Extending the Lexicon by Exploiting Subregularities*

Robert Wilensky

Computer Science Division
Department of EECS
University of California, Berkeley
Berkeley, CA 94720
wilensky@teak.berkeley.edu

1. Introduction

This paper is concerned with the acquisition of the lexicon. In particular, we propose a method that uses analogical reasoning to hypothesize new polysemous word senses. This method is one of a number of knowledge acquisition devices to be included in DIRC (Domain Independent Retargetable Consultant). DIRC is a kind of intelligent, natural language-capable consultant kit that can be retargeted at different domains. DIRC is essentially "empty-UC" (UNIX Consultant, Wilensky et al., 1988). DIRC is to include the language and reasoning mechanisms of UC, plus a large grammar and a general lexicon. The user must then add domain knowledge, user knowledge and lexical knowledge for the area of interest.

2. Previous Work in Acquisition of the Lexicon.

There have been numerous attempts to build systems that automatically acquire word meanings. Mostly, these have been either dictionary readers or attempts to hypothesize meanings of completely unfamiliar words from context (e.g., Selfridge (1982), Granger (1977)). In contrast, we have focussed on the problem of acquiring word senses that are related to ones already known.

2.1. A Note on Word Senses

For the purposes at hand, we are only concerned with word senses that are synchronically related. These may be polysemous senses of individuals words, as well as related senses of different words. In addition, we distinguish *meanings* or *conception structures* of a word from senses. (We will use the term "meaning" and "conceptual structure" interchangely in this context.) A single meaning may be involved in any number of senses, each of which has grammatical or other differences. Typically, a word has at least one *core meaning* from which the meanings involved in other senses are in some sense synchronically based.

For example, the word "open" has adjectival and verbal senses; the verbal senses include some whose meaning is, roughly, making physical access available to an enclosed region by moving some object (e.g., "open a jar", "open a draw", "open the door"). This is probably a core meaning of the word. There are several senses involving this meaning, just among the verbal senses. These senses are differentiated from one another by how the components of the meaning relate to the verb's valence. For example, one sense has the object moved as the patient, and hence as the direct object of the transitive verb (as in "open the door"); another uses the container itself as the direct object (e.g., "open the jar"); perhaps another involves some sort of aperture that widens (e.g., "open your throat" or "open the pupil of your eye"). Additionally, each of these components of the meaning can be realized as patients by appearing as the subject of the intransitive version of the verb. We consider each differentiable valence structure for both the transitive and intransitive verb forms as constituting different senses, although we presume that the same conceptual structure is in all of these examples.

Yet other senses of "open" have the meaning of causing an information-containing item to come into existence (e.g., "open a bank account" or "open a file on someone"). This second meaning is probably based on the first one. Also, the various adjective uses (e.g., "the open door") are separate senses in this view having some as yet unspecified relation to the senses described above. Finally, other words, e.g., "close", have senses that we presume to be related to the various senses of "open" just discussed.

2.2. MIDAS

Previously, we have succeeded in doing some automatic lexical acquisition by exploiting conventional metaphors as motivations for linