# Metaphor as an Emergent Property of Machine-Readable Dictionaries

**William B. Dolan**
Microsoft Research
One Microsoft Way
Redmond, WA 98052-6399
billdol@microsoft.com

**Abstract**

Previous computational attempts to handle non-literal word usage have been restricted to "toy" systems that combine hand-coded lexicons with restricted sets of metaphor types that can be used to sanction specific classes of semantic subcategorization violations. These hand-coded efforts are unlikely to ever scale up to the rigors of real, free text. We describe an example-based approach to metaphor interpretation which exploits a large lexical knowledge base derived from a machine-readable dictionary. We first present an extended example which is meant to demonstrate the scope of evidence about common English metaphors which is encoded in ordinary dictionaries. We then go on to demonstrate how novel instances of metaphor can be interpreted by accessing this information. A given violation of some default semantic expectation can be checked against the LKB to determine whether this violation is an instance of some more systematic mapping of English word meanings from one semantic domain to another. We argue that this approach to metaphor interpretation obviates the need for the traditional "metaphor-handling component" in natural language understanding systems, and will allow these systems to overcome the brittleness of hand-coded approaches.

## Introduction

This paper argues that a machine-readable dictionary (MRD) can provide the basis for a robust and conceptually simple approach to the computational understanding of non-literal language, and that a lexicon-based metaphor processing strategy has significant advantages over other strategies.

Previous work on this difficult problem (e.g. Martin 1990; Fass 1988; DeJong & Waltz 1983) has centered around a model in which a natural language understanding (NLU) system identifies and passes potential instances of metaphorical usage to a special metaphor-handling component. This component operates according to rules which are significantly different in character from the rules which are used to interpret "normal", literal usage. The metaphor component is assumed to contain an exhaustively specified set of metaphor types, each reflecting a mapping between two otherwise distinct semantic domains (e.g. *people* and *plants*). "Interpreting" an instance of non-literal usage involves linking the usage to one of these pre-specified metaphorical equations and, in more ambitious systems, using information about this linkage to account for why the input violates default semantic expectations. For instance, in processing the sentence *Robbie's running ate up the distance*, Martin's system attempts to interpret the unexpected pairing of *eat up* with *distance* by linking this verb-object pair to the hand-coded metaphor *Eat-Up/Reduce-Money*. Finding evidence in its lexicon that *distance* and *money* are analogous in that both can be *reduced* in *amount*, the program "interprets" the input sentence by substituting *distance* for *money* in the equation *eat up:food::reduce: money*.

However, such efforts to address the metaphor interpretation problem share a number of serious flaws. Most important is the fact that none can be considered more than a "toy" system, since each exploits a small, hand-constructed, restricted-domain lexicon and each handles a relatively small set of hand-coded metaphor types. Scaling up to the rigors of real, free text would involve far more than simply adding more metaphors and word senses, and in fact the task looks impossible. A major difficulty is the bootstrapping problem: which metaphorical equations will be relevant for handling unrestricted text, and how — apart from exhaustively enumerating these equations through some process of introspection — can they be incorporated into the system's metaphorical model? How can we ensure, for every word and sense in the language, that the lexicon encodes adequate information about the kinds of semantic properties that can be analogically equated through processes of English metaphor?

Using evidence from a large lexical knowledge base (LKB) that we have derived from the on-line version of the Longman Dictionary of Contemporary English (LDOCE), we will argue that explicit enumeration of metaphors like *Eat-Up/Reduce-Money* is both unnecessary for and ultimately detrimental to the goal of understanding non-literal language. Instead, we show that the lexicon itself is a sufficiently rich source of information about metaphorical sense extensions to allow the understanding of a wide range of non-literal uses *without any explicit model of metaphor*. In effect,

the strategy that we suggest involves treating the lexicon as a huge example base of possible word sense extensions, and allowing the lexicon itself to directly determine whether or not a particular meaning extension is licensed in a particular context. In demonstrating this claim, we show that a large, richly-structured LKB contains far more information about possible English metaphors than could ever be expressed by an approach which lists prespecified classes of possible meaning extensions.

In order to illustrate this point, we informally describe an algorithm which examines the LKB and attempts to identify systematic mappings between distinct semantic domains. The result of this algorithm is that many interesting metaphorical connections "emerge" from the data contained the LKB. This is followed by a brief discussion of how this information about English metaphors can be exploited to interpret novel uses of words.

## Metaphors from MRDs

The dictionary entry for almost any polysemous word will show that MRDs constitute a very rich source of evidence concerning the metaphorical extension of word meanings. Frequently, such entries will include first a central, core, literal sense of the word (e.g. the animal sense of *pig*), while one or more of the subsequent senses will represent some extension of that central sense to some other domain (e.g. the *person* sense of *pig*). However, dictionaries typically fail to explicitly characterize the relationships between different word senses; there is no indication, for example, either that this *animal/human* sense alternation is metaphorical, or that it is just one instance of a very common and productive kind of systematic polysemy which allows (usually negative) characteristics of animals to be extended to humans.

How can we overcome this limitation of MRDs and make explicit the systematic nature of such sense extensions across the lexicon of English? Our approach relies on a large LDOCE-derived LKB and on tools for navigating through the semantic network implicit in this LKB. This LKB is the product of automatically processing the definition strings of the 33,000 single word noun definitions and the 12,000 single word verb definitions (45,000 definitions total) in the on-line version of the LDOCE (For further details, see Dolan et al. 1993 and Richardson et al. 1993.) Parsing and semantic analysis yield more than 100,000 instances of approximately 25 different types of semantic relations that link a headword to a word in the text of its definition. This work builds on earlier work by Jensen & Binot (1987) and Montemagni & Vanderwende (1992). Among the relation types currently identified are *Hypernym* (*Is_a*), *Part_of*, *Typical_Object*, *Means_of*, and *Location_of*. A second stage of processing involves sense-disambiguating the words which are linked by semantic relations; the resulting LKB thus encodes semantic links between individual word senses, rather than dictionary headwords. Each prose definition is mapped into a logical form representation, or semantic relation structure, in which word senses are linked by labeled semantic arcs, and the lexicon as a whole is modeled as a huge directed graph. Network navigation software (Richardson, forthcoming) allows efficient and fast identification of paths linking any arbitrary pair of words or word senses within the lexical network.

Our strategy for identifying and exploiting common "conventional" metaphors (Lakoff & Johnson 1980) within this LKB is conceptually quite simple. We begin with the observation that sequences of word senses within a dictionary entry frequently reflect a metaphorical shift of some kind, as in the case of *pig*. The algorithm examines the different senses of a polysemous word, focusing on semantic alternations between the sets of word senses which appear as the values of a given semantic relation for each. For example, it will discover that the words *game* occurs as the value of a *Location_of* relation on one sense of the verb *score*, while the word *argument* occurs as a *Location_of* on a different sense of this verb. A reasonable hypothesis is that if two distinct classes of *Location_of* values are linked through the different senses of a word, then perhaps there are other links within the LKB that connect the same two semantic domains. The apparently random correspondence between *game* and *argument* in this entry can thus be used to point the way to the huge set of metaphorical associations between English words closely linked to *(sporting) games* and words having to do with *(verbal) arguments*. Hypotheses like these can be directly tested by searching for paths through the lexicon which pass through these words. In effect, then, we use individual dictionary entries to seed a discovery process that searches the entire knowledge base for evidence corroborating the hypothesized metaphorical connection.

We will now show in more detail how such an hypothesis is formulated and tested. Consider how the algorithm identifies an interesting metaphorical connection indicated by the different semantic classes of *Typical_Objects* associated with two senses of the transitive verb *plant*. The first step involves considering the set of *Typical_Objects* that occur within LDOCE for the first verb sense of *plant*, "to put plants or seeds in the ground to grow". This verb sense occurs in dozens of the semantic relation structures extracted from the 45,000 definitions we consider. Within this set of structures, seven distinct nouns occur as the direct object of this verb sense of *plant* (Figure 1).

| Object of *plant* | | Source | Definition of source |
|---|---|---|---|
| seed, n 1 | < | *plant*, v 1 | to put (plants or seeds) in the ground to grow |
| | < | seed, v 2 | to *plant* seeds (in a piece of ground) |
| | < | drill, n 4.2 | a row of seeds often *planted* in this way |
| plant, n 1 | < | *plant*, v1 | to put (plants or seeds) in the ground to grow |
| shrub, n 1 | < | shrubbery, n1 | (part of a garden *planted* with) shrubs forming a mass or group |
| plantation, n 2 | < | plantation, n2 | a large group of growing trees *planted* esp. to produce wood |
| grove, n 1 | < | grove, n1 | a group of fruit trees, *planted* or natural, sometimes in a garden |
| blackthorn, n 1 | < | blackthorn, n1 | a type of European plant...often *planted* in rows... |
| box, n 1 | < | box, n1 | a small type of tree... often *planted* in rows as a wall or a fence |

**Figure 1: Nouns which occur as the object of *plant, v 1* in the set of 45,000 LDOCE definitions**

Note that these seven nouns—things which can be *planted*—form a tightly coherent semantic class: all are botanical words. (See Dolan 1994, for more complex methods of comparing senses for overall semantic similarity.) Identifying the class to which these word senses belong constitutes the second step of our algorithm, which uses our system's path identification tools to discover the shortest network of links within our LKB capable of connecting all these words. As Figure 2 shows, the shortest paths all pass through the zero-derived noun sense *plant, n 1*. Even more important is the fact that the *nature* of these paths is highly restricted: in each case, the path is short and passes exclusively through *Hypernym* or *PartOf* relations, indicating a that a direct ancestor relationship holds between each of these seven noun senses and the core noun sense of *plant*.

shortest path connecting them is a direct *Hypernym* link: a *belief* is a kind of *idea*. Once again, then, we have found a case in which the *Typical_Objects* of a verb sense form tightly linked semantic class.

Next comes the most interesting step. As mentioned above, the fact that two distinct classes of *Typical_Objects* for the two senses of *plant* have been identified suggests that there might be a metaphorical connection between *plants* on the one hand and *ideas/beliefs* on the other hand. Intuitively, it is obvious that this is true: *ideas* are like *plants* in that they can *grow*, *spring up*, *wither*, and so on. The only question is how to confirm this hypothesis without appeals to human intuition: how can the lexicon itself be used to confirm or deny the existence of this metaphorical association between two distinct semantic domains?

The solution involves using our system's network

| | | | | |
|---|---|---|---|---|
| seed, n 1 | - *PartOf* -> | plant, n 1 | | |
| blackthorn, n 1 | - *Hypernym*-> | plant, n 1 | | |
| plantation, n 2 | - *Hypernym* -> | tree, n 1 | - *Hypernym* -> | plant, n 1 |
| grove, n 1 | - *Hypernym*-> | tree, n 1 | - *Hypernym* -> | plant, n 1 |
| box, n 1 | - *Hypernym* -> | tree, n 1 | - *Hypernym* -> | plant, n 1 |
| shrub, n 1 | - *Hypernym*-> | bush, n 1 | - *HypernymOf*-> | tree, n 2 - *Hypernym* -> plant, n 1 |

**Figure 2: Lexical paths linking *Typical_Objects* of *plant, v1* to *plant, n1* (the core noun sense of *plant*)**

So far, then, the algorithm has discovered that the core verb sense of *plant* takes a well-defined semantic class of nouns as its direct object. This is a typical scenario; the set of *Typical_Objects* associated with a given transitive verb sense in LDOCE are usually closely related semantically. In particular, it is often possible to identify either a shared *Hypernym* for all of these objects, or else a chain of *Part_of* relations which lead to a common *Hypernym*.

