CorMet: A Computational, Corpus-Based Conventional Metaphor Extraction System

Zachary J. Mason*
Brandeis University

CorMet is a corpus-based system for discovering metaphorical mappings between concepts. It does this by finding systematic variations in domain-specific selectional preferences, which are inferred from large, dynamically mined Internet corpora.

Metaphors transfer structure from a source domain to a target domain, making some concepts in the target domain metaphorically equivalent to concepts in the source domain. The verbs that select for a concept in the source domain tend to select for its metaphorical equivalent in the target domain. This regularity, detectable with a shallow linguistic analysis, is used to find the metaphorical interconcept mappings, which can then be used to infer the existence of higher-level conventional metaphors.

Most other computational metaphor systems use small, hand-coded semantic knowledge bases and work on a few examples. Although CorMet's only knowledge base is WordNet (Fellbaum 1998) it can find the mappings constituting many conventional metaphors and in some cases recognize sentences instantiating those mappings. CorMet is tested on its ability to find a subset of the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991).

1. Introduction

Lakoff (1993) argues that rather than being a rare form of creative language, some metaphors are ubiquitous, highly structured, and relevant to cognition. To date, there has been no robust, broadly applicable computational metaphor interpretation system, a gap this article is intended to take a first step toward filling.

Most computational models of metaphor depend on hand-coded knowledge bases and work on a few examples. CorMet is designed to work on a larger class of metaphors by extracting knowledge from large corpora without drawing on any hand-coded knowledge sources besides WordNet.

A method for computationally interpreting metaphorical language would be useful for NLP. Although metaphorical word senses can be cataloged and treated as just another part of the lexicon, this kind of representation ignores regularities in polysemy. A conventional metaphor may have a very large number of linguistic manifestations, which makes it useful to model the metaphor's underlying mechanisms. CorMet is not capable of interpreting any manifestation of conventional metaphor but is a step toward such a system.

CorMet analyzes large corpora of domain-specific documents and learns the selectional preferences of the characteristic verbs of each domain. A **selectional preference** is a verb's predilection for a particular type of argument in a particular role. For instance, the object of the verb *pour* is generally a liquid. Any noun that *pour* takes as an

* Computer Science Department, Waltham, MA 02134. E-mail: zmason@amazon.com.

an object is likely to be intended as a liquid, either metaphorically or literally. CorMet finds conventional metaphors by finding systematic differences in selectional preferences between domains. For instance, if CorMet were to find a sentence like *Funds poured into his bank account* in a document from the *FINANCE* domain, it could infer that in that domain, *pour* has a selection preference for financial assets in its subject. By comparing this selectional preference with *pour*'s selectional preferences in the *LAB* domain, CorMet can infer a metaphorical mapping from money to liquids. By finding sets of co-occuring interconcept mappings (like the above mapping and a mapping from investments to containers, for instance), Cormet can articulate the higher-order structure of conceptual metaphors. Note that Cormet is designed to detect higher-order conceptual metaphors by finding *some* of the sentences embodying *some* of the interconcept mappings constituting the metaphor of interest but is not designed to be a tool for reliably detecting all instances of a particular metaphor.

CorMet's domain-specific corpora are obtained from the Internet. In this context, a domain is a set of related concepts, and a domain-specific corpus is a set of documents relevant to those concepts. CorMet's input parameters are two domains between which to search for interconcept mappings and, for each domain, a set of characteristic keywords.

CorMet is tested on its ability to find a subset of the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991), a manually compiled catalog of metaphor. CorMet works on domains that are specific and concrete (e.g., the domain of *finance*, but not that of *actions*). CorMet's discrimination is relatively coarse: It measures trends in selectional preferences across many documents, so common mappings are discernible. CorMet considers the selectional preferences only of verbs, on the theory that they are generally more selectively restrictive than nouns or adjectives.

It is worth noting that WordNet, CorMet's primary knowledge source, implicitly encodes some of the metaphors CorMet is intended to find; Peters and Peters (2000) use WordNet to find many *artifact/cognition* metaphors. Also, WordNet enumerates some metaphorical senses of some verbs. CorMet does not use any of WordNet's information about verbs and ignores regularities in the distribution of noun homonyms that could be used to find some metaphors.

The article is organized as follows: Section 2 describes the mechanisms by which conventional metaphors are detected. Section 3 walks through CorMet's process in two examples. Section 4 describes how the system's performance is evaluated against the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991), and Section 5 covers select related work.

2. The Metaphor Extraction Engine

2.1 Searching the Net for Domain Corpora

Ideally, CorMet could draw on a large quantity of manually vetted, highly representative domain-specific documents. The precompiled corpora available on-line (Kucera 1992; Marcus, Santorini, and Marcinkiewicz 1993) do not span enough subjects. Other on-line data sources include the Internet's hierarchically structured indices, such as Yahoo's ontology (www.yahoo.com) and Google's (www.google.com). Each index entry contains a small number of high-quality links to relevant Web pages, but this is not helpful, because CorMet requires many documents, and those documents need not be of more than moderate quality. Searching the Internet for domain-specific text seems to be the only way to obtain sufficiently large, diverse corpora.

CorMet obtains documents by submitting queries to the Google search engine. There are two types of queries: one to fetch any domain-specific documents and another to fetch domain-specific documents that contain a particular verb. The first kind of query consists of a conjunction of from two to five randomly selected **domain keywords**. Domain keywords are words characteristic of a domain, supplied by the user as an input. For the *FINANCE* domain, a reasonable set of keywords is *stocks, bonds, NASDAQ, Dow, investment, finance*. Each query incorporates only a few keywords in order to maximize the number of distinct possible queries.

Queries for domain-specific documents containing a particular verb are composed of a conjunction of domain-specific terms and a disjunction of forms of the verb that are more likely to be verbs than other parts of speech. For the verb *attack*, for instance, acceptable forms are *attacked* and *attacking*, but not *attack* and *attacks*, which are more likely to be nouns. The syntactic categories in which a word form appears are determined by reference to WordNet.

Some queries for the verb *attack* in the *FINANCE* domain are:

1. (attacked OR attacking) AND (bonds AND Dow AND investment)
2. (attacked OR attacking) AND (NASDAQ AND investment AND finance)
3. (attacked OR attacking) AND (stocks AND bonds AND NASDAQ)
4. (attacked OR attacking) AND (stocks AND NASDAQ AND Dow)

Queries return links to up to 10,000 documents, of which CorMet fetches and analyzes no more than 3,000. In the 13 domains studied, about 75% of these documents are relevant to the domain of interest (as measured through a randomly chosen, hand-evaluated sample of 100 documents per domain), so the noise is substantial. The documents are processed to remove embedded scripts and HTML tags.

The mined documents are parsed with the apple pie parser (Sekine and Grishman 1995). Case frames are extracted from parsed sentences using templates; for instance, *(S (NP & OBJ) (VP (were | was | got | get) (VP WORDFORM-PASSIVE))* is used to extract roles for passive, agentless sentences (where *WORDFORM-PASSIVE* is replaced by a passive form of the verb under analysis).

2.2 Finding Characteristic Predicates

Learning the selectional preferences for a verb in a domain is expensive in terms of time, so it is useful to find a small set of important verbs in each domain. CorMet seeks information about verbs typical of a domain, because these verbs are more likely to figure in metaphors in which that domain is the metaphor's source. *Besiege*, for instance, is characteristic of the *MILITARY* domain and appears in many instances of the *MILITARY* → *MEDICINE* mapping, such as *The antigens besieged the virus.*

To find domain-characteristic verbs, CorMet dynamically obtains a large sample of domain-relevant documents, decomposes them into a bag-of-words representation, stems the words with an implementation of the Porter (1980) stemmer, and finds the ratio of occurrences of each word stem to the total number of stems in the domain corpus. The frequency of each stem in the corpus is compared to its frequency in general English (as recorded in an English-language frequency dictionary [Kilgarriff 2003]).

The 400 verb stems with the highest relative frequency (computed as a ratio of the stem's frequency in the domain to its frequency in the English frequency dictionary) are considered characteristic. CorMet treats any word form that may be a verb (according to WordNet) as though it is a verb, which biases CorMet toward verbs with common nominal homonyms. Word stems that have high relative frequency in more than one

Table 1
Characteristic stems for *LAB* and *FINANCE* domains.

Rank	*LAB*	*FINANCE*
1	oxidiz	amortiz
2	sulfat	arbitrag
3	fluorin	labor
4	vapor	overvalu
5	titrat	outsourc
6	adsorb	escrow
7	electropl	repurchas
8	valenc	refinanc
9	atomiz	forecast
10	anneal	invest
11	sinter	discount
12	substitu	stock
13	compound	certify
14	hydrat	bank
15	frit	credit
16	ionize	yield
17	deactiv	bond
18	intermix	rate
19	halogen	reinvest
20	solubl	leverag

domain, like *e-mail* and *download*, are eliminated on the suspicion that they are more characteristic of documents on the Internet in general than of a substantive domain. Table 1 lists the 20 highest-scoring stems in the *LAB* and *FINANCE* domains.

2.3 Selectional Preference Learning

There are three constraints on CorMet's selectional-preference-learning algorithm. First, it must tolerate noise, because complex sentences are often misparsed, and the case frame extractor is error prone. Second, it should be able to work around WordNet's lacunae. Finally, there should be a reasonable metric for comparing the similarity between selectional preferences.

CorMet first uses the selectional-preference-learning algorithm described in Resnik (1993), then clustering over the results. Resnik's algorithm takes a set of words observed in a case slot (e.g., the subject of *pour* or the indirect object of *give*) and finds the WordNet nodes that best characterize the selectional preferences of that slot. (Note that WordNet nodes are treated as categories subcategorizing their descendants.) A case slot has a preference for a WordNet node to the extent that that node, or one of its descendants, is more likely to appear in that case slot than it is to appear at random.

An overall measure of the choosiness of a case slot is **selectional-preference strength**, $S_R(p)$, defined as the relative entropy of the posterior probability $P(c|p)$ and the prior probability $P(c)$ (where $P(c)$ is the a priori probability of the appearance of a WordNet node c, or one of its descendants, and $P(c|p)$ is the probability of that node or one of its descendants appearing in a case slot p.) Recall that the relative entropy of two distributions X and Y, $D(X||Y)$, is the inefficiency incurred by using an encoding optimal for Y to encode X.

$$S_R(p) = D(P(c|p)||P(c)) = \sum_c P(c|p) \log \frac{P(c|p)}{P(c)}$$

The degree to which a case slot selects for a particular node is measured by **selectional association**. In effect, the selectional associations divide up the selectional preference strength for a case slot among that slot's possible fillers. Selectional association is defined as

$$\Lambda_R(p,c) = \frac{1}{S_R(p)} P(c|p) \log \frac{P(c|p)}{P(c)}$$

To compute $\Lambda_R(p,c)$, what is needed is a distribution over word classes, but what is observed in the corpus is a distribution over word forms. Resnik's algorithm works around this problem by approximating a word class distribution from the word form distribution. For each word form observed filling a case slot, credit is divided evenly among all of that word form's possible senses (and their ancestors in WordNet). Although Resnik's algorithm makes no explicit attempt at sense disambiguation, greater activation tends to accumulate in those nodes that best characterize a predicate's selectional preferences.

CorMet uses Resnik's algorithm to learn domain-specific selection preferences. It often finds different selectional preferences for predicates whose preferences should, intuitively, be the same. In the *MILITARY* domain, the object of *assault* selects strongly for *fortification* but not *social group*, whereas the selectional preferences for the object of *attack* are the opposite. Taking the cosine of the selectional preferences of these two case slots (one of many possible similarity metrics) gives a surprisingly low score. In order to facilitate more accurate judgments of selectional-preference similarity, CorMet finds clusters of WordNet nodes that, although not as accurate, allow more meaningful comparisons of selectional preferences.

Clusters are built using the nearest-neighbor clustering algorithm (Jain, Murty, and Flynn 1999). A predicate's selectional preferences are represented as vectors whose nth element represents the selectional association of the nth WordNet node for that predicate. The similarity function used is the dot product of the two selectional-preference vectors. Empirically, the level of granularity obtained by running nearest-neighbor clustering twice (i.e., clustering over the sets of nodes constituting selectional preferences, then clustering over the clusters) produces the most conceptually coherent clusters.

There are typically fewer than 100 second-order clusters (i.e., clusters of clusters) per domain. In the *LAB* domain there are 54 second-order clusters, and in the *FINANCE* domain there are 67. The time complexity of searching for metaphorical interconcept mappings between two domains is proportional to the number of pairs of salient domain objects, so it is more efficient to search over pairs of salient clusters than over the more numerous individual salient nodes.

Table 2 shows a *MILITARY* cluster. These clusters are helpful for finding verbs with similar, but not identical, selectional preferences. Although *attack*, for instance, does not select for *fortification*, it does select for other elements of *fortification*'s cluster, such as *building* and *defensive structure*.

The fundamental limitation of WordNet with respect to selectional-preference learning is that it fails to exhaust all possible lexical relationships. WordNet can hardly be blamed: The task of recording all possible relationships between all English words is prohibitively large, if not infinite. Nevertheless, there are many words that intuitively should have a common parent but do not. For instance, *liquid body substance* and *water* should both be hyponyms of *liquid*, but in WordNet their shallowest common ancestor is *substance*. One of the descendants of *substance* is *solid*, so there is no single node that represents all liquids.

