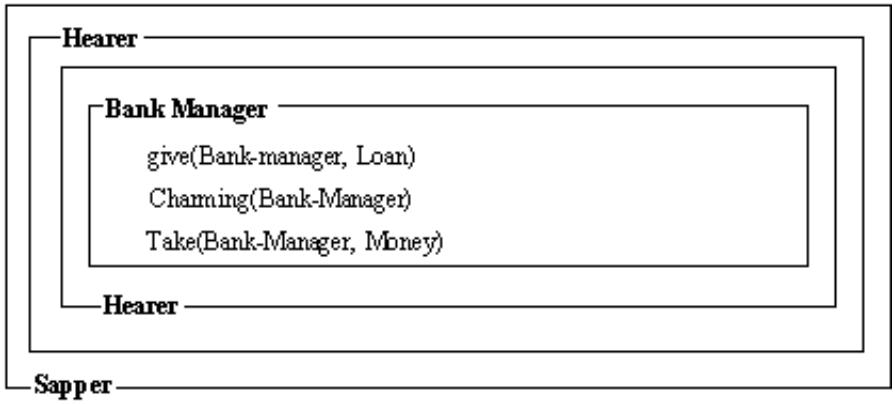

Figure 3: A belief-model of a neutral hearer's beliefs regarding bank managers.

The belief model of Figure 3 presents a much more reasonable picture of a prototypical Bank Manager - that of a charming professional that gives loans and takes back money. At first blush, then, there does not seem to exist a conversational ground between the descriptions offered by Figures 2 and 3. However, if the belief system is willing (and capable) to pursue the comparison through the realms of domain incongruence, cross-domain linkages can be established between the concepts Charming & Seductive, Money & Blood, and Take & Suck; Figure 1 graphically illustrates the first two of these mappings. Additionally, if the hearer also believes that bank managers are greedy to some degree, this will augment the conversational ground with the additional mapping Greedy: Blood-Thirsty; this higher-order mapping is a product of the squaring rule, and is depicted in Figure 1 with the label "2". In effect then, Sapper brings to this belief space model an additional layer of metaphoric robustness, introducing a capacity for recognising and exploiting domain incongruences in the amalgamation of belief spaces and the determination of conversational ground.

The extended Sapper memory model thus comprises two interlinked architectures: a belief-space model composed of viewpoint and topic environments, and a localist connectionist network comprising weighted links, each of which represents a distinct proposition. Each link is annotated with one or more topic environment identifiers, as the same proposition may be held by different agents, and a corresponding credibility weighting for each topic, which represents the credence each belief agent places upon the proposition. A viewpoint is considered *primed* whenever that viewpoint is currently adopted by the system, allowing Sapper to consider a proposition from within the belief spaces of different agents; the priming of a viewpoint environment automatically causes its constituent topic environments, and the propositions contained therein, to be primed also. To support this *viewpoint-switching* capability, activation is therefore constrained by the connectionist component to occur across only those proposition links contained within a currently primed topic environment.

4. Ameliorative and Pejorative Shift in Comprehension

The existence of multiple viewpoints within semantic memory allows for considerable *meaning* shift, or *conceptual parallax*, to occur in the comprehension of metaphor, as an utterance may be interpreted differently within conflicting agent belief-spaces. This issue is of particular relevance in quantifying the overall ameliorative/pejorative content of an utterance: consider for instance the metaphor of Figure 4, "Surgeons are Butchers", in which an ameliorative shift occurs when considering the metaphor from the viewpoint of the vehicle, while a complementary pejorative shift occurs when a tenor viewpoint is adopted.

The Sapper mappings generated for this metaphor are as follows:

[.86] If Butcher is like Surgeon

[.58] **Then** White-Apron is like White-Smock

[.25] *and* Abattoir is like Operating-Theatre

[.75] *and* Meat is like Human-Flesh

[.94] *and* Cleaver is like Scalpel

[.98] *and* Carcass is like Corpse

[.95] *and* Slaughter is like Surgery

In drawing two conceptual domains closer together, the metaphor attempts to reconcile the myriad associations of the tenor and vehicle into a cohesive whole. However, while a systematic 1-to-1 mapping between these domains is achieved, the metaphor nevertheless fails to marry the connotative meaning of the two domains completely, but this is as it should be. What makes a good metaphor interesting is the volatile nature of the juxtaposition - while there are good structural reasons why both domains should be united, as demonstrated by the mappings above, there also exist conflicting associations that imbue the metaphor with a certain unsettling quality. Thus, bubbling below this neat structural reconciliation is a tension that threatens to

tear the domains apart, rather than bring them together, and it is this tension that lends metaphor its characteristic flavour of disturbing unorthodoxy.

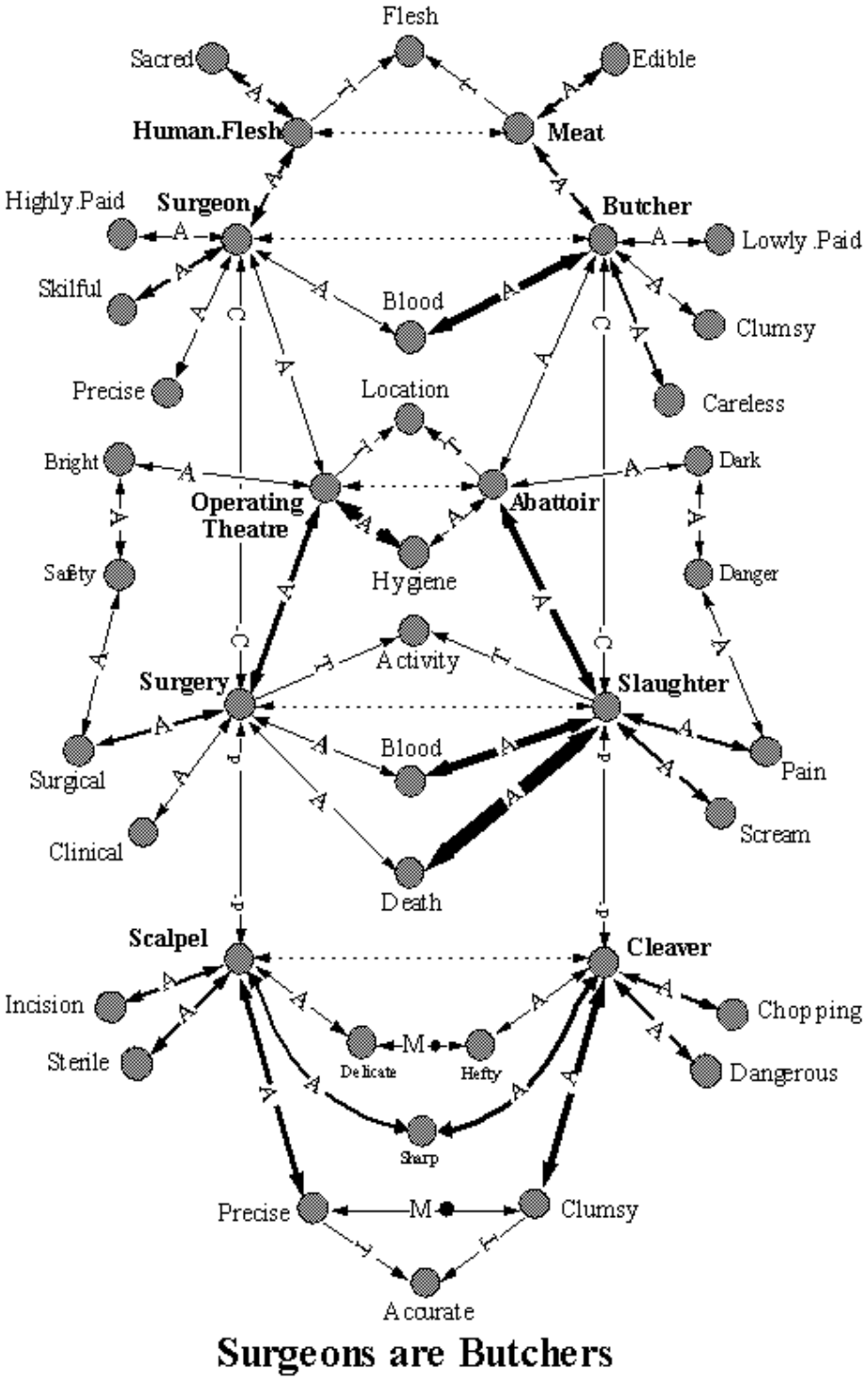

Figure 4: Sapper description of the metaphor "Surgeons are Butchers".

For instance, in the metaphor above, Human-Flesh is seen to resemble Meat, which consequently diminishes its connotation of sacred inviolability. Furthermore, this analogue opens the door to anthropophagy, allowing human flesh to be viewed as edible, and perhaps more disturbingly, as even tasty. Likewise, an operating theatre is seen to resemble an abattoir, which serves to accentuate the associations of pain and death which are normally so understated in the stereotypical conceptualization of surgery. The creation of this mapping also diminishes the connotations of brightness, and indirectly, safety, associated with operating theatres, highlighting instead the connotations of darkness and danger commonly associated (rightly or wrongly) with abattoirs. This of course assumes a reading from the tenor perspective - from the vehicle perspective, the concept of abattoir would seem to be lightened somewhat by the association with clinical and bright operating theatres, places where pain, darkness, screaming and death assume a less salient cast. Likewise, from the butcher's perspective, the apposition of meat with human-flesh imbues the former with a sacred overtone that is not conventionally attributed to the victualler's profession.

A means of gauging this ameliorative/pejorative shift between the tenor and vehicle viewpoints is therefore necessary if the system is to accurately determine the *cognitive, or emotive force* of a metaphor, that is, a context-independent measure of the persuasive power of the metaphor regardless of the speaker's intent. Such a measure is of course a principal determiner in updating the belief-space of the hearer in response to the speaker's metaphor, and indicates much of what should be considered the *imparted ground* of the metaphor. A connectionist means of obtaining such a measure is simply implemented within Sapper, as illustrated in Figure 5:

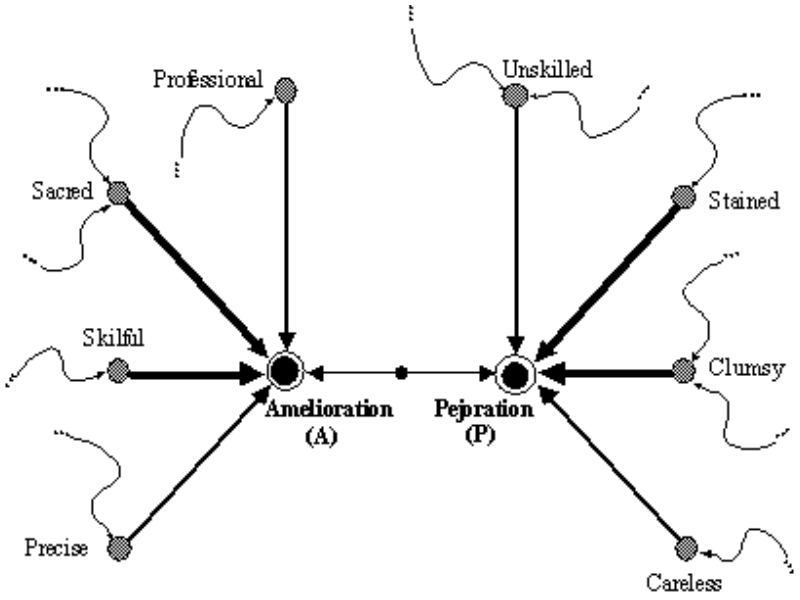

Figure 5: Two collector nodes (depicted black on white), which accept all incoming activation, but do not propagate it onwards, are used to measure the relative ameliorative/pejorative content of a concept within a given network.

As shown in Figure 5, Sapper designates two concept nodes, Amelioration (denoted A) and Pejoration (denoted P), to serve as special activation collectors, which accumulate but do not redistribute activation from adjoining nodes. Descriptive

concepts throughout the network which possess any measure of ameliorative/ pejorative connotation are connected to these nodes, whereby the association strength of each linkage reflects the extent of the connotation. In this respect these nodes resemble the *pragmatic node* that tempers the ACME model with contextual activation.

So, by priming the viewpoint environment of an observer O, and subsequently initiating the spread of activation, or *zorch* Z, from a concept node T, the activation levels of the collectors A and P will provide a measure of the relative ameliorative/pejorative view toward T by O:

• $Ameliorative(O, T) = B_O Z_T (A)$

• $Pejorative(O, T) = B_O Z_T (P)$

The zorch arriving at collector node A from node T, denoted $Z_T (A)$, in the viewpoint environment of believer O, denoted, B_O, represents a measure of the ameliorative view felt by O for T, denoted Ameliorative(O, T). The overall view (or *esteem*) held by O for T is denoted View(O, T); a positive value for this function implies O holds an ameliorative opinion of T, while a negative value suggests a pejorative opinion.

• $View(O, T) = B_O Z_T (A) - B_O Z_T (P)$

It follows therefore that T's self-esteem is defined by:

• $Self\text{-}Esteem(T) = View(T, T) = B_T Z_T (A) - B_T Z_T (P)$

The system is thus in a situation to gauge the effect upon a hearer H's self-esteem should H adopt the proposition set X (i.e., a set of one or more new propositions):

• $\Delta Self\text{-}Esteem(H) = View(H+X, H) - View(H, H)$

$$= B_{H+X} Z_H (A) - B_{H+X} Z_H (P) - B_H Z_H (A) + B_H Z_H (P)$$

The proposition set of interest to Sapper, of course, is the set of bridges awakened in comprehending a metaphor, as Sapper is required to attach a credibility weighting to each of these new propositions in the belief space of the Hearer H. Calculation of the ensuing effects upon the self-esteem of the Hearer is therefore essential if the system is to ascertain the likelihood of the hearer adopting X, and determine the certainty / credibility rating that H will attach to X. However, it is also necessary to determine the perceived change in self-esteem experienced by the speaker S as observed by the hearer H:

• $B_H \Delta Self\text{-}Esteem(S) = B_H [View(S+X, S) - View(S, S)]$

$$= B_H [B_{S+X} Z_S (A) - B_{S+X} Z_S (P) - B_S Z_S (A) + B_S Z_S (P)]$$

Thus, for a metaphoric utterance M, in which speaker S describes a tenor concept T via the vehicle concept V to a hearer H, the emotive force of M, *Ef(M)*, is defined:

• $Ef(M) = |View(H, V)| = |B_H Z_V(A) - B_H Z_V(P)|$

i.e., the absolute ameliorative/pejorative content of the vehicle chosen to describe the tenor. However, as suggested above, the system must take into consideration the effect the utterance M has upon the hearer H, and the effect M has upon S as perceived by H, if the persuasive power of the metaphor is to be determined. The *Pragmatic force*, Pf(M) of a metaphor reflects this perspective shift that exists between speaker and hearer:

• $Pf(M) = \Delta Self\text{-}Esteem(H) - B_H \Delta Self\ Esteem(S)$

A context-dependent, or pragmatically motivated, measure of the persuasive force of the metaphor can now be formulated. This context-dependent measure is termed the *communicative force* of the metaphor, or Cf(M). If the speaker S, in uttering the *metaphor M, communicates the set of propositions X to the hearer H, then Cf(M) is:*

• $Cf(M) = f\,.\,Ef(M) + (1 - f).\,Pf(M)$

$= f\,.\,Ef(M) + (1 - f).\,[\Delta Self\text{-}Esteem(H) - B_H \Delta Self\ Esteem(S)]$

$= f\,.\,|View(H, V)| + (1 - f).View(H+X, H) - (1 - f).View(H, H)$

$- (1 - f).B_H\,View(S+X, S) + (1 - f).B_H\,View(S, S)$

$= |f\,.B_H Z_V(A) - f\,.B_H Z_V(P)|$

$+ (1 - f).B_{H+X} Z_H(A) - (1 - f).B_{H+X} Z_H(P)$

$- (1 - f).B_H Z_H(A) + (1 - f).B_H Z_H(P)$

$- (1 - f).B_H B_{S+X} Z_S(A) + (1 - f).B_H B_{S+X} Z_S(P)$

$+ (1 - f).B_H B_S Z_S(A) - (1 - f).B_H B_S Z_S(P)$

The scaling factor f represents the readiness, or *susceptibility*, of the system to accept the propositions conveyed through an utterance without recourse to the relative pragmatic position of the hearer and speaker. As defined above, the communicative force of a metaphor is thus not only dependent upon the emotive force of the utterance, but also upon the likely effect that acceptance of the metaphor is to have upon the hearer H, and upon the *openness* of the speaker S as perceived by H. If M does not relate to H in any descriptive fashion, then ΔSelf-Esteem(H) is zero and not a factor in the final interpretation. However, if M describes H positively, the communicative force is strengthened, causing H to attribute more certainty to the propositions conveyed therein, while if M describes H negatively, the communicative

force is diminished and less certainty is attributed to the propositions it conveys. Similarly, if M is seen by H to describe S positively, that is, in a self-serving manner, then H is less likely to accept fully the propositions conveyed by S, while if M is seen to describe S negatively, that is, in a self-deprecating manner, H is more likely to attach greater credence to the propositional content of M. Recalling the example metaphor of Figure 4, "Surgeons are butchers", this utterance has more communicative power when either uttered *by* a surgeon, or *to* a butcher, and less communicative power when uttered *to* a surgeon or *by* a butcher.

What has been achieved by this algebraic manipulation? Building upon a simple means of ascertaining the relative ameliorative/pejorative view of a concept from the perspective of a particular belief agent, Sapper is capable of quantifying the *persuasive power*, or communicative force, of a metaphor as it relates to the beliefs of the hearer.

5. Determination of Speaker Intentionality

As described and formulated in the previous section, a metaphoric utterance may effect changes in the belief space of the hearer, and some measure of the likelihood that such changes are wrought is the communicative force of the utterance. Of course, the utterance is communicated by the speaker with precisely such a persuasive goal, and following Davidson (1979), metaphors must by analysed in terms of the pragmatic goals of the speaker. When a speaker utters "My wife's cooking is a disaster", the speaker's intention is to convey a dislike of his wife's cooking. However, a propositional expression of this dislike is not contained within the metaphor itself, but arises from a pragmatic analysis of the utterance as a speech act.

Given a metaphor M uttered by S, which describes T as V, and in doing so conveys the proposition set X to H, then the following belief ascription heuristics apply:

Like/Dislike:

• When M is perceived to be pejorative toward T, i.e., View(H, V) < 0,

Infer: $\Delta B_H\, B_S\, Z_T(P) \propto Cf(M)$

i.e., modify the link strengths to ensure that more *zorch* arrives at P from T in B_H B_S

• When M is perceived to be ameliorative toward T, i.e., View(H, V) > 0,

Infer: $\Delta B_H\, B_S\, Z_T(A) \propto Cf(M)$

Trust/Distrust:

• When M is perceived to be pejorative toward T, i.e., View(H, V) < 0,

Infer: $\Delta B_H\, B_S\, Z_T(Untrustworthy) \propto Cf(M)\, .\, max[0, B_H\, B_S\, Z_V(Power)]$

i.e., modify the link strengths to ensure that more zorch arrives at Untrustworthy from T in B_H B_S - the extent of the modification is related to H's certainty that S holds V to be in a position of power.

• When M is perceived to be ameliorative toward T, i.e., View(H, V) > 0,

Infer: $\Delta B_H\, B_S\, Z_T\,(Trustworthy) \propto Cf(M)\,.\, max[0,\, B_H\, B_S\, Z_V(Power)]$

Respect/Fear

• When M is perceived to be ameliorative toward T, i.e., View(H, V) > 0,

Infer: $\Delta B_H\, B_S\, Z_T(Respect) \propto Cf(M)\,.\, max[0,\, B_H\, B_S\, Z_V(Power) - B_H\, B_S\, Z_S(Power)]$

i.e., modify the link strengths to ensure that more zorch arrives at Respect from T in B_H B_S - the extent of the modification is related to H's certainty that S holds V to be in a position more powerful than himself/herself.

• When M is perceived to be pejorative toward T, i.e., View(H, V) < 0,

Infer: $\Delta B_H\, B_S\, Z_T(Fear) \propto Cf(M)\,.\, max[0,\, B_H\, B_S\, Z_V(Power) - B_H\, B_S\, Z_S\,(Power)]$

These heuristics provide a principled basis for modifying the hearer's beliefs concerning the speaker's beliefs as communicated in a metaphor. In each case, the proposed modification alters the activation dynamics of the belief viewpoint environment $B_H B_S$, in a manner which is consistent with the perceived communicative force of the utterance, as formulated in the previous section.

6. Metaphors of Mind in Belief Ascription

Representing fundamental metaphors of mind is essential not only to the effective modelling of belief ascription, but also to the full treatment of metaphor as a speech-act. For in dealing with metaphors which relate to a sentient tenor, that is, a tenor which is a belief agent in its own right, the system must do more than determine those beliefs the hearer will ascribe to the speaker (the perlocutive effect); it is also necessary to determine those beliefs that the hearer perceives the speaker to ascribe to the tenor. To model this ascription from speaker to tenor, Sapper employs a fundamental *metaphor of mind*, that of concept association and disassociation (such basic metaphors of cognition are discussed in Lakoff & Johnson 1980, and given a computational treatment in Veale & Keane 1992a,b). The essential principle is this: when a speaker S utters a metaphor M to insult a sentient tenor T, such that B_HView(S, V) < 0, then H perceives S to *distance* himself conceptually from T; however, when a speaker S utters a metaphor M to praise a sentient tenor T, such that B_HView(S, V) > 0, then H perceives S to move himself conceptually *closer* to T. Consider then Figures 6 and 7:

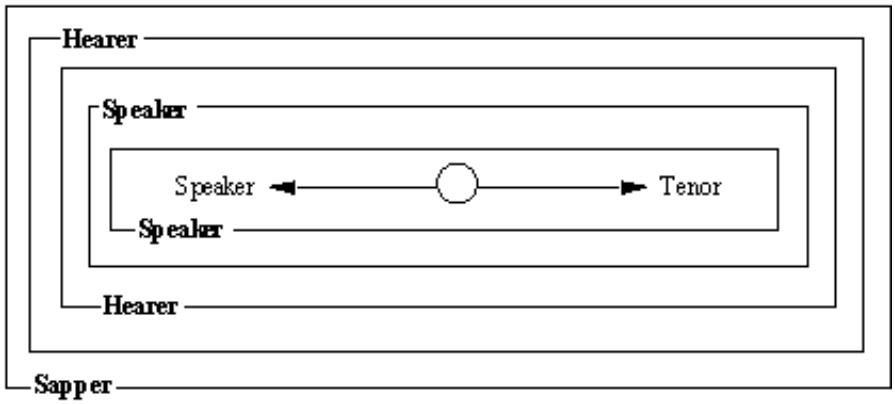

Figure 6: Belief-Space representation of a Speaker uttering a metaphoric insult against a given Tenor. In the belief-space of the hearer, the Speaker is seen to conceptually distance himself from the Tenor.

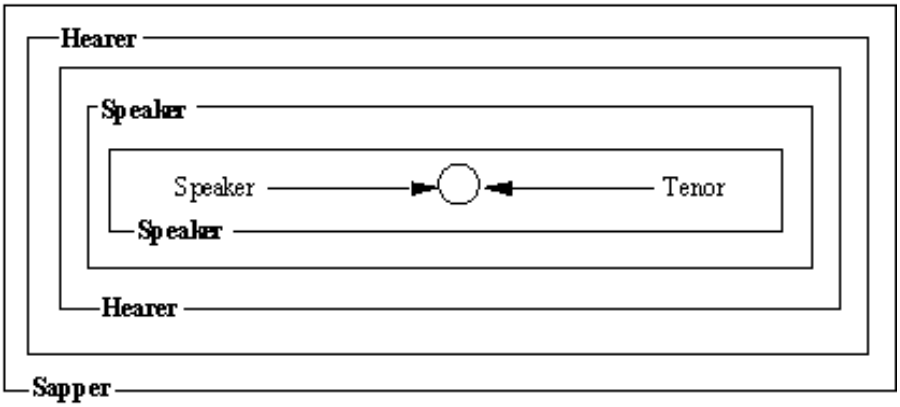

Figure 7: Belief-Space representation of a Speaker conveying metaphoric praise for a given Tenor. In the belief-space of the hearer, the Speaker is seen to move himself conceptually closer to the Tenor.

The metaphors of mind illustrated in Figures 6 and 7 can be given more formal expression as follows: if M is a metaphor uttered by S, describing T as V to H, and in doing so imparting the proposition set X to H (i.e., newly awakened bridges), then

• When M insults T, that is, $B_H\text{View}(S, V) < 0$

For every newly awakened bridge $T_x{:}V_x$ in X

When $\text{View}(T, T_x) > \text{View}(T, V_x)$ {Adopting $T_x{:}V_x$ is detrimental to T}

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto -\text{Cf}(M)$ {T distances himself from V_x in $B_H B_S$}

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto \text{Cf}(M)$ {S moves himself closer to V_x in $B_H B_S$}

When $\text{View}(T, T_x) < \text{View}(T, V_x) > 0$ {Adopting $T_x{:}V_x$ is beneficial to T}

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto Cf(M)$ *{T moves himself closer to* V_x *in* $B_H B_S$*}*

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto -Cf(M)$ *{S distances himself from* V_x *in* B_H*}*

• When M praises T, that is, B_HView(S, V) > 0

For every newly awakened bridge T_x:V_x in X

When View(T, Tx) > View(T, Vx) {Adopting T_x:V_x is detrimental to T}

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto -Cf(M)$ *{T distances himself from* V_x *in* $B_H B_S$*}*

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto -Cf(M)$ *{S distances himself from* V_x *in* B_H*}*

When View(T, T_x) < View(T, Vx) {Adopting T_x:V_x is beneficial to T}

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto Cf(M)$ *{T moves himself closer to* V_x *in* $B_H B_S$*}*

$\Delta B_H B_S B_T Z_{Tx}(V_x) \propto Cf(M)$ *{S moves himself closer to* V_x *in* B_H*}*

Given these heuristics for belief update/revision in response to metaphor interpretation, the pejorative metaphor "Surgeons are Butchers" can be analysed in terms of the speaker's desire to distance himself from a surgeon's beliefs. The system makes the natural assumption that a surgeon does not consider himself to be a butcher, in the interests of maintaining his self view, or esteem, and will therefore not lend any credibility to the comparison of surgery with slaughter, flesh with meat, or scalpels with cleavers. The speaker, however, in attacking surgeons, is perceived by the hearer to align himself fully with these beliefs.