The third step taken by our algorithm involves repeating the first and second steps, this time examining the set of *Typical_Objects* associated with the second verb sense of plant, "to put (an idea, belief, etc.) in the mind." In this case, *idea* and *belief* turn out to be the only *Typical_Objects* of this sense within LDOCE. The exploration tools to identify any paths in the lexical network which link *Typical_Objects* from the first sense of *plant* to *Typical_Objects* from the second verb sense of plant. Are there, for instance, significantly ranked paths linking *idea* to *seed*, *idea* to the core verb sense of *plant*, or *belief* to *seed*? We will focus here on just the set of paths returned for the first of these pairings, *idea/seed*. As it turns out, these two words are linked by 188 paths within the network currently defined by our LDOCE-derived LKB. Crucially, many of these paths are extremely salient, based on a ranking measure which takes into the length of the path and other factors (see Richardson, forthcoming, for a description of how paths are computed and weighted.) Five highly-ranked paths are given in Figure 3.

| | | | | | |
|---|---|---|---|---|---|
| seed, n 4 | *-Hypernym->* | germ, n2 | *-PartOf->* | idea, n1 | **< seed, n4 & germ, n2** |
| seed, n 1 | *-TypicalSubjOf->* | begin, v1 | *-TypicalObject->* | idea, n 1 | **< seedbed, n1 & author, n2** |
| seed, n 1 | *-TypicalObjOf->* | form | *-TypicalObject->* | idea, n 1 | **< stigma, n2 & foresee, v1** |
| seed, n 2 | *-Hypernym->* | form | *-TypicalObject->* | idea, n 1 | **< seed, n3 & foresee, v1** |
| seed, n 1 | *-TypicalObjOf ->* | develop, v 1 | *-Synonym->* | germinate, v 2 *-TypicalSubject->* | idea, n 1 **< set, v17 & germinate, v2** |

**Figure 3: Top-ranked lexical paths linking *seed* and *idea*, along with their source definitions**

A number of interesting points can be made about this set of paths. In particular, consider the words which constitute the "intersection" of *idea* and *seed*—those which are, in effect, the pivot point on the path which connects the semantic relation structures derived from two different LDOCE definitions. For instance, the most highly ranked path between the two is one which connects the structures derived from the following definitions of the words *seed* and *germ*. The path through these definitions connects via (different senses of) the word *germ*.

*seed, n4*: "something from which growth or development begins; beginning; GERM"
*germ, n2*: "a beginning point, esp. of an idea"

The intersection words which occur most frequently and on the shortest paths in this set are *produce*, *plant*, *germ*, *germinate*, *form*, *set*, *begin*, *develop*, and *start*. (The word *produce*, for example, appears as the intersection for 39 of the 188 paths.) These intersection words fall precisely into the class that interests us: words which reflect a pervasive metaphorical association between *plants* and *ideas* in the English lexicon. Both *seeds* and *ideas* are kinds of *germs*, they can *begin*, *form*, *develop*, *germinate*, and so on. (Once adjective definitions are folded into our LKB, we will also find links through words like *fertile*, *green*, and *ripe*.) Words which provide the source definitions of these paths also yield interesting support for the notion of a metaphorical connection between *ideas* and *plants*. Among these words are *pollinate*, *start*, *evolution*, *kernel*, *development*, *grow*, *seedbed*, *bear*, *resurgence*, *force*, and *stigma*.

A point that should emphasized is that because the path search mechanism traverses labeled links in our LKB, the resulting sets of paths do more than simply confirm or refute the existence of a metaphorical link between two domains. These paths also allow us to infer something about the semantic nature of the metaphorical shift in question. Consider two LDOCE senses of the verb *fight*:

*fight, v1* "*to use violence against (esp. others of one's kind)...*"
*fight, v3* "*to use argument against (someone, or each other)*"

The first of these senses occurs in a number of LDOCE definitions, always with a handheld weapon as its *Means* (e.g., *poleaxe*, *hand*, *sword*, *lance*, *fist*), while the definition of the other sense of *fight* provides its only *Means* value, *argument*. Examining the paths which connect *weapon* (the shared hypernym of most of the *Means* of *fight, v1*) to *argument* confirms that the vocabulary of English includes a huge number of words which reflect an association between physical warfare and verbal disagreements. Among the intersection words in this set of paths are *point*, *defence*, *fight*, and *attack*; among the words whose definitions contribute to these paths are *harm*, *wound*, *aim*, and *fray*. However, we can go beyond the simple conclusion that English displays a metaphorical association between these two types of disputes. In particular, we can infer from the fact that what differs between the first and third senses of the verb *fight* is the value of the *Means* relation that the "violent disagreement" component of the core *warfare* meaning is carried over into the *verbal disagreement* sense. What crucially differs between these two senses is the *Means* of *fighting*—in one case with weapons, and in the other with words.

The metaphor discovery process described in this section forms the basis of a straightforward strategy for handling novel metaphorical uses in on-line NLU system. We sketch this approach in the next section.

## Interpreting Novel Instances of Metaphor

Consider the problem of understanding a sentence like *The idea flourished*. There is no sense of the intransitive verb *flourish* in LDOCE which would allow this sentence to be directly interpreted; the only available senses have as their *Typical_Subjects* the words *plant* and *person*. The input sentence thus violates the default semantic preferences encoded in the dictionary. Using exactly the same path discovery process sketched in the previous section, however, the system can readily test whether these defaults are being violated in a way that is

consistent with the huge example base of metaphorical extensions captured by the dictionary-derived LKB. A search for paths linking the expected subject *plant* to the actual subject *idea* yields a huge set of highly-ranked paths, providing ample evidence for the hypothesis that *ideas* and *plants* are related through metaphor. In this case, the alternation we encounter is in the *Typical_Subject* relation tied to the first intransitive sense of this verb: "(esp. of plants) to grow healthily". We can go on to make the inference that those parts of the definition of this verb which are not involved in the metaphorical extension (i.e., "to grow healthily") will be equally well applicable to the subject *idea*.

In cases where evidence for a metaphorical connection between two semantic domains is weak, further corroboration can sometimes be obtained by iterating the path-discovery process. On the second pass, a highly-ranked intersection or source word will be substituted for one of the words used in the initial sense. Consider the problem of understanding a sentence like *The idea withered.* The words *idea* and *wither* are linked by three paths in the dictionary network, each passing through the word *development* and exhibiting the core noun sense of *plant* as their intersection points. However, each of these paths is assigned a low rank, indicating that they cannot be regarded as useful corroborating evidence for the assumption that there is a significant metaphorical connection between *ideas* and *withering*. However, if we now iterate the path search using *idea* and the newly-discovered intersection word *plant*, the huge set of paths linking *idea* and *plant* suddenly emerges. Many of these paths are highly salient, including one which links the two words directly through *plant*.

Iterating paths in this way frequently makes it possible to discover a set of very strong associations to support a metaphorical connection even when there is no direct one. In effect, this involves exploring the semantic space around two (or more) word senses in the input string, looking for ways to align them with some highly salient cross-domain mapping that is more directly reflected in the lexicon. This process very similar to the kind of inferencing strategies that our system uses to find associations between words in processing sentences which do not involve metaphor.

In addition to allowing for the interpretation of an arbitrary set of metaphor types, the approach to metaphor interpretation that we have described here also has the advantage of subtly weighting the strength of different metaphors. While a hand-coding approach which lists correspondences like *people=plants* would imply that each of these equations has equal status, the lexicon-based model we have proposed will produce different levels of confidence for a novel use which exploits an extremely pervasive correspondence like this, versus one which appears to exploit a much more sporadic correspondence like *mountains = human bodies* (*foot*, *flank* and *shoulder* are among the few lexicalized instances of this metaphor.)

## Conclusions

By relying on a richly structured lexicon as an example base for the interpretation of metaphor, natural language understanding systems should be able to overcome the brittleness of previous approaches to this problem. Rather than relying on a special component to handle non-literal language, we have suggested that metaphors can be processed using exactly the same machinery which is used to process literal usage: tools which search the lexicon for similarities between words. No single set of "conventional" metaphors needs to be hard-coded in the NLU system, and at no stage of processing is it necessary to explicitly distinguish between "literal" and "non-literal" speech. This latter point is especially important, since in any broad-coverage lexicon—whether derived from an MRD or built by hand—distinguishing between literal and metaphorical word meanings is a difficult exercise with no obvious utility for an NLU system.

In the model described in this paper, the interpretation of novel instances of non-literal language is handled by consulting the lexicon and dynamically searching for evidence that might link this usage to an existing class of meaning extensions. As a result, a potentially open-ended number of different metaphors can be handled by the system, with this set limited only by the number of words and senses contained in the lexicon, and by the level of detailed information associated with each of these senses. Increasing the coverage of the lexicon will only increase its ability to determine whether a particular lexical alternation is an instance of some broader class of metaphorical usage.

A final advantage of our approach over previous efforts is that it eliminates the need to elaborate by hand the set of semantic features that might potentially be isolated and exploited in a novel metaphorical usage. Rather than attempt this daunting task, we have chosen to rely on lexicographers' intuitions about which properties of word meanings are salient enough to be included in their dictionary definitions. It is exactly these properties, we suggest, that are the fundamental building blocks of English metaphor.

### Acknowledgments

Thanks to Diana Peterson, Karen Jensen, Joseph Pentheroudakis, Lucy Vanderwende, George Heidorn, John Messerly, and Deb Coughlin. Special gratitude is due to Steve Richardson, who built the path identification software that made this work possible.

## References


DeJong, G. F. and D. Waltz. 1983. *Understanding Novel Language.* Computers and Mathematics with Applications, 9.

Dolan, W., L. Vanderwende, and S. Richardson (1993). "Automatically Derived Structured Knowledge Bases from On-line Dictionaries", Proceedings of the First Conference of the Pacific Association for Computational Linguistics, April 21-24, Simon Fraser University, Vancouver, Canada.

Dolan, W. 1994. "Word sense ambiguation: clustering related senses." Proceedings of COLING-94, Kyoto, Japan.

Fass, D. 1988. Collative Semantics: A Semantics for Natural Language. Ph.D. thesis, New Mexico State University, Las Cruces, New Mexico. CRL Report No. MCCS-88-118.

Jensen, K. and J.-L. Binot (1987). "Disambiguating prepositional phrase attachments by using on-line dictionary definitions". Computational Linguistics 13: 3-4, pp. 251-260.

Lakoff, G. and M. Johnson. 1980. *Metaphors We Live By.* University of Chicago Press, Chicago, Illinois.

Martin, J. H. 1990. *A Computational Model of Metaphor Interpretation.* Academic Press, Boston.

Montemagni, S. and L. Vanderwende (1992). "Structural patterns vs. string patterns for extracting semantic information from dictionaries." In Proceedings of COLING92, pp. 546-552.

Richardson, S., L. Vanderwende and W. Dolan. 1993. "Combining dictionary-based and example-based methods for natural language analysis." Proceedings of the Fifth International Conference on Theoretical and Methodological Issues in Machine Translation.

Richardson, S. (forthcoming) Ph.D. dissertation, CUNY.