Li and Abe (1998) describe another method of corpus-driven selectional-preference learning that finds a tree cut of WordNet for each case slot. A tree cut is a set of

Table 2
The elements of a cluster of WordNet nodes characteristic of the *MILITARY* domain.

penal institution-1	*fortification-1*
correctional institution-1	*defensive structure-1*
institution-2	*housing-1*
structure-1	*room-1*
establishment-4	*prison-1*
building-1	*tower-1*
area-3	

nodes that specifies a partition of the ontology's leaf nodes, where a node stands for all the leaf nodes descended from it. The method chooses among possible tree cuts according to minimum-description-length criteria. The description length of a tree cut representation is the sum of the size of the tree cut itself (i.e., the minimum number of nodes specifying the partition) and the space required for representing the observed data with that tree cut. For CorMet's purposes, the problem with this approach is that it is difficult to find clusters of (possibly hypernymically related) nodes representing a selectional preference using its results (because the tree cut includes exactly one node on each path from each leaf node to the root). There are similar objections to similar approaches such as that of Carroll and McCarthy (2000).

2.4 Polarity

Polarity is a measure of the directionality and magnitude of structure transfer between two concepts or two domains. Nonzero polarity exists when language characteristic of a concept from one domain is used in a different domain of a different concept. The kind of characteristic language CorMet can detect is limited to verbal selectional preferences.

Say CorMet is searching for a mapping between the concepts *liquids* (characteristic of the *LAB* domain) and *assets* (characteristic of the *FINANCE* domain), as illustrated in Figure 1. There are verbs in *LAB* that strongly select for *liquids*, such as *pour*, *flow*, and *freeze*. In *FINANCE*, these verbs select for *assets*. In *FINANCE* there are verbs that strongly select for *assets* such as *spend*, *invest*, and *tax*. In the *LAB* domain, these verbs select for nothing in particular. This suggests that *liquid* is the source concept and *asset* is the target concept, which implies that *LAB* and *FINANCE* are the source and target domains, respectively. CorMet computes the overall polarity between two domains (as opposed to between two concepts) by summing over the polarity between each pair of high-salience concepts from the two domains of interest.

Interconcept polarity is defined as follows: Let α be the set of case slots in domain X with the strongest selectional preference for the node cluster A. Let β be the set of case slots in domain Y with the strongest selectional preferences for the node cluster B. The degree of structure flow from A in X to B in Y is computed as the degree to which the predicates α select for the nodes B in Y, or $\mathit{selection\_strength}(Y, \alpha, B)$. Structure flow in the opposite direction is $\mathit{selection\_strength}(X, \beta, A)$. The definition of $\mathit{selection\_strength}(\mathit{Domain}, \mathit{case\_slots}, \mathit{node\_cluster})$ is the average of the selectional-preference strengths of the predicates in *case_slots* for the nodes in *node_cluster* in *Domain*. The polarity for α and β is the difference in the two quantities. If the polarity is near zero, there is not much structure flow and no evidence for a metaphoric mapping.

In some cases a difference in selectional preferences between domains does not indicate the presence of a metaphor. To take a fictitious but illustrative example, say

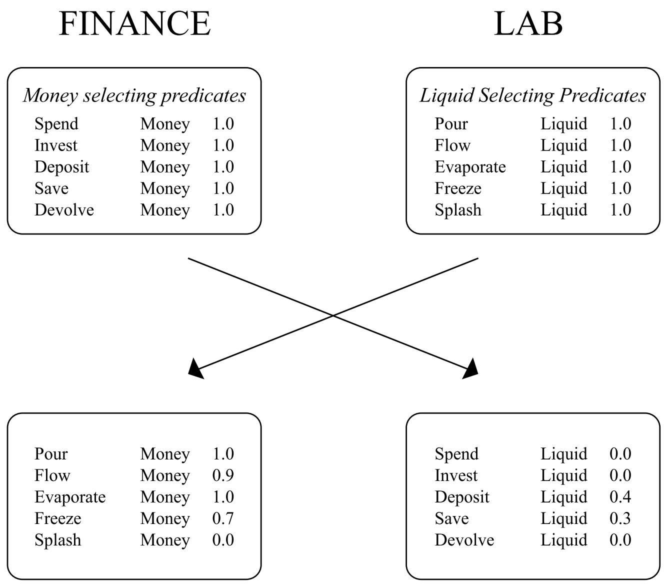

Figure 1
Asymmetric structure transfer between *LAB* and *FINANCE*. Predicates from *LAB* that select for liquids are transferred to *FINANCE* and select for money. On the other hand, predicates from *FINANCE* that select for money are transferred to *LAB* and do not select for liquids.

that in the *LAB* domain the subject of *sit* has a preference for chemists whereas in the *FINANCE* domain it has a preference for investment bankers. The difference in selectional preferences is caused by the fact that chemists are the kind of person more likely to appear in *LAB* documents and investment bankers in *FINANCE* ones. Instances like this are easy to filter out because their polarity is zero.

A verb is treated as characteristic of a domain X if it is at least twice as frequent in the domain corpus as it is in general English and it is at least one and a half times as frequent in domain X as in the contrasting domain Y (these ratios were chosen empirically). *Pour*, for instance, occurs three times as often in *FINANCE* and twenty-three times as often in *LAB* as it does in general English. Since it is nearly eight times as frequent in *LAB* as in *FINANCE*, it is considered characteristic of the former. This heuristic resolves the confusion than can be caused by the ubiquity of certain conventional metaphors—the high density of metaphorical uses of *pour* in *FINANCE* could otherwise make it seem as though *pour* is characteristic of that domain.

A verb with weak selectional preferences (e.g., *exist*) is a bad choice for a characteristic predicate even if it occurs disproportionately often in a domain. Highly selective verbs are more useful because violations of their selectional preferences are more informative. For this reason a predicate's salience to a domain is defined as its selectional-preference strength times the ratio of its frequency in the domain to its frequency in English.

Literal and metaphorical selectional preferences may coexist in the same domain. Consider the selectional preferences of *pour* in the chemical and financial domains. In the *LAB* domain, *pour* is mostly used literally: People pour liquids. There are occasional metaphorical uses (e.g., *Funding is pouring into basic proteomics research*), but the literal sense is more common. In *FINANCE*, *pour* is mostly used metaphorically, although there are occasionally literal uses (e.g., *Today oil poured into the new Turkmenistan pipeline*).

Algorithms 1–3 show pseudocode for finding metaphoric mappings between concepts.

Algorithm 1: FIND_INTER_CONCEPT_MAPPINGS(*domain1, domain2*)

comment: Find mappings from concepts in *domain*1 to concepts in *domain*2 or vice versa

Domain_1_Clusters ← GET_BEST_CLUSTERS(*domain1*)

Domain_2_Clusters ← GET_BEST_CLUSTERS(*domain2*)

for each *Concept_1* ∈ *Domain_1_Clusters*

do {
 for each *Concept_2* ∈ *Domain_2_Clusters*
 do {
  Polarity_Score ← DETECT_INTER_CONCEPT_MAPPING(*Concept_*1, *Concept_*2, *domain*1, *domain*2)
  if *Polarity_score* > *NOISE_THRESHOLD*
   then OUTPUT_MAPPING(*Concept_*1 → *Concept_*2)
  if *Polarity_score* < −*NOISE_THRESHOLD*
   then OUTPUT_MAPPING(*Concept_*2 → *Concept_*1)

Algorithm 2: DETECT_INTER_CONCEPT_MAPPING(*Concept_1, Concept_2, domain1, domain2*)

*polarity_from_*1*_to_*2 ← INTER_CONCEPT_POLARITY(*Concept_*1, *Concept_*2, *domain*1, *domain*2)

*polarity_from_*2*_to_*1 ← INTER_CONCEPT_POLARITY(*Concept_*2, *Concept_*1, *domain*2, *domain*1)

if ABSOLUTE_VALUE(*polarity_from_*1*_to_*2 − *polarity_from_*2*_to_*1) < *C*1

 then return (0);

if *polarity_from_*1*_to_*2 > *C*2 **and** *polarity_from_*2*_to_*1 > *C*2

 then return (0);

return (*polarity_from_*1*_to_*2 − *polarity_from_*2*_to_*1)

Algorithm 3: INTER_CONCEPT_POLARITY(*Concept_A, Concept_B, domain_A, domain_B*)

polarity ← 0

domain_A_predicates ← GET_PREDICATES_SELECTING_FOR_CONCEPT(*Concept_A*, *domain_A*)

for each *Predicate_from_A* ∈ *domain_A_predicates*

 do {*polarity* ← *polarity* + SELECTION_STRENGTH(*Predicate_from_A*, *Concept_B*, *domain_B*)

return (*polarity*)

2.5 Systematicity

According to the thematic-relation hypothesis (Grubner 1976), many domains are conceived of in terms of physical objects moving along paths between locations in space. In the money domain, assets are mapped to objects and asset holders are mapped to locations. In the idea domain, ideas are mapped to objects, minds are mapped to locations, and communications are mapped to paths. Axioms of inference from the target domain usually become available for reasoning about the source domain, unless there is an aspect of the source domain that specifically contradicts them. For instance, in the domain of material objects, a thing moved from point X to point Y is no longer at X, but in the idea domain, it exists at both locations.

Thematically related metaphors may consistently co-occur in the same sentences. For example, the metaphors *LIQUID → MONEY* and *CONTAINERS → INSTITUTIONS* often co-occur, as in the sentence *Capital flowed into the new company*. Conversely, co-occurring metaphors are often components of a single metaphorical conceptualization. A metaphorical mapping is therefore more credible when it is a component of a system of mappings.

In CorMet, systematicity measures a metaphorical mapping's tendency to co-occur with other mappings. The systematicity score for a mapping X is defined as the number of strong, distinct mappings co-occurring with X. This measure goes only a little way toward capturing the extent to which a metaphor exhibits the structure described in the thematic-relations hypothesis, but extending CorMet to find the entities that correspond to objects, locations, and paths is beyond the scope of this article.

2.6 Confidence Rating

CorMet computes a confidence measure for each metaphor it discovers. Confidence is a function of three things. The more verbs mediating a metaphor (as *attack* and *assault* mediate *ENEMY → DISEASE* in *The antigen attacked the virus* and *Chemotherapy assaults the tumor*), the more credible it is. Strongly unidirectional structure flow from source domain to target makes a mapping more credible. Finally, a mapping is more likely to be correct if it systematically co-occurs with other mappings. The confidence measure should not be interpreted as a probability of correctness: The data available for calibrating such a distribution are inadequate. The weights of each factor, empirically assigned plausible values, are given in Table 3.

The confidence measure is intended to wrap all the available evidence about a metaphor's credibility into one number. A principled way of doing this is desirable, but unfortunately there are not enough data to make meaningful use of machine-learning techniques to find the best set of components and weights. There is substantial arbitrariness in the confidence rating: The components used and the weights they are

Table 3
Factors used in evaluating a mapping M and their weights.

Component	Weight
$\frac{\lvert supporting\_predicates(M)\rvert}{max\_num\_of\_support\_preds\_in\_domain}$	0.25
$\frac{polarity(M)}{max\_polarity\_in\_domain}$	0.5
$\frac{\lvert co\_occurring\_mappings(M)\rvert}{max\_number\_of\_cooccurring\_mappings}$	0.25

Table 4
Characteristic keywords for *LAB* and *FINANCE* domains.

LAB	beaker experiment cylinder chemical precipitate mixture reaction valence molarity pressure
FINANCE	money stocks bonds equity trading inflation arbitrage capital investment market

assigned could easily be different and are best considered guesses that give reasonable results.

3. Two Examples

This section provides a walk-through of the derivation and analysis of the concept mapping *LIQUID* → *MONEY* and components of the interconcept mapping *WAR* → *MEDICINE*. In the interests of brevity only representative samples of CorMet's data are shown. See Mason (2002) for a more detailed account.

3.1 *LIQUID* → *MONEY*

CorMet's inputs are two domain sets of characteristic keywords for each domain (Table 4). The keywords must characterize a cluster in the space of Internet documents, but CorMet is relatively insensitive to the particular keywords.

It is difficult to find keywords characterizing a cluster centering on money alone, so keywords for a more general domain, *FINANCE*, are provided. It is also difficult to characterize a cluster of documents mostly about liquids. Chemical-engineering articles and hydrographic encyclopedias tend to pertain to the highly technical aspects of liquids instead of their everyday behavior. Documents related to laboratory work are targeted on the theory that most references to liquids in a corpus dedicated to the manipulation and transformation of different states of matter are likely to be literal and will not necessarily be highly technical. Tables 5 and 6 show the top 20 characteristic verbs for *LAB* and *FINANCE*, respectively.

CorMet finds the selectional preferences of all of the characteristic predicates' case slots. A sample of the selectional preferences of the top 20 verbs in *LAB* and *FINANCE* are shown in Tables 7 and 8, respectively. The leftmost columns of these two tables have the (stemmed form of the) characteristic verb and the thematic role characterized. The right-hand sides have clusters of characteristic nodes. The numbers associated with the nodes are the bits of uncertainty about the identity of a word x resolved by the fact that x fills the given case slot, or $P(x \leftarrow N) - P(x \leftarrow N|case\_slot(x))$ (where $x \leftarrow N$ is read as x is N or a hyponym of N).

All of the 400 possible mappings between the top 20 concepts (clusters) from the two domains are examined. Each possible mapping is evaluated in terms of polarity, the number of frames instantiating the mapping, and the systematic co-occurrence of that mapping with different, highly salient mappings. The best mappings for *LAB* × *FINANCE* are shown in Table 9.

Mappings are expressed in abbreviated form for clarity, with only the most recognizable (if not necessarily the most salient) node of each concept displayed. The foremost mapping characterizes money in terms of liquid, the mapping for which the two domains were selected. The second represents a somewhat less intuitive mapping from liquids to institutions. This metaphor is driven primarily by institutions' capacity

Table 5
Characteristic verbs 1–20 of the *LAB* domain.

Rank	Stem	Ratio of frequencies	Frequency in domain	Frequency in English
1	oxidiz	3,073.608	0.0003	1.0e−07
2	sulfat	2,301.591	0.0003	1.3e−07
3	fluorin	1,452.467	0.0001	1.0e−07
4	vapor	1,325.237	0.0007	5.2e−07
5	titrat	831.007	0.0006	8.3e−07
6	adsorb	433.721	5.6e−05	1.2e−07
7	electropl	392.986	3.1e−05	7.9e−08
8	valenc	349.522	0.0004	1.4e−06
9	atomiz	324.696	1.9e−05	5.9e−08
10	anneal	312.406	8.1e−05	2.5e−07
11	sinter	264.322	3.6e−05	1.3e−07
12	substitu	251.511	3.7e−05	1.4e−07
13	compound	99.632	0.002	2.0e−05
14	hydrat	238.017	0.0001	6.5e−07
15	frit	237.08	1.6e−05	6.9e−08
16	ionize	221.372	9.2e−05	4.1e−07
17	deactiv	207.629	1.4e−05	6.9e−08
18	intermix	84.18	5.0e−06	5.9e−08
19	halogen	195.701	0.0001	6.9e−07
20	solubl	192.204	0.0007	4.1e−06

Table 6
Characteristic verbs 1–20 of the *FINANCE* domain.

Rank	Stem	Ratio of frequencies	Freqency in domain	Frequency in English
1	amortiz	807.531	5.6e−05	6.9e−08
2	arbitrag	305.836	0.0006	2.0e−06
3	labor	302.797	0.0004	1.6e-06
4	overvalu	296.945	4.7e−05	1.5e−07
5	outsourc	260.625	2.8e−05	1.0e−07
6	escrow	248.192	2.9e−05	1.1e−07
7	repurchas	241.309	9.4e−05	3.8e−07
8	refinanc	213.369	3.4e−05	1.5e−07
9	forecast	27.007	0.0004	1.4e−05
10	invest	72.604	0.0019	2.7e−05
11	discount	22.59	0.0005	2.2e−05
12	stock	70.172	0.0067	9.5e−05
13	certify	21.08	5.7e-05	2.7e−06
14	bank	20.624	0.0045	0.0002
15	credit	20.432	0.0016	7.9e−05
16	yield	56.144	0.001	1.8e−05
17	bond	122.467	0.0045	3.7e−05
18	rate	17.563	0.0055	0.0003
19	reinvest	104.197	0.0001	1.1e−06
20	leverag	100.576	0.0002	2.2e−06

Table 7
Sample selectional preferences for *LAB* verbs.

vapor	obj	substance-1	0.0116
		liquid-1	0.0478
		fluid-1	0.0473
anneal	with	metallic element-1	0.0217
		substance-1	0.0101
		chemical element-1	0.0112
compound	subj	substance-1	0.0123
		compound-2	0.036
		organic compound-1	0.0431
adsorb	obj	matter-3	0.0145
		substance-1	0.014
		physical object-1	0.0087
hydrat	subj	substance-1	0.0181
		compound-2	0.0401

Table 8
Sample selectional preferences for *FINANCE* verbs.

invest	obj	income-1	0.0118
		financial gain-1	0.0114
		security-8	0.0069
		currency-1	0.034
		sum-1	0.0136
		transferred property-1	0.0036
		fund-1	0.008
		asset-1	0.1183
		gain-4	0.0113
		medium of exchange-1	0.0415
		money-1	0.0375
discount	obj	cost-1	0.0269
		financial loss-1	0.0263
		transferred property-1	0.0237
		loss-2	0.0262
		outgo-1	0.0269
credit	subj	cost-1	0.0211
		financial loss-1	0.0206
		transferred property-1	0.0182
		loss-2	0.0205
		outgo-1	0.0211

Table 9
Mappings *LAB* → *FINANCE*.

Mapping	Frames	Polarity	Systematicity	Final score
liquid-1 → income-1	61	11.8	2	.56
liquid-1 → institution-1	59	3.83	2	.55
container-1 → institution-1	11	3.16	1	.35
liquid-1 → information-1	56	4.29	2	.54

to dissolve. Of course, this mapping is incorrect insofar as solids undergo dissolution, not liquids. CorMet made this mistake because of faulty thematic-role identification; it frequently failed to distinguish between the different thematic roles played by the subjects in sentences like *The company dissolved* and *The acid dissolved the compound.* The third mapping characterizes communication as a liquid. This was not the mapping the author had in mind when he chose the domains, but it is intuitively plausible: One speaks of information flowing as readily as of money flowing. That this mapping appears in a search not targeted to it reflects this metaphor's strength. It also illustrates a source of error in inferring the existence of conventional metaphors between domains from the existence of interconcept mappings. The fourth mapping is from containers to organizations. This mapping complements the first one: As liquids flow into containers, so money flows into organizations. Another good mapping, not present here, is *money flows into equities and investments*. CorMet misses this mapping because, at the level of concepts, money and equities are conflated. This happens because they are near relatives in the WordNet ontology and because there is very high overlap between the predicates selecting for them.

Compare the mappings CorMet derived with the Master Metaphor List's (Lakoff, Espenson, and Schwartz 1991) characterization of the *MONEY IS A LIQUID* metaphor:

1. Cash is a Liquid.
   (a) liquid assets
   (b) currency
   (c) liquidating assets
   (d) My money is all dried up
   (e) He's just sponging off you (absorbing cash)
   (f) He's solvent/insolvent
2. Gain/Loss is Movement of a Liquid.
   (a) cash flow
   (b) influx and outflux of money
   (c) Don't pour your money down the drain
3. Money Which Cannot be Accessed is Frozen
   (a) frozen assets
   (b) price freeze
4. Control in Financial Situation is Control in Liquid
   (a) keep your head above water, financially
   (b) get in over your head
   (c) stay afloat
   (d) the business went under/sunk
   (e) drowning in debts

The Master Metaphor List also describes *INVESTMENTS ARE CONTAINERS FOR MONEY*, as exemplified in the following:

1. Put your money in bonds.
2. The bottom of the economy dropped out.

Table 10
A sample of frames from *FINANCE* instantiating *liquid* → *income.*

vb	subj	obj	into	from	with
dissolv		stakes			
pour	investors	cash			
pour	investors	cash			
pour	profits		market		
pour	Cash		shares		
pour	Earnings				
pour	stake	brand			
pour	cash				
pour	flight	money	stocks		
pour	investors		stocks		
cool	stocks				
cool	Reserve	economy			
evapor	profit				
evapor	mortgages				
evapor	profit	turn			
pump	Reserve	reserves			
pump	stocks		them		
vapor	stock				
vapor	profits				
melt	profit		nothing		
melt	stocks				

3. I'm down to my bottom dollar.
4. This is an airtight investment.

CorMet has found mappings that can reasonably be construed as corresponding to these metaphors. Compare the mappings from the Master Metaphor List with frames mined by this system and identified as instantiating *liquid* → *income*, shown in Table 10. It is important to note that although CorMet can list the case frames that have driven the derivation of a particular high-level mapping, it is designed to discover high-level mappings, not interpret or even recognize particular instances of metaphorical language. Just as in the Master Metaphor List, there are frames in the CorMet listing in which money and equities are characterized as liquids, are moved as liquids (i.e., pouring earnings and pumping reserves) and change state as liquids (i.e., melting stocks, dissolving stakes, evaporating profits, frozen money).

