Epistemological Pitfalls in Metaphor Comprehension:

A Comparison of Three Models and a New Theory of Metaphor

Tony Veale[1] & Mark Keane[2]
[1]*Hitachi Dublin Laboratory, Dublin, Ireland.*
[2]*Computer Science Department, Trinity College, Dublin, Ireland.*

1. Introduction

If metaphor is to be viewed as a fundamental cognitive agency, as recent work suggests, what ramifications does this view have for a model of semantic memory? This paper presents a computational treatment of metaphor comprehension, named Sapper (see Veale & Keane 1993, 1994), which is built upon a parallel, adaptive, and learning network model of semantic memory. Sapper is a hybrid symbolic/connectionist model which views the interpretation of novel metaphors as a process of connectionist *bridge-building*, a process which subsequently alters the activation dynamics between different conceptual schemata in semantic memory, thereby causing these schemata to *interact* (following Black, 1962) in a *representationally dynamic* fashion.

Sapper employs a bottom-up approach to metaphor comprehension, one which encourages the existing structure of semantic memory to shape and accommodate the most natural interpretation for each concept juxtaposition. In this way, the entirety of contingent background knowledge is brought to bear on the interpretation process. However, the Sapper mechanism combines the base-filtering stage of interpretation with the formation of initial match hypotheses in a single connectionist phase, thereby significantly curtailing the sweep of the matching process and side-stepping the factorial death that models such as the Structure Mapping Engine (SME - see Gentner 1983) can all too easily fall victim to. In fact, this paper will provide empirical evidence that SME is fundamentally unsuited to the interpretation of a broad class of metaphors that rely on an object-centred, as opposed to predicate-centred, representation. These metaphors, which often find linguistic expression as noun: noun comparisons, depend mainly upon the adequate representation of object partonomies and taxonomies, rather than the representation of actions and events toward which models such as SME are inherently biased. This evidence casts serious epistemological doubts on the validity of these models as cognitive theories of human metaphor comprehension.

Structure of This Paper

This paper addresses a single, but broadly sweeping, question: What representational demands are placed upon a model of semantic memory by a theory which views metaphor as an active and dynamic conceptual agency? This paper provides an answer in the form of a hybrid model of memory which marries the complementary strengths of both the symbolic and connectionist paradigms. Our discussion of this model will observe the following course: section two outlines the principles and mechanics of a hybrid model of semantic memory, emphasising the connectionist mechanics of network representation, and providing an elegant computational account of that appealing but infinitely nebulous view of metaphor, Black's Interaction View; various epistemological issues in metaphor and analogy are raised by this style of bottom-up analysis, and a discussion of such issues is presented in section three. A quantitative evaluation of the model is then presented in section four, where a comparative analysis with other models of metaphor and analogy is reported. The paper then concludes with an overall summary in section five.

2. Sapper: A Hybrid Model of Metaphor Interpretation

The Sapper framework, as described in Veale & Keane (1993), is a hybrid symbolic / connectionist model, embodying a basic philosophy which views the interpretation of novel metaphor as a process of connectionist *bridge-building*. From the Sapper perspective, metaphor comprehension involves the construction (or more accurately, the *awakening*) of new cross-domain linkages, which serve as bridges to bind the *analog-pairs* established by the metaphor. The novelty of a metaphor may be measured by the extent to which it adds to the structure of the network, as it is accommodated by the system. This philosophy thus views metaphor as a dynamic, constructive, conceptual phenomenon, which evokes a response in a reactive fashion from an adaptive and accommodating knowledge-base.

A Sapper network is essentially a localist graph in which nodes represent concepts, and where arcs represent semantic relations between these concept nodes. While localism is often pejoratively labelled *grandmother-cell coding*, two significant architectural features arise out of a localist rather than distributed architecture: (i) *knowledge isomorphism* - the structure of a localist network directly mirrors the semantic composition of the knowledge represented therein; and (ii) *knowledge parallelism* - in a localist network any number of different concepts (i.e., nodes) may be co-active simultaneously.

Now, while this first feature is simply a matter of design convenience, the second is arguably fundamental to the operation of metaphor, the heart of which is the apt juxtaposition of different concepts (see Harnad 1982). Sapper exploits knowledge parallelism to enable the creative reconciliation of different concept domains to occur, in a bottom-up fashion, by inferring cross-domain linkages, or *bridges*, where previously none existed (e.g., AHA! if generals are surgeons, then an operating theatre can be seen as a battlefield!).

New cross-domain bridges are laid down along *dormant* inter-concept connections, which have been established a priori by the rule-based component of the model. These dormant connections provide the possible routes along which metaphor creativity can arise. The primary role of the rule-based, symbolic, component is therefore the analysis of the network organization for particular consistencies of structure, and the augmentation of the network with new (but dormant) connections on the basis of these consistencies. Such dormant network linkages represent merely *plausible*, rather than fully *established*, semantic relations, and are thus not operative carriers of activation. The connectionist component of Sapper is responsible for the controlled propagation of activation energy, or *zorch* (see Hendler 1989), throughout the network, as it flows from the matriarch concept nodes as evoked by the metaphor. The matriarch concepts of a metaphor, following Richards (1936), are known as the *tenor*, or metaphorized subject, and the *vehicle*, or metaphorizing object. The task of the connectionist component, then, is to predict which dormant linkages should eventually be *awakened*, to become active bridges linking the domains of the tenor and vehicle concepts.

Activation waves in Sapper possess a unique signature *frequency* as well as an *amplitude* (zorch). Thus, for any given activation wave, the connectionist component can determine the matriarch node of origin, whether tenor or vehicle. As illustrated in Figure 1, when two activation waves from different matriarchs meet at a dormant linkage, the linkage is recognized as a potential cross-domain bridge between the tenor and vehicle, and the end-points of the linkage thus form the basis of an initial match hypothesis.

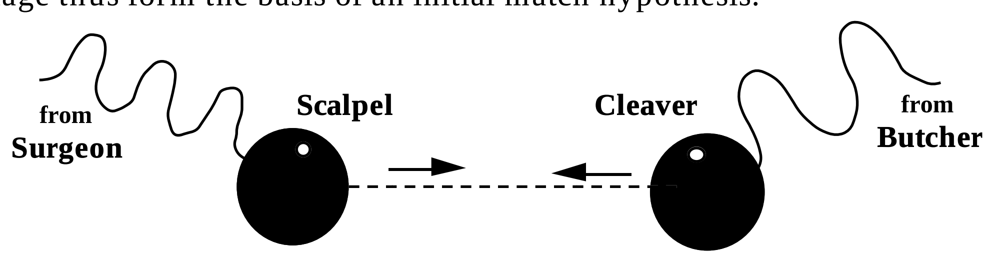


Figure 1: A dormant linkage between the concepts Scalpel and Cleaver is deemed to provide a plausible match hypothesis when it becomes a cross-over path for competing activation waves from the Surgeon (tenor) and Butcher (vehicle) concept nodes.

Figure 1 depicts such a scenario during the interpretation of the metaphor "Surgeons are butchers." Whenever waves of competing activation meet at a dormant linkage, that linkage forms the basis of an initial match hypothesis. Because dormant linkages are laid down by the triangulation and squaring rules on the basis of local consistencies of structure, the hypothesis stage is thus, in a strong sense, driven by both literal similarity and higher-order structural constraints. And because activation originates at the concept nodes of the tenor and vehicle, the connectionist phase successfully integrates the base-filtering and hypothesis formation stages of processing.

Returning to our current metaphor example, the Sapper system awakens a range of initial cross-domain bridges in response to the juxtaposition of Surgeon and Butcher, some of which are eventually recognized to be globally inconsistent, and therefore rejected (i.e., returned to a dormant state). The bridges that survive produce the following mappings:

[.86] If **Butcher** is like **Surgeon**

[.25] Then 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

The number specified in square brackets to the left of each mapping is a numeric measure, between -1 and +1, of the perceived similarity of the related concepts *after* the metaphor has been comprehended. This measure combines a metric for both the literality similarity of the related items (e.g., Cleavers and Scalpels are sharp and metallic), and the higher-order similarity that is now seen to exist between the two (e.g., the relation between Scalpel and Cleaver supports the second-order mapping Slaughter: Surgery, which in turn supports the mapping Abattoir: Operating-Theatre).

The Sapper rule component employs two distinct constructor rules to augment the knowledge-base with dormant conceptual bridges --- the *Triangulation Rule* and the *Squaring Rule*. In essence, these rules compile into the knowledge-base the top-down knowledge that is necessary to infer systematic cross-domain binding; and because this knowledge is automatically pre-compiled into the network, the connectionist phase is spared the necessity of performing global structural analysis, and is therefore adequately modelled as a spreading activation process. The *Triangulation* rule is invoked whenever two concept nodes share a common association or superclass, as is the case with the concepts SURGEON:BUTCHER & BLOOD, MEAT:HUMAN-FLESH & FLESH, SURGERY:SLAUGHTER & DEATH, OPERATING-THEATRE:ABATTOIR & LOCATION, and SCALPEL:CLEAVER & SHARP, laying down dormant linkages between the schemata BUTCHER and SURGEON, HUMAN-FLESH and MEAT, CLEAVER and SCALPEL, ABATTOIR and OPERATING-THEATRE, and SLAUGHTER and SURGERY. The *Squaring* rule, which is a second-order constructor inasmuch at it may act upon the linkages laid down by the triangulation rule, is used to *reinforce* the bridges SURGEON:BUTCHER & MEAT:HUMAN-FLESH, and SURGERY:SLAUGHTER & SCALPEL:CLEAVER. The Squaring rule builds bridges upon bridges, each new linkage extending the inter-domain *reach* of the last; in this way Sapper accounts for the phenomenon of domain incongruence discussed in Tourangeau & Sternberg (1981). The mechanics of these two rules are illustrated in Figure 2:

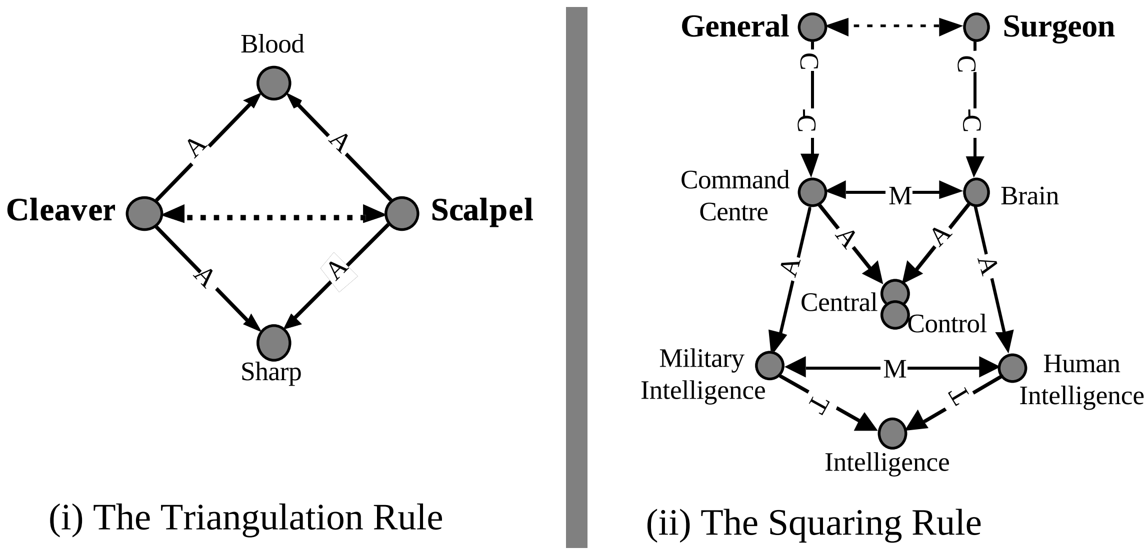
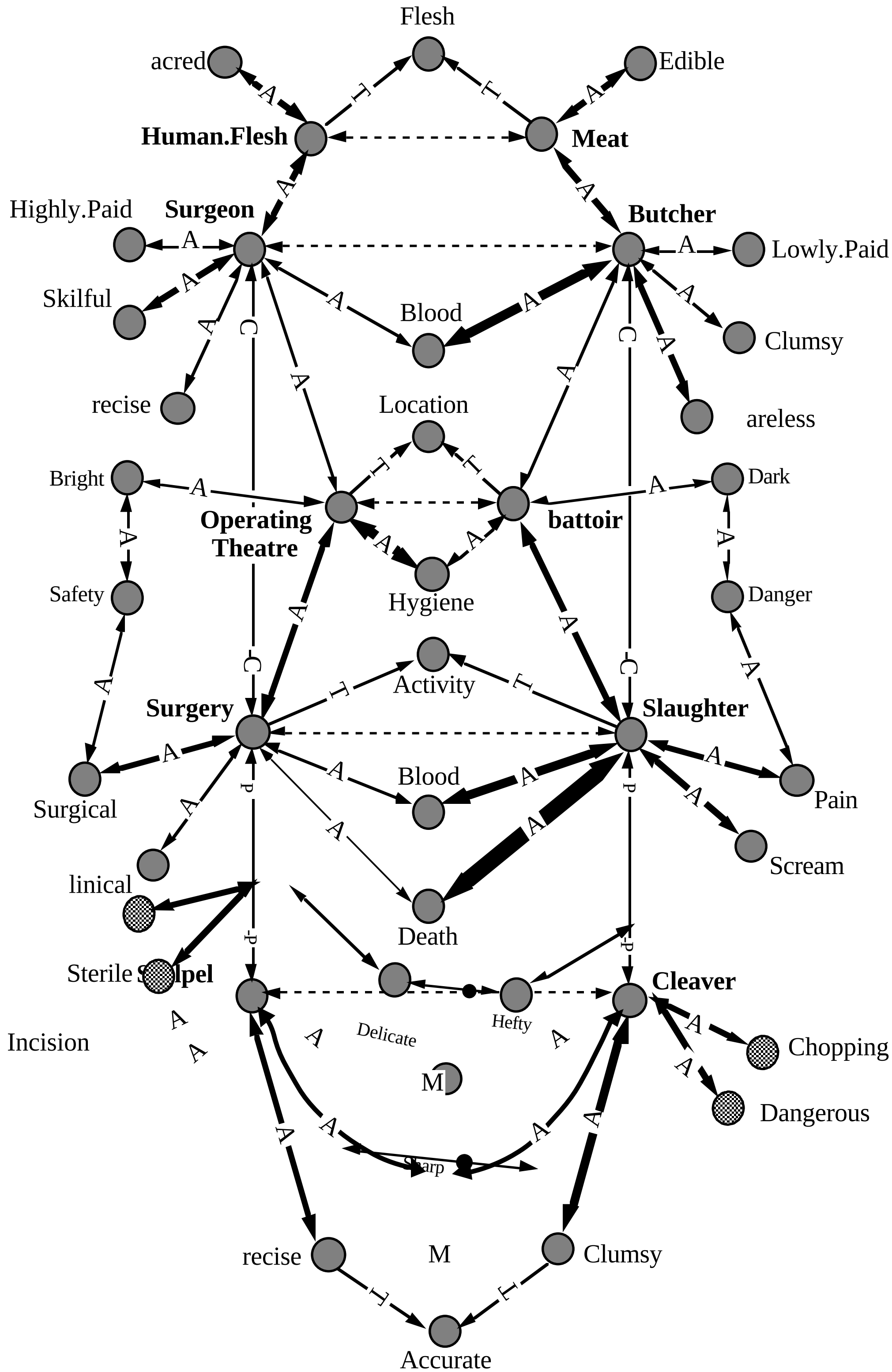


Figure 2: The Triangulation Rule (i) and the Squaring Rule (ii) augment the knowledge base with additional dormant structures, precompiled pathways that may later be used to form cross-domain analog bindings. The Triangulation rule infers new structure on the basis of shared associations, while the squaring rule employs shared metaphor bridges (previously awakened dormant bridges, or established metaphors) as an evidential basis. Key: Dashed bi-directional arrows depict dormant ***conceptual bridges****, while unbroken bi-directional arrows labelled "M" depict established metaphors, or awakened bridges; bi-directional arrows labelled "A" depict* ***Attributive / Associational*** *relations; bi-directional arrows labelled "C -C" depict* ***Control*** *relations; unidirectional arrows labelled "T" depict* ***Taxonomic*** *relations.*

The Sapper memory network fragment corresponding to the Surgeon: Butcher metaphor is illustrated in Figure 3:



Surgeons are Butchers

Figure 3: Sapper description of the metaphor "Surgeons are Butchers". Key as in Figure 2, with the following additions: • depicts an ***inverter****, the network equivalent of logical negation, which transforms inhibitory activation into excitation, and excitatory activation into inhibition; Line thickness indicates relational salience, or strength.*

The triangulation and squaring rules are wholly independent of any particular comprehension task. Full metaphor comprehension is only initiated once the matriarch nodes corresponding to the tenor (Surgeon) and vehicle (Butcher) have been *clamped*, and the connectionist component proceeds to propagate activation from these source nodes; in our current example, this causes the dormant linkages between SURGEON:BUTCHER, MEAT:HUMAN-FLESH, SURGERY:SLAUGHTER, SCALPEL:CLEAVER, and OPERATING-THEATRE: ABATTOIR to be recognized as initial match hypotheses. Sapper then directs these linkages to be *provisionally awakened*, thereby providing an evidential basis for the squaring rule to infer even higher-level structural hypotheses.

The opening of these bridges allows activation to flow freely between the tenor and vehicle domains, altering the activation dynamics of the network in such a way that the tenor *actually* interacts with the vehicle at a conceptual level. This network effect is illustrated in Figure 4. The activation patterns of BUTCHER, SLAUGHTER, MEAT and CLEAVER interact with those of SURGEON, SURGERY, HUMAN-FLESH and SCALPEL to produce a response to the metaphor. Overall, a Surgeon is seen, through the lens of the metaphor, to be an altogether less skilful and precise tradesman, performing surgery which is akin to the slaughter of innocents, while amidst the pain and screams of fear, wielding a blood-stained scalpel to slash and chop liberally into human *meat*. A graphic interpretation to be sure, but one that, in following the interaction view of Richards (1936) and Black (1962), operates both ways. The metaphor denigrates surgeons, but elevates butchers, who are now seen to be that much less clumsy, careless and imprecise, and altogether more professional.

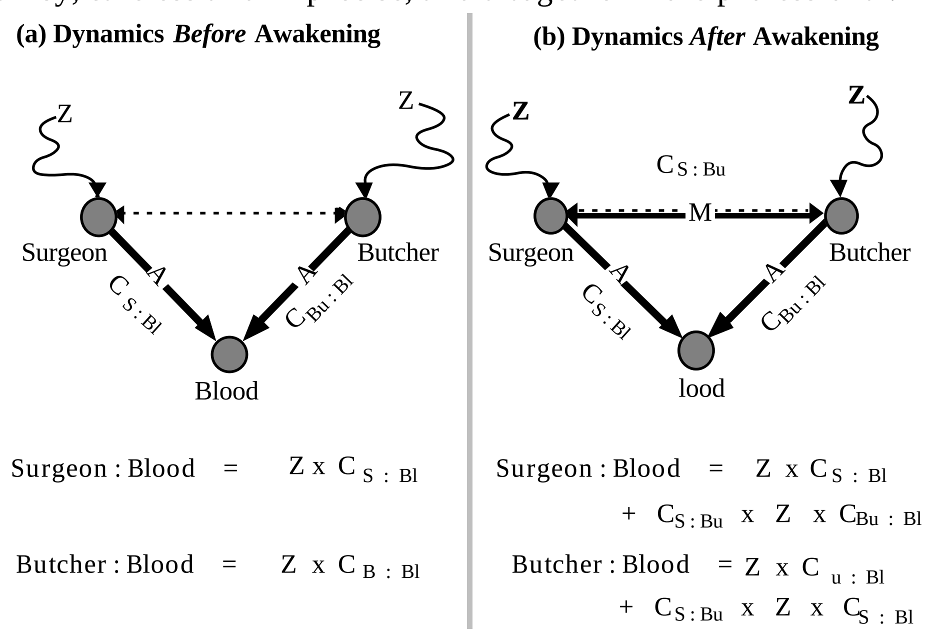


Figure 4: The awakening of a new bridge alters the activation dynamics of the network, allowing zorch (Z) to flow freely between two domains.

2.1 Mapping Systematicity

Due to the heuristic nature of the triangulation and squaring rules, and the local nature of the hypothesis formation process, the connectionist phase will most likely yield more match hypotheses than are needed to produce a final interpretation of the metaphor. Many of these match hypotheses will be *ghosts* - analog mappings that are valid in some other metaphor but not in the one currently under consideration - or *noise*, mappings which contradict other, more systematic mappings and thereby diminish the overall coherence of the interpretation. It is therefore necessary to weed out these undesirable mappings, to produce an interpretation that is wholly coherent with itself and with the contents of semantic memory.

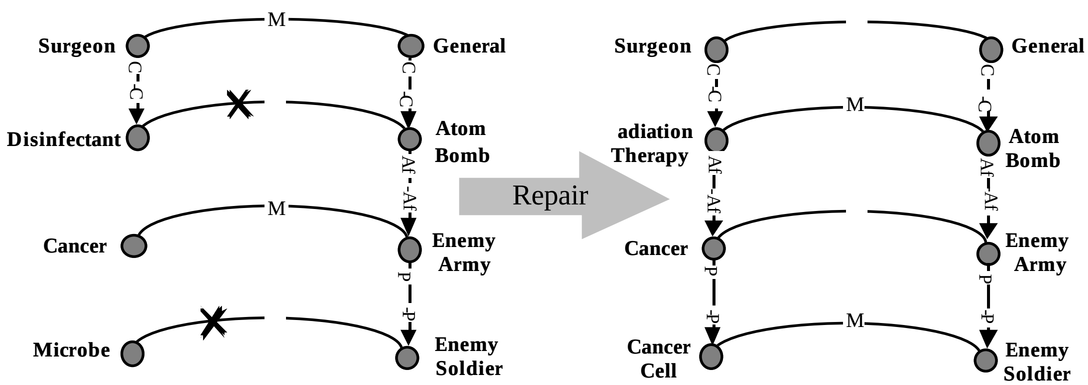


Figure 5: In analysing the Surgeon: General metaphor, an initial seed mapping-lattice is created by heuristically ordering individual match hypotheses; this lattice may be globally inconsistent, so unsystematic mappings are recognized and repaired. Key: Arcs labelled “Af -Af” depict Affect relations, while “P -P” depict partonomic links.

This task is achieved in the following manner: first, a seed interpretation is produced, by ranking each match hypothesis according to a heuristic measure of systematicity and literal similarity, and by provisionally rejecting any mapping which competes with a better ranked mapping (for instance, in comparing Generals to Surgeons, Scalpel: Snub-Fighter beats out Scalpel: B-52). This seed interpretation thus contains a range of mappings which are, individually, the best possible mappings for each concept of the tenor domain, but taken collectively as a meaning structure, these mappings may not work at all coherently together. A repair phase is thus evoked, whereby the weakest unsystematic mapping (according to the previously mentioned heuristic measures) is undone, and replaced with the next best mapping for the tenor concept concerned. This repair phase, illustrated in Figure 5, continues until all mappings are wholly consistent with each other, and a maximally systematic interpretation is produced.

3. Epistemological Issues in Metaphor Comprehension

Linguistic metaphors come in different syntactic guises: qualification metaphors of the adjective: noun variety, object-centred metaphors of the noun: noun variety, and predicate-centred metaphors of the verb: verb variety. While it is to be expected that each juxtapositional form is interpreted relative to the same knowledge structures, full comprehension of each form may stress different aspects of those structures than others, and overall, adopt a different epistemological perspective on the organization of memory.

Surprisingly, even metaphor theories which claim to be generalized structure matchers, such as SME, the Structure Mapping Engine of Falkenhainer, Forbus and Gentner (1989), embody hard-wired epistemological biases which make them more suited to one guise of metaphor over another (the claim that SME is amenable to metaphor comprehension is made in Gentner, Falkenhainer & Skorstad 1989). This section will set the stage for the empirical demonstration of the next section, arguing that objected-centred metaphors of the noun: noun variety comprise a significant Achille's heel of SME. As our analysis will reveal, this weakness derives from an epistemological bias in SME's design, which leads the unwary matcher into factorial death for certain types of domain structure in which hierarchical organization is present but implicit. This factorial demise is also exacerbated by SME's refusal to employ literal similarity as a constraint on the creation of initial match hypotheses. The empirical results also suggest that such a refusal to effectively limit the number of initial match hypotheses also ham-strings the connectionist-flavoured ACME model of Holyoak & Thagard (1989). This evidence casts serious epistemological doubts on the validity of these models both as generalized matching systems, and as cognitive theories of human metaphor comprehension.

Object-Centred Representations

Metaphors which contrive the juxtaposition of two object concepts, such as Surgeon & Butcher, or Car & Rocket, will more readily exploit the object-centred aspects of these concepts in memory. That is to say, those aspects of conceptual representation which emphasise objects and entities over actions and events, such as meronomic, taxonomic and associational structures, make the greatest contribution to the analysis of noun: noun metaphors. This form of representation is illustrated in the example metaphor of Figure 6:

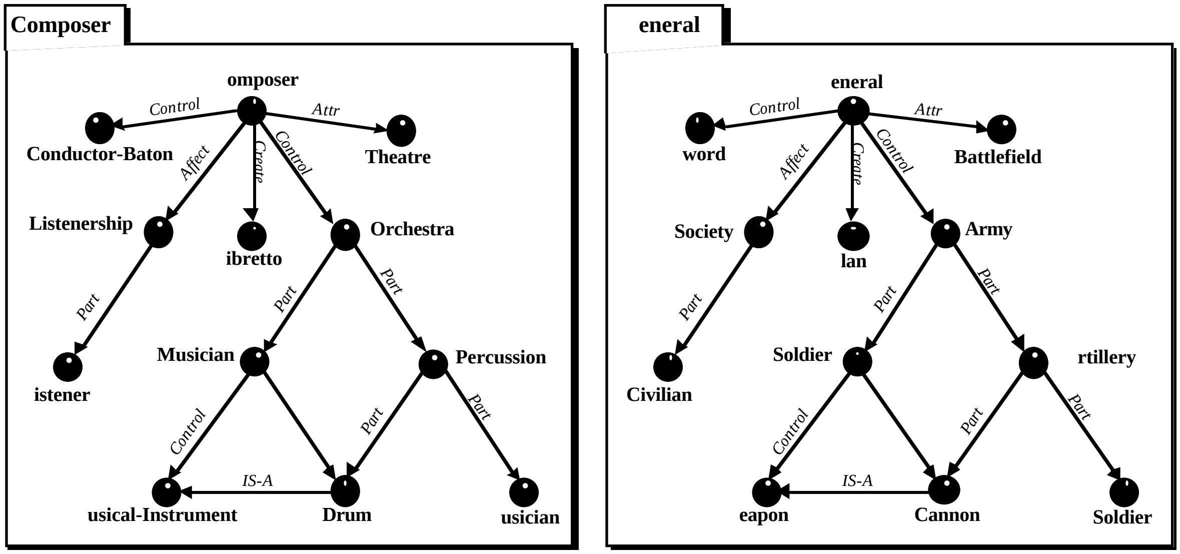


Figure 6: Object-Centred Representations for the concept schemata of Composer and General. Notation: The labelling scheme differs from that employed in previous diagrams, as our discussion is to be a general one concerning representation issues, and is not specific to the Sapper mechanism.

The epistemological style of Figure 6 may be considered object-centred inasmuch as object/entity nodes establish the foreground of the representation, while inter-concept relations are strung between these nodes like tinsel on a Christmas tree. The significant point to note about this representational form is that linkages as well as nodes are labelled (as in Sapper), and thus, both types of graph component may be considered to be first-class representational elements.