7. Summary & Conclusions

Metaphor is an elegant and concise communicative form which is employed by a speaker as a means of conveying a state of affairs to a hearer; as such, it deserves to be analysed as a speech-act, with a particular illocutionary intent and perlocutionary effect. The state of affairs conveyed by metaphor is not always expressible in what is traditionally termed *literal language*, and thus the comprehension of metaphor is often a learning experience for the hearer, inasmuch as it requires him/her to reorganise his conceptual structures to accommodate the novel *analog-bindings* of the metaphor. Neither is this state of affairs always completely inherent in the propositional structure of the utterance; often much of the meaning conveyed by metaphor is pragmatic in nature, inasmuch as the metaphor provides the hearer with a glimpse into the belief-space of the speaker.

This paper has described an extension to the Sapper framework, a hybrid symbolic/connectionist model of metaphor previously outlined in Veale & Keane

(1993), which incorporates elements of the ViewGen belief framework presented in Wilks, Barnden & Wang (1991). This extended framework provides a suitable computational environment for analysing the illocutionary intent of the speaker, and perlocutionary effect upon the hearer, of a broad class of metaphors with an observable ameliorative/pejorative connotation. By concentrating on this class of metaphor, it is possible to speak of the overall *impression* imparted by a metaphor, avoiding commitment to particular propositional structures. This impression is easily quantified in terms of the activation dynamics of the memory network, and forms the basis for quantifying both the emotive and communicative force of a metaphor. These measures in turn underpin various belief ascription heuristics which allow the system to extract some of the pragmatic meaning conveyed by the metaphor.

In essence, this treatment represents merely a first salvo in attacking the computational modelling of metaphor, not only as a propositional form, but as a communicative act which both conveys speaker beliefs, and revises hearer beliefs. Future ground is yet to be made on this issue, however, in arriving at a model of metaphor which acknowledges the interplay between hearer and speaker belief-spaces as being essential to comprehension.

References

Black, M. (1962). Models and Metaphor: studies in language and philosophy. Ithaca, NY: Cornell University Press.

Davidson, D. (1979). What Metaphors Mean, in The Philosophy of Language, ed. A. P. Martinich. Oxford University Press: London.

Lakoff, G. & M. Johnson. (1980). Metaphors We Live By. Chigaco, Illinois: University of Chicago Press.

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

Ortony, A. (1979). The role of similarity in similes and metaphors, in Metaphor and Thought, ed. A. Ortony. Cambridge, MA: Cambridge University Press.

Richards, I. A. (1936). The Philosophy of Rhetoric. NY: Oxford University Press.

Tourangeau, R. & Sternberg, R. J. (1981). Aptness in Metaphor, Cognitive Psychology 13.

Veale, T. & M. T. Keane. (1992a). Conceptual Scaffolding: Using metaphors to build knowledge structures, in the Proc. of the 10th European Conference on Artificial Intelligence, Chichester: John Wiley & Sons.

Veale, T. & M. T. Keane. (1992b). Conceptual Scaffolding: A spatially founded meaning representation for metaphor comprehension, Computational Intelligence 8(3), 494-519.

Veale, T. & M. T. Keane. (1993). A Connectionist Model of Semantic Memory for Metaphor Interpretation, presented at the 1993 Workshop on Neural Architectures and Distributed AI, October 19-20, the Centre for Neural Engineering, U.S.C. California.

Way, E. C. (1991). Knowledge Representation and Metaphor, Studies in Cognitive systems, Kluwer Academic Publishers.

Weiner, J. E. (1984). A knowledge representation approach to understanding metaphors, Computational Linguistics, 10(1), 1-15.

Wilks, Y., J. Barnden & J. Wang. (1991). Your Metaphor or Mine: Belief Ascription and Metaphor Interpretation, in the Proc. of the 12th International Joint Conference on Artificial Intelligence, IJCAI'91. San Mateo, CA: Morgan Kaufmann.