3.2 *MILITARY* → *MEDICINE*

This subsection describes the search for mappings between the *MEDICINE* and *MILITARY* domains. The domain keywords for *MEDICINE* and *MILITARY* are shown in Table 11. The characteristic verbs of the *MILITARY* and *MEDICINE* domains are given in Tables 12 and 13, respectively. Their selectional preferences are given in Tables 14 and 15, respectively.

The highest-quality mappings between the *MILITARY* and *MEDICINE* domains are shown in Table 16. This pair of domains produces more mappings than the the *LAB* and *FINANCE* pair. Many source concepts from the *MILITARY* domain are mapped to body parts. The heterogeneity of the source concepts seems to be driven by the heterogeneity of possible military targets. Similarly, many source concepts are mapped to drugs. The case frames supporting this mapping suggest that this is because of

Table 11
Characteristic keywords for the *MEDICINE* and *MILITARY* domains.

MEDICINE	doctor surgeon hospital operate pharmaceutical medicine recuperate organ tissue bacteria virus diagnose cancer sickness nurse research
MILITARY	army navy soldier battle war attack bombing destruction infantry tactics siege invasion troops barracks

Table 12
Characteristic verbs for *MILITARY.*

Rank	Stem	Ratio of frequencies	Frequency in domain	Frequency in English
0	nuke	372.494	2.2e−05	5.9e−08
1	harbor	714.253	0.0004	6.9e−07
2	strafe	156.471	5.1e−05	3.2e−07
3	honor	626.577	0.0003	4.7e−07
4	combat	105.121	0.001	9.6e−06
5	torpedo	96.519	0.0002	2.0e−06
6	stonewal	382.93	3.0e−05	7.9e−08
7	bombard	54.602	0.0002	5.1e−06
8	skirmish	56.105	0.0001	2.4e−06
9	bomb	49.341	0.0019	3.9e−05
10	favor	169.023	0.0001	6.5e−07
11	envision	158.417	1.4e−05	8.9e−08
12	attack	31.661	0.0034	0.0001
13	cannonad	117.742	1.0e−05	8.9e−08
14	rearm	115.601	1.2e−05	1.0e−07
15	sieg	107.732	0.0008	7.8e−06
16	raid	20.817	0.0004	2.1e−05
17	highlight	77.358	0.0014	1.9e−05
18	enlist	74.138	0.0002	3.4e−06
19	infest	17.725	1.3e−05	7.4e−07

the heterogeneity of military aggressors (fortifications do not generally fall into this category; this mapping is an error caused by the frame extractor's frequent confusion of subject and object). These mappings can be interpreted as indicating that things that are attacked map to body parts and things that attack map to drugs.

The mapping *fortification* → *illness* represents the mapping of targetable strongholds to disease. Illnesses are conceived of as fortifications besieged by treatment.

Compare this with the Master Metaphor List's characterization of *TREATING ILLNESS IS FIGHTING A WAR*:

1. The Disease is an Enemy.
2. The Body is a Battleground.
   (a) The body is not immune to invasion.
   (b) The disease infiltrates your body and takes over.
3. Infection is an Attack by the Disease.
   (a) His body was under siege by AIDS.
   (b) He was attacked by an unknown virus.
   (c) The virus began an attack on the organ systems.

Table 13
Characteristic verbs for *MEDICINE.*

Rank	Stem	Ratio of frequencies	Frequency in domain	Frequency in English
1	immuniz	304.704	0.0001	4.0e−07
2	diaper	110.023	2.6e−05	2.3e−07
3	detoxify	106.181	2.0e−05	1.8e−07
4	oxidiz	104.006	1.1e−05	1.0e−07
5	pasteur	102.149	3.5e−05	3.4e−07
6	palpat	89.38	1.4e−05	1.5e−07
7	misdiagnos	87.394	7.8e−06	8.9e−08
8	metastas	87.049	4.0e−05	4.5e−07
9	expector	86.826	8.6e−06	9.9e−08
10	implant	85.263	0.0001	2.3e−06
11	decoct	82.996	6.6e−06	7.9e−08
12	vaccin	81.157	0.0007	8.8e−06
13	transplant	78.7	0.0005	7.1e−06
14	labor	77.016	0.0001	1.6e−06
15	infect	69.575	0.0003	5.4e−06
16	deactiv	67.126	4.6e−06	6.9e−08
17	detox	63.417	7.6e−06	1.1e−07
18	recuper	62.588	7.3e−05	1.1e−06
19	heal	61.753	0.0005	8.9e−06
20	clot	58.416	7.4e−05	1.2e−06

Table 14
Selectional preferences for *MILITARY* verbs.

combat	subj	social group-1	0.005
		body-3	0.0123
		gathering-1	0.0053
combat	obj	military unit-1	0.0156
		social group-1	0.01
enlist	subj	unit-3	0.0135
		military unit-1	0.0603
		social group-1	0.0475
muster	subj	military unit-1	0.0164
		social group-1	0.0131
		military unit-1	0.0368
		military unit-1	0.0397
		company-6	0.0101
		gathering-1	0.0013
		unit-3	0.0196
		social gathering-1	0.0049
		force-4	0.0052
bomb	subj	district-1	0.0046
		seat-5	0.0046
		region-3	0.0022
		administrative district-1	0.0051
		country-1	0.0021
		capital-3	0.0046
		city-2	0.0073
		national capital-1	0.0056

Table 15
Selectional preferences for *MEDICINE* verbs.

immuniz	subj	descendant-1	0.0246
		child-2	0.0198
		relative-1	0.0137
		offspring-1	0.0193
		child-4	0.0246
oxidiz	subj	food-1	0.0513
		substance-1	0.0158
implant	subj	organ-1	0.0204
		gland-1	0.0303
		body part-1	0.0238
		tissue-1	0.0151
		part-7	0.0225

Table 16
Mappings *MILITARY* → *MEDICINE.*

Mapping	Frames	Polarity	Systematicity	Final score
military unit-1 → body part-1	285	65.55	33	0.95
fortification-1 → body part-1	298	55.12	33	0.88
vehicle-1 → body part-1	238	35.2	32	0.67
military action-1 → body part-1	207	35	25	0.6
region-3 → body part-1	57	30.9	5	0.31
skilled worker-1 → body part-1	127	17.3	11	0.31
military unit-1 → drug-1	84	51.77	28	0.64
vehicle-1 → drug-1	63	35.7	28	0.5
military action-1 → drug-1	71	30.91	27	0.47
fortification-1 → drug-1	67	24.64	22	0.38
weaponry-1 → drug-1	58	10.8	24	0.28
military action-1 → medical care-1	71	28.21	20	0.4
fortification-1 → medical care-1	78	16.37	20	0.32
weaponry-1 → medical care-1	48	9.64	20	0.24
fortification-1 → illness-1 243	.21	38	.45	

4. Medicine is a Weapon.
   (a) The so-called cure is no magic bullet.
5. Medical Procedures are Attacks by the Patient.
   (a) The doctors tried to wipe out the infection.
6. The Immune System is a Defense.
   (a) The body normally has its own defenses.
7. Winning the War is being Cured of the Disease.
   (a) Beating measles takes patience.
8. Being Defeated is Dying.
   (a) The patient finally gave up the battle.