Scenario-Centred or Predicate-Centred Representations

Whereas object-centred metaphors tend to focus upon the entities of a domain, scenario-based or predicate-centred metaphors are naturally more disposed toward the relational structure that exists between these entities. Curiously enough, this shift in emphasis is achieved by denying labelling privileges to the linkages which connect the nodes of the representation, forcing the knowledge-base to encode both actions and objects as nodes. The linkages of a predicate-centred representation exist only to tie the argument nodes of a predicate to the predicate node itself, in effect then establishing a graph notation for the predicate calculus. For in much the same way that arguments go unlabelled in a predicate calculus expression, relying wholly upon ordering constraints to define their relationship to the governing predicate, the nodes of a predicate-centred representation also rely upon explicit surface ordering to convey implicit deep relationships. This style of representation is illustrated in Figure 7:

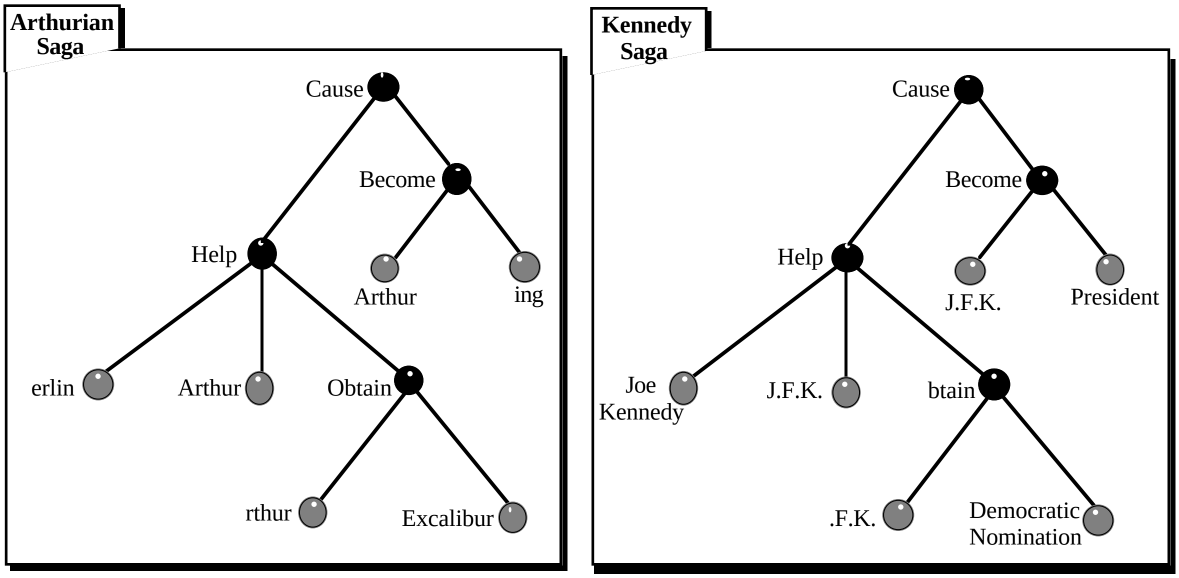


Figure 7: Predicate-centred representations for the (partial) concept schemata of Arthurian-Saga and Kennedy-Saga. Notation: Black nodes represent predicates, while grey nodes represent entities.

As conveyed in Figure 7, the means by which Arthur became King of England are analogical to those pursued in Kennedy's ascension to the American presidency, for the same abstract causal structure is present in both domains. It is toward this form of mapping problem that predicate-centred represents are best suited, problems where one hierarchical relational structure, essentially a nested predication, is matched with another. In a predicate-centred model of analogy such as Gentner's SME, predicate nodes (depicted in black) are mapped under a strict identicality constraint, while entity nodes (depicted in grey) may conceivably map onto any entity node in the target domain, provided the 1-to-1 coherence of the overall mapping is preserved. Entity nodes are therefore almost incidental to the mapping process in a predicate-centred representation, as it is the predicates and their relationship to each other that ultimately determine the basis of the analogy.

Support Relations

While different models of metaphor and analogy may place different degrees of significance on the role of structural isomorphism, most theories (such as SME, ACME and Sapper) ultimately acknowledge that analogy and metaphor are, by and large, structure-preserving processes. Philosophically, this is an unavoidable, indeed almost tautological, position to assume, for in any formal/computational system, meaning is necessarily explicated in structural terms, so where structure is not respected, neither is meaning. In the ACME approach, the structure of the tenor and vehicle domains dictates the structure of the constraint network that is especially constructed for the interpretative task at hand; the conceptual structure of the tenor and vehicle is thus the major source of constraints upon the mapping process. In the Sapper approach, the conceptual structure of the tenor and vehicle domains not only provides the evidential basis of the triangulation and squaring rules, but also defines the pathways along which activation energy will flow to awaken new cross-domain bridges. And in the SME approach, the hierarchical structure of the tenor and vehicle domains is used to constrain, in a top-down fashion, the combinatorial possibilities of cross-domain entity mappings. Each model therefore, in its own way, looks to the structural make-up of the tenor and vehicle to indicate what partial mappings support other partial mappings, and which partial mappings are irreconcilable with each other.

A metaphor interpretation system must, therefore, take its support relations wherever it can find them. However, certain models, such as SME and ACME, place far too much emphasis on the importance of hierarchical support as manifested by nested predicate structures, to the detriment of other forms of support relation. These models exhibit a serious *epistemological blind-side*, an oversight which makes them inherently unsuited to the interpretation of noun: noun metaphors and any other form of comparison that requires an object-centred representation. To see why this must be so, consider the predicate-centred re-representation of a domain that is more amenable to object-centred organisation, as illustrated in Figure 8:

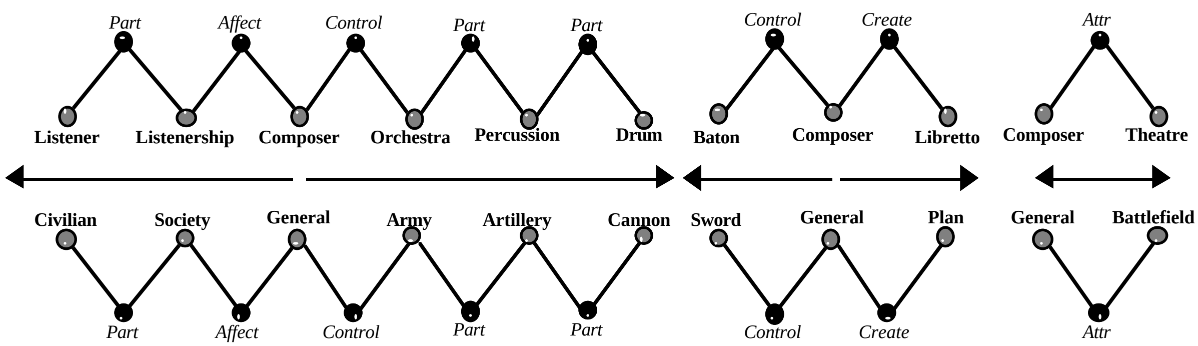


Figure 8: A Predicate-centred view of the domains of Composer and General. Arrows are provided to indicate the direction in which predicate arguments should be read. Compare these flat structures with the hierarchical depth of the equivalent object-centred views illustrated in Figure 6.

The most striking quality of the representations in Figure 8 must surely be the obvious lack of hierarchical depth - gone are the tree structures of Figure 6, rich in vertical support relations, to be replaced by a shallow and unnested representation that favours horizontal rather than vertical expansion. However, all that has been performed here is an alternative depiction of the same knowledge structures - a graph re-labelling in which no knowledge has been added or removed - and thus, on a wholly syntactic level, the form may have changed but the meaning content has been preserved. Effectively, the support structures which found vertical expression in an object-centred representation have simply been *sheared* to find horizontal expression in a predicate-centred representation, while the support relations which were manifest hierarchically in the former are now present in the latter in a *sideways systematic* fashion. However, models such as SME and ACME which are locked into a particular form of processing, and are algorithmically predisposed to seeking support structure in hierarchical rather than horizontal organization, fail completely to acknowledge this sideways systematicity.

In the terminology of SME, each predicate node in Figure 8 (illustrated in black) comprises a *root mapping* in the overall analogy, inasmuch as each predicate is ungoverned by a higher-order relation (in contrast, say, with the structures of Figure 7). Unfortunately, the SME algorithm includes a *root merging* stage whose complexity is factorially dependent upon the number of such roots (this merge stage is described in Falkenhainer, Forbus, & Gentner 1989). Clearly, if every predicate in the tenor and vehicle domain is to be considered a root, then SME quickly descends into factorial hell. Likewise, the excitatory and inhibitory linkages which codify the support relations in an ACME constraint network are laid down on the basis of hierarchical support; if this support is not perceived in a sideways fashion, the constraint network may become underdetermined. However, unlike the structure-bound SME, the connectionist ACME employs a form of spreading activation which can horizontally *cut across* structures, thereby enforcing in some astructural manner the sideways systematicity inherent in the representation. In effect, spreading activation facilitates a form of *lateral thinking*, in which one part of the network may reinforce another, even when these regions are not connected by some overarching hierarchical relation. Consider again the example of Figure 8: activation may propagate upwards and sideways, allowing the network to transcend the limitations of simple vertical support, to additionally exploit horizontal support relations. However, while more epistemological sure than SME, a more secure case against ACME can nevertheless be made on grounds of scalability, and such a case is addressed in the next section.

Epistemological Commitments

In the final analysis, then, noun: noun metaphors demand an object-centred representation if the support relations manifest in the structural make-up of the tenor and vehicle domains are to be made explicit. If the sample metaphor of Figure 7 illustrates the suitability of a predicate-centred representation for verb: verb comparisons, the example of Figure 8 clearly demonstrates the error of shoehorning noun: noun metaphors into such a representation. Indeed, while an object-centred representation is readily augmented to accommodate verb: verb metaphors, no simple additions can be made to a predicate-centred representation to accommodate noun: noun metaphors. Any such changes would shake at the very roots of the host theory, and substantially reformulate that theory to the extent that it bore no resemblance to the original. It would seem then that the representation of choice in the Sapper model, an object-centred localist network, is best suited to the rigours of a metaphor interpretation system that is to tackle figurality in its variety of linguistic forms. This view is borne out in the experimental evaluation described in the next section, in which such doubts regarding the dubious epistemological foundations of the SME and ACME approaches are given quantitative form.

4. Quantitative Evaluation and Analysis

The quantitative evaluation of Sapper described in this section was carried out within a sample memory network containing 284 concept nodes and 1597 user-specified inter-concept relation links. This network represents a description of the profession domain, and contains conceptual schemata for fifteen different profession types, such as Surgeon, Butcher, Scientist, and so on. The total number of automatic inferences generated by Sapper while assimilating this network description is 2299 dormant conceptual bridges, all of which are created using the triangulation rule, as higher-order inferences only occur during metaphor comprehension. On average then, each profession schemata comprises 19 localist concept nodes, and 106 inter-node linkages. Of these linkages, 36 codify high-level relations (such as Cause and Depend), and 70 codify attributions and taxonomic orderings; while this distinction is an artificial one which means nothing at all to Sapper, the distinction is nevertheless considered valid in the context of SME. The following profession types are defined in this network:

Surgeon	Priest	Politician	Chef	Magician
Butcher	Scientist	Criminal	Accountant	Composer
General	Architect	Hacker	Author	Sculptor

The following network linkages are used to encode semantic relations within and amongst these conceptual schemata:

Attribute	Effect	Affect	Down	Agent
Control	Instrument	Create	Manner	Patient
Part	Predicate	Connect	Source	Event
Substance	Aspect	Disconnect	Target	Character
Purpose	Perform	Up	Method	Metaphor

The linkages Effect, Manner, Method, Target, Purpose, Event, Character, Agent, Patient and Instrument are provided for the description of verb-centred scenarios, for example, the JFK historical setting, or the King Arthur saga (the latter being a good metaphor for the former), while the linkage Metaphor is provided to allow the knowledge engineer to encode conventional metaphors into the network (such as the *Family Tree* and *Gene Pool* metaphors).

A Quantification of Automatic Inference

Given such a test-bed environment, the scene is set to ask some basic empirical questions regarding the profligacy of automatic inference, as performed by the triangulation and squaring rules. For instance, what kind of worst-case and average-case scenarios can be imagined here? Even if Sapper were to push the limits of prodigality in the creation of dormant linkages, the worst case scenario involves the generation of $n*(n-1)/2$ dormant linkages, where n is the number of localist nodes in semantic memory; large to be sure, but hardly nightmarish in its extent. And as it happens, the average case performance is much lower than this ceiling, the effect of well-defined domain structure being to considerably reduce the possibilities of inter-node bridging. This situation is illustrated in Figure 9, in which the bridging overhead of our profession test-bed network is graphed:

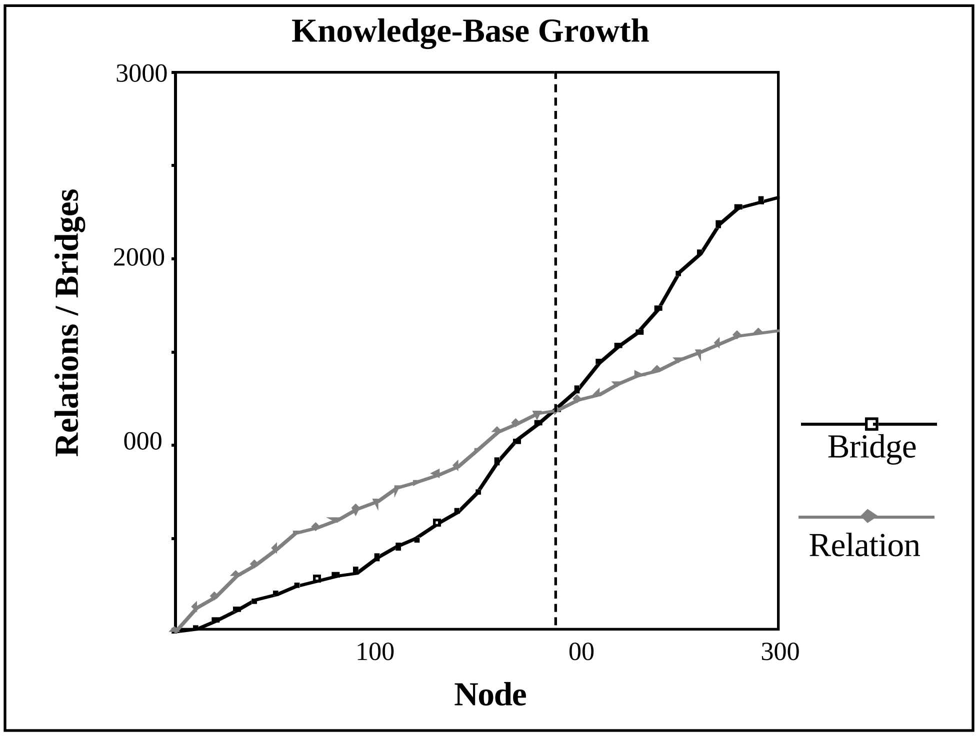


Figure 9: Graph of user-specified concept relations against system-inferred conceptual bridges.

Note that the network seems to reach a critical mass at just under 200 nodes, at which point the system begins to out-produce the knowledge-engineer in adding linkages to the network. But when the network reaches 300 concept nodes in size, the worst-case model predicts that 44850 dormant linkages will have been automatically inferred, while the real situation presents a more tractable picture: just over 2000 such linkages, less that five percent of the worst case scenario, are actually added to the network. Sapper would thus seem to exhibit remarkable thrift in its automatic inference capabilities.

Of course, these bridges do not necessarily have to be held in the system until they are used, if ever; rather, they might be created dynamically as the metaphoric context dictates. This is achievable by delaying application of the triangulation rule until the spreading activation phase of metaphor interpretation - as each node is newly activated, the triangulation rule is applied to that node and to those other nodes which are currently active within the scope of the rule. This is the classic computational trade-off in space versus time, whereby the storage costs of maintaining many, potentially useless, bridges is translated into processing costs at run-time.

An Experiment

This profession network provides a suitable test environment in which to perform a comparative evaluation of Sapper, SME and ACME; disregarding issues of metaphor symmetry, focusing instead on the concerns of structural consistency and mapping coherence, fifteen different concept descriptions yield one hundred and five different figurative comparisons. Running upon this pool of 105 test metaphors, Sapper achieves a respectable average interpretation time of 12.5 seconds per metaphor. The incremental case, wherein the same metaphors are presented to the system a second time, shows that this time drops to 7 seconds for metaphors that have been previously encountered. Neither SME or ACME provide any results whatsoever, for any of the test metaphors, even when they are allowed to run for days at a time, yielding an effective destruct-test of Allegro Common Lisp[1] but little else. Eventually then, unable to obtain performance ratings for SME or ACME on this test network, a tabulation the number of initial match hypotheses generated by each model is instead provided. This data proves to be revelatory in that it clearly demonstrates the cause of failure in SME and ACME - too many initial match hypotheses cause each model to become intractable in later stages of processing. This experimental finding is illustrated in Figure 10:

[1] These systems, LISP implementations of which were obtained from the approved FTP sites, were allowed to run on a SPARC 2 platform in the Allegro Common-Lisp development environment. Tests were abandoned when it was realized that these systems would continuously exhaust the garbage collection facilities of the environment before ever returning a result for a single metaphor. For instance, SME was allowed run on the Surgeon: General metaphor for over fifty hours before crashing the LISP host.

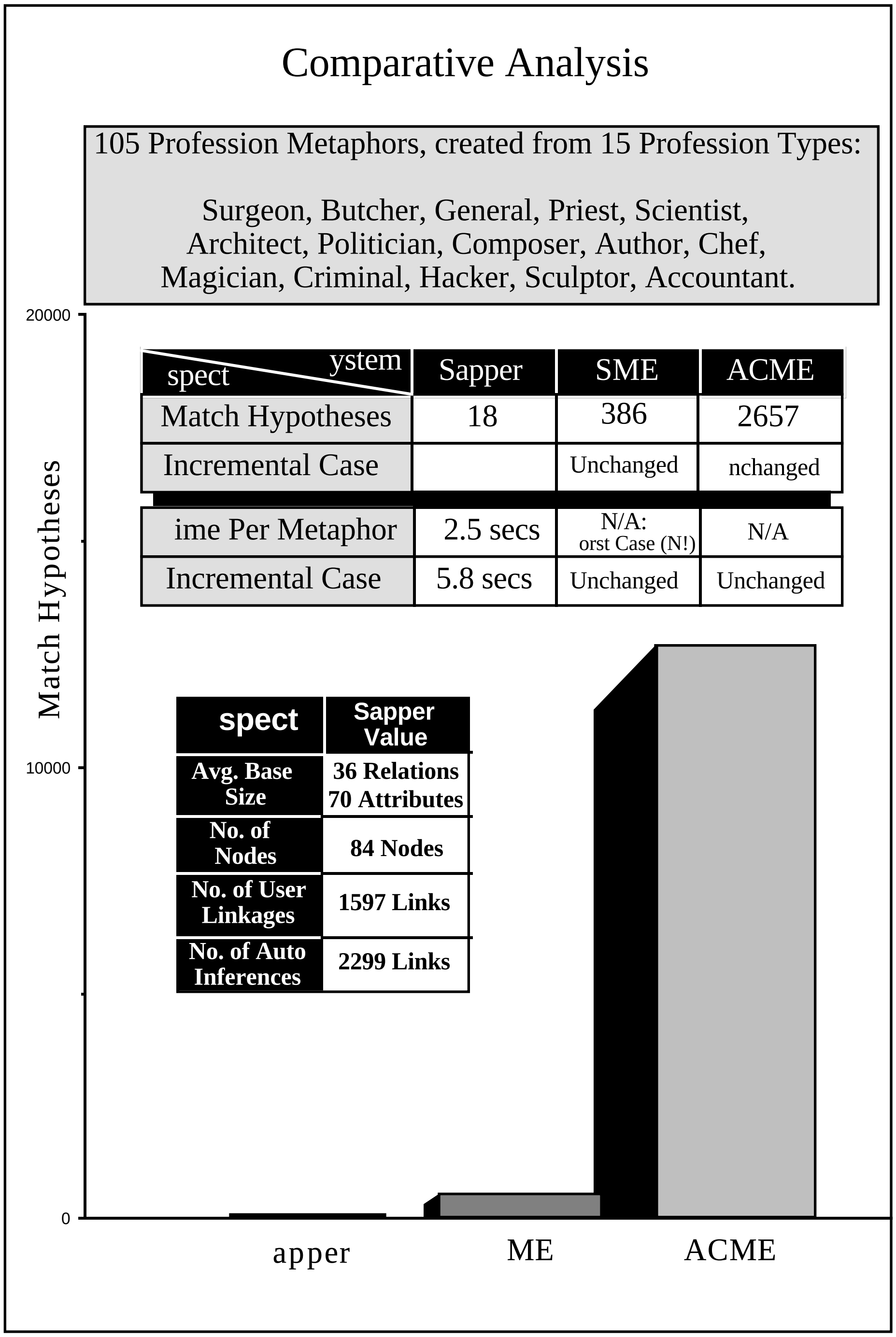


Figure 10: Comparative evaluation of Sapper, SME and ACME as determined within the Profession test network. Note that the unavailability of time figures for SME and ACME reflects the inability of these models to generate a result in real time.

As can be seen in Figure 10, Sapper on average generates considerably less match hypotheses than SME (18 versus 386), while SME in turn is significantly more thrifty than ACME (386 versus 12657). Sapper's hypothesising thrift arises from its dependence on conceptual bridges, which are only ever created when there exists a first-order (literal) or higher-order (creative) similarity between two concepts. SME, however, employs no such notion of similarity and instead prefers to temper match selection using the notion of predicate identicality. ACME is the prodigal son of the trio, eschewing even predicate identicality and instead favouring the weaker constraint of arity-matching. No wonder then that ACME fails to generate a mapping for any metaphor; while not as prone to the epistemological blind-sight that so afflicts SME, ACME instead suffers from *network bloat* - the constraint network developed for each metaphoric comparison is simply to large to be resolved in any reasonable amount of time.

Although considerably less prodigal than ACME in generating match hypotheses, SME is nevertheless inherently factorial in its root gmap merge stage of processing. Now, because of its inability to determine sideways support relations in an object-centred domain (of which the profession test-bed is an exemplar), each linkage in every concept description forms the basis of a root gmap, and thus, SME is factorial over the number of linkages in each concept description. To appreciate the complexity implications of this, consider the example metaphor General as Surgeon - this metaphor causes SME to generate 398 root gmaps. The worst case scenario for the merge stage is that it is order 398!, a nightmarish scenario if ever there was one. To further appreciate the enormity of this number, consider that if an analogy machine were capable of performing one million gmap merges a second, the General-Surgeon example might still occupy this machine for U^{50} years, where U is a conservative estimate of the age of the universe at fourteen billion years.

Counter-Arguments from the SME Camp

Consider the following rebuttal: SME does not work upon these examples simply because they are structurally-impoverished and causally-deficient[2]. In short, the test metaphors are contrived and thus invalid because any hierarchical causal structure has been removed, but if such causal structure were correctively added to these examples, SME would perform more than adequately upon them. This line of argument is fallacious for the following reasons:

(i) In no way should these object-centred domains be considered structurally or causally lacking. The network linkages Effect, Control, Perform, Method, Purpose and so on are provided to explicitly represent causal relations, and

[2] Personal communication from Ronald Ferguson, an active worker in the SME *tradition*, during Cognitive Science 1994 held in Georgia Tech, Atlanta.

these may be stringed together, in a sideways fashion, to construct complex causal structures.

(ii) A suitable graphical representation in which both nodes and links are labelled demonstrates that there is indeed hierarchical structure inherent in these object-centred domains. It is simply a matter for the analogy system in question to seek out this structure using the correct *filter* (e.g., vertical versus horizontal support).

(iii) Even if these examples *were* structurally impoverished, would this excuse SME's apparent ultra-sensitivity to representation? No, at best the SME theory should provide an algorithmic basis for determining the suitability of its inputs in advance, rather than having to spend geological time demonstrating the deficiency of the representation.

(iv) An empirical demonstration of the presence of this structure, and its sufficiency for analogical mapping, is provided by the Sapper mechanism and its ability to recognize and map this structure.

In summary then, these results show the SME approach to be seriously lacking in mapping competence, and not at all the general-purpose matching algorithm it purports to be (see Falkenhainer, Forbus & Gentner 1989).

5. Summary and Conclusions

This chapter has presented both the driving motivations and basic philosophy underlying the Sapper model of memory for metaphor comprehension. As stated in the introduction, a computational treatment of metaphor as a first-class cognitive phenomenon, which addresses the various signature phenomena debated in the literature, raises a number of interesting issues concerning knowledge representation. Namely, how is systematicity among different conceptual schemata from different domains to be enforced? How best should these schemata be organized to ensure that the most natural interpretation of a metaphor emerges from, or is shaped by, the existing knowledge-base, rather than being eked out by a dedicated metaphor processor? What characterises *learning* in the interpretation of novel metaphors, and how should this learning be constrained to occur in a systematic manner? How can metaphors be reified to the status of active conceptual entities, such that they dynamically strive to impose themselves upon incoming schemata, and in doing so elaborate themselves further?

A hybrid model has been presented to address these issues, marrying the complementary strengths of the symbolic and connectionist paradigms to combine both high-level structural inference with low-level opportunistic activation flow. The Sapper model of connectionist bridge-building, it is argued, provides a computational framework that is truest to the interaction view of metaphor, as advocated by Richards and Black, while explicating the manner in which metaphors move from *attention-winning* novelty to trite *conventionality*. Sapper operates in a bottom-up fashion, and thereby complements the top-down strategy of conceptual scaffolding developed in Veale & Keane (1992a,b) to yield a comprehensive account of the metaphor phenomenon, at both a lexical semantics and a deep conceptual level.

Sapper also provides a computational account of metaphor creativity that is essentially based upon the exaggeration of domain incongruences. Local similarities of a literal nature, initially established using the triangulation rule, are magnified by repeated application of the squaring rule, in the appropriate network contexts, to generate higher-order similarities that simply did not exist before interpretation of the metaphor. In this way, Sapper offers a cognitively appealing view of metaphor as a creative force that invents, rather than simply observes, new associations between concepts.

Sapper also provides a flexible network model that is amenable to the interpretation of metaphors in different linguistic guises, whether noun: noun, adjective: noun or verb: verb. Each of these forms are interpreted relative to an object-centred representation which highlights some epistemological concerns about the structural foundations of the SME and ACME models. While Sapper is capable of handling metaphors that prefer object-based descriptions (such as General: Surgeon) and predicate-based descriptions (such as Kennedy-Saga: King-Arthur-Saga), SME is shown to competent with the former only, becoming totally unhinged when expected to deal with metaphors of the latter variety. This result casts a dark shadow not only over SME's claim to be a generalised structure mapping engine, but also upon any serious pretensions to cognitive plausibility its creators may have entertained.

References

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

Falkenhainer, B., Forbus, K. D., & D. Gentner. (1989). Structure-Mapping Engine. *Artificial Intelligence*, 41, pp 1-63.

Gentner, D. (1983). Structure Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7(2), pp 155-170.

Gentner, D., B. Falkenhainer & J. Skorstad. (1989). Metaphor: The Good, The Bad and the Ugly, in *Theoretical Issues in Natural Language Processing*, ed. Yorick Wilks. Hillsdale, NJ: Lawrence Erlbaum Associates.

Harnad, S. (1982). Metaphor and Mental Duality, in *Language, Mind and Brain*, T. Simon & R. Scholes, eds.

Hendler, J. A. (1989). Marker Passing over Micro-Features: Toward a Hybrid Symbolic / Connectionist Model, *Cognitive Science* **13**(1).

Holyoak, K. J. and P. Thagard. (1989). Analogical Mapping by Constraint Satisfaction, *Cognitive Science* 13, pp 295-355.

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

Tourangeau, R. & R. Sternberg (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 *Neural Architectures and Distributed AI Workshop*, October 19-20 1993, Centre for Neural Engineering, U.S.C. California.

Veale, T. & M. Keane. (1994). Belief Modelling, Intentionality and Perlocution in Metaphor Comprehension, in *the Proc. of the Sixteenth Annual Meeting of the Cognitive Science Society*, Atlanta, Georgia. Hillsdale, NJ: Lawrence Erlbaum.