Table 17
Selected frames supporting {*fortification, vehicle, military action, region, skilled worker*} → *body part*.

vb	subj	obj	into	from	with
attack	system	receptors			
attack	pain	joints			
attack	immunosuppressants	kidney			
besieg	flood	abdomen			
besieg	scars	thigh			
destroy	organs	bacteria			
destroy	Microtubules	agents			
destroy		ganglion			
destroy	therapy	tissue			
destroy	cancer	bone			
destroy	virus	liver			
destroy	internist	stomach			
target		organ			
target	vaccine	intestines			

CorMet's results can reasonably be interpreted as matching all of the mappings from the Master Metaphor List except winning-is-a-cure and defeat-is-dying. CorMet's failure to find this mapping is caused by the fact that *win*, *lose*, and their synonyms do not have high salience in the *MILITARY* domain, which may be a reflection of the ubiquity of *win* and *lose* outside of that domain.

Table 17 shows sample frames from which the *body part* → {*fortification, vehicle, military action, region, skilled worker*} mapping was derived.

4. Testing against the Master Metaphor List

This section describes the evaluation of CorMet against a gold standard, specifically, by determining how many of the metaphors in a subset of the Master Metaphor List (Lakoff, Espenson, and Schwartz 1991) can be discovered by CorMet given a characterization of the relevant source and target domains. The final evaluation of the correspondence between the mappings CorMet discovers and the Master Metaphor List entry is necessarily done by hand. This is a highly subjective method of evaluation; a formal, objective evaluation of correctness would be preferable, but at present no such metric is available.

The Master Metaphor List is the basis for evaluation because it is composed of manually verified metaphors common in English. The test set is restricted to those elements of the Master Metaphor List with concrete source and target domains. This requirement excludes many important conventional metaphors, such as *EVENTS ARE ACTIONS*. About a fifth of the Master Metaphor List meets this constraint. This fraction is surprisingly small: It turns out that the bulk of the Master Metaphor List consists of subtle refinements of a few highly abstract metaphors. The concept pairs and corresponding domain pairs for the target metaphors in the Master Metaphor List are given in Table 18.

A mapping discovered by CorMet is considered correct if submappings specified in the Master Metaphor List are nearly all present with high salience and incorrect submappings are present with comparatively low salience. The mappings discovered that best represent the targeted metaphors are shown in Table 19.

Some of these test cases are marked successes. For instance, *ECONOMIC HARM IS PHYSICAL INJURY* seems to be captured by the mapping from the *loss-3* cluster to

Table 18
Master Metaphor List mappings and the domain pairs in which they are sought.

Master Metaphor List mapping	Domains
Theories are Fortifications	Theory & Architecture
Emotion is a Fluid	Emotion & Lab
People are Containers for Emotions	Emotion & Lab
Love is War	Love & Military
Effects of Humor are Injuries	Humor & Military
Treating Illness is Fighting a War	Medicine & Military
Love is a Journey	Love & Journey
Economic Harm is Physical Injury	Finance & Medicine
Machines are People	Mechanical & Body
Money is a Liquid	Finance & Lab
Investments are Containers for Money	Finance & Lab
Bodies are Buildings	Body & Architecture
Society is a Body	Society & Body

Table 19
Best mappings for domain pairs.

Master Metaphor List mapping	Empirical mapping	Score
Fortifications → Theories	none	0
Fluid → Emotion	liquid-1 → feeling-1	.25
Containers for Emotions → People	container-1 → person-1	.13
War → Love	feeling-1 → military unit-1	.34
Injuries → Effects of Humor	weapon-1 → joke-1	.18
Fighting a War → Treating Illness	military action-1 → medical care-1	.4
Journey → Love	travel-1 → feeling-1	.17
Physical Injury → Economic Harm	harm-1 → loss-3	.20
Machines → People	none	0
Liquid → Money	liquid-1 → income-1	.56
Containers for Money → Investments	container-1 → institution-1	.35
Buildings → Bodies	none	0
Body → Society	body part-1 → organization-1	.14

the *harm-1* cluster. CorMet found reasonable mappings in 10 of 13 cases attempted. This implies 77% accuracy, although in light of the small test and the subjectivity of judgment, this number must not be taken too seriously.

Some test cases were disappointing. CorMet found no mapping between *THEORY* and *ARCHITECTURE*. This seems to be an artifact of the low-quality corpora obtained for these domains. The documents intended to be relevant to architecture were often about zoning or building policy, not the structure of buildings. For theory, many documents were calls for papers or about university department policy. It is unsurprising that there are no particular mappings between two sets of miscellaneous administrative and policy documents. The weakness of the *ARCHITECTURE* corpus also prevented CorMet from discovering any *BODY* → *ARCHITECTURE* mappings. Accuracy could be improved by refining the process by which domain-specific corpora are obtained to eliminate administrative documents or by requiring documents to have a higher density of domain-relevant terms.

Is it meaningful when CorMet finds a mapping, or will it find a mapping between any pair of domains? To answer this question, CorMet was made to search for

Table 20
Arbitrarily selected domains and the mapping strengths between them.

Domain 1	Domain2	Polarity
Medicine	Plants	0
Military	Society	0
Medicine	Society	0
Finance	Body	0
Lab	Theory	0
Society	Journey	0

mappings between randomly selected pairs of domains. Table 20 lists a set of arbitrarily selected domain pairs and the strength of the polarization between them. In all cases, the polarization is zero. This can be interpreted as an encouraging lack of false positives. Another perspective is that CorMet should have found mappings between some of these pairs, such as *MEDICINE* and *SOCIETY*, on the theory that societies can be said to sicken, die, or heal. Although this is certainly a valid conventional metaphor, it seems to be less prominent than those metaphors that CorMet did discover.

5. Related Work

Two of the most broadly effective computational models of metaphor are Fass (1991) and Martin (1990), in both of which metaphors are detected through selectional-preference violations and interpreted using an ontology. They are distinguished from CorMet in that they work on both novel and conventional metaphors and rely on declarative hand-coded knowledge bases.

Fass (1991) describes Met*, a system for interpreting nonliteral language that builds on Wilks (1975) and Wilks (1978). Met* discriminates among metonymic, metaphorical, literal, and anomalous language. It is a component of collative semantics, a semantics for natural language processing that has been implemented in the program meta5 (Fass, 1986, 1987, 1988). Met* treats metonymy as a way of referring to one thing by means of another and metaphor as a way of revealing an interesting relationship between two entities.

In Met*, a verb's selectional preferences are represented as a vector of types. The verb *drink*'s preference for an animal subject and a liquid object are represented as $(animal, drink, liquid)$. Metaphorical interpretations are made by finding a sense vector in Met*'s knowledge base whose elements are hypernyms of both the preferred argument types and the actual arguments. For example, *the car drinks gasoline* maps to the vector $(car, drink, gasoline)$. But *car* is not a hypernym of *animal*, so Met* searches for a metaphorical interpretation, coming up with $(thing, use, energy\_source)$.

Martin (1990) describes the Metaphor Interpretation, Denotation, and Acquisition System (MIDAS), a computational model of metaphor interpretation. MIDAS has been integrated with the Unix Consultant (UC), a program that answers English questions about using Unix. UC tries to find a literal answer to each question with which it is presented. If violations of literal selectional preference make this impossible, UC calls on MIDAS to search its hierarchical library of conventional metaphors for one that explains the anomaly. If no such metaphor is found, MIDAS tries to generalize a known conventional metaphor by abstracting its components to the most-specific senses that encompass the question's anomalous language. MIDAS then records the most concrete metaphor descended from the new, general metaphor that provides an explanation for the query's language.

MIDAS is driven by the idea that novel metaphors are derived from known, existing ones. The hierarchical structure of conventional metaphor is a regularity not captured by other computational approaches. Although MIDAS can quickly understand novel metaphors that are the descendants of metaphors in its memory, it cannot interpret compound metaphors or detect intermetaphor relationships besides inheritance. *INVESTMENTS → CONTAINERS* and *MONEY → WATER*, for instance, are clearly related, but not in a way that MIDAS can represent. Since not all novel metaphors are descendants of common conventional metaphors, MIDAS's coverage is limited.

MetaBank (Martin 1994) is an empirically derived knowledge base of conventional metaphors designed for use in natural language applications. MetaBank starts with a knowledge base of metaphors based on the Master Metaphor List. MetaBank can search a corpus for one metaphor or scan a large corpus for any metaphorical content. The search for a target metaphor is accomplished by choosing a set of probe words associated with that metaphor and finding sentences with those words, which are then manually sorted as literal, examples of the target metaphor, examples of a different metaphor, unsystematic homonyms, or something else. MetaBank compiles statistics on the frequency of conventional metaphors and the usefulness of the probe words. MetaBank has been used to study container metaphors in a corpus of UNIX-related e-mail and to study metaphor distributions in the *Wall Street Journal*.

Peters and Peters (2000) mine WordNet for patterns of systematic polysemy by finding pairs of WordNet nodes at a relatively high level in the ontology (but still below the root nodes) whose descendants share a set of common word forms. The nodes *publication* and *publisher*, for instance, have *paper*, *newspaper*, and *magazine* as common descendants. This is a metonymic relationship; the system can also capture metaphoric relationships, as in the nodes *supporting structure* and *theory*, among whose common descendants are (for example) *framework*, *foundation*, and *base*. Peters and Peters' system found many metaphoric relationships between node pairs that were descendants of the unique beginners *artifact* and *cognition.*

Goatly (1997) describes a set of linguistic cues of metaphoricality beyond selectional-preference violations, such as *metaphorically speaking* and, surprisingly, *literally*. These cues are generally ambiguous (except for *metaphorically speaking*) but could usefully be incorporated into computational approaches to metaphor.

6. Conclusion

CorMet embodies a method for semiautomatically finding metaphoric mappings between concepts, which can then be used to infer conventionally metaphoric relationships between domains. It can sometimes identify metaphoric language, if it manifests as a common selectional-preference gradient between domains, but is far from being able to recognize metaphoric language in general. CorMet differs from other computational approaches to metaphor in requiring no manually compiled knowledge base besides WordNet. It has successfully found some of the conventional metaphors on the Master Metaphor List.

CorMet uses gradients in selectional preferences learned from dynamically mined, domain-specific corpora to identify metaphoric mappings between concepts. It is reasonably accurate despite the noisiness of many of its components. CorMet demonstrates the viability of a computational, corpus-based approach to conventional metaphor but requires more work before it can constitute a viable NLP tool.

References

Carroll, J., and D. McCarthy. 2000. Word sense disambiguation using automatically acquired verbal preferences. *Computers and the Humanities*, 34(1–2).

Cho, See-Young. 1993. Metaphor and cultural coherence. In *Proceedings of the 27th Conference on Cross-Language Studies and Contrastive Linguistics*.

Fass, Dan. 1986. Collative semantics: An approach to coherence. Memorandum in Computer and Cognitive Science MCCS-86-56, New Mexico State University, New Mexico.

Fass, Dan. 1987. Collative semantics: An overview of the current meta5 program. Memorandum in Computer and Cognitive Science MCCS-87-112, New Mexico State University, NM.

Fass, Dan. 1988. Collative semantics: A semantics for natural language processing. Memorandum in Computer and Cognitive Science MCCS-88-118, New Mexico State University, NM.

Fass, Dan. 1991. Met: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, 17(1):49–90.

Fellbaum, Christiane, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Goatly, Andrew. 1997. *The Language of Metaphors*. Routledge, London.

Gruber, Jeffrey. 1976. *Lexical Structures in Syntax and Semantics*. Amsterdam, North-Holland.

Jain, Anil K., M. Narasimha Murty, and Patrick J. Flynn. 1999. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323.

Kilgarriff, Adam. 2003. BNC word frequency list. Available online at http://www.itri.brighton.ac.uk/Adam.Kilgarriff/bnc-readme.html.

Kucera, Henry. 1992. Brown corpus. In S. Shapiro, editor, *Encyclopedia of Artificial Intelligence*, volume 1. Wiley, New York, pages 128–130.

Lakoff, George. 1993. The contemporary theory of metaphor. In Andrew Ortony, editor, *Metaphor and Thought*. Cambridge University Press, Cambridge.

Lakoff, George, Jane Espenson, and Alan Schwartz. 1991. The master metaphor list. Draft 2nd ed. Technical Report, University of California at Berkeley.

Lenat, Douglas. 1995. Cyc: A large-scale investment in knowledge infrastructure. In *Communications of the ACM*, 38:11.

Li, Hang and Naoki Abe. 1998. Generalizing case frames using a thesaurus and the MDI principle. *Computational Linguistics*, 24(2):217–244.

Marcus, Mitchell P., Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19:313–330.

Martin, James. 1990. *A Computational Model of Metaphor Interpretation*. Academic Press.

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

Mason, Zachary. 2002. *A Computational, Corpus-Based Metaphor Extraction System*. Ph.D. thesis, Brandeis University.

Peters, Winn and Ivonne Peters. 2000. Lexicalised systematic polysemy in WordNet. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, Athens.

Porter, Martin F. 1980. An algorithm for suffix stripping. *Program*, 14(3):130–137.

Resnik, Philip. 1993. *Selection and Information: A Class Based Approach to Lexical Relationships*. Ph.D. thesis, University of Pennsylvania.

Sekine, Satoshi and Ralph Grishman. 1995. A corpus-based probabilistic grammar with only two non-terminals. In *Proceedings of the Fourth International Workshop on Parsing Technology*, Prague, Czech Republic.

Wilks, Yorick. 1975. A preferential, pattern-seeking, semantics for natural language inference. *Artificial Intelligence*, 6:53–74.

Wilks, Yorick. 1978. Making preferences more active. *Artificial Intelligence*, 11(3):197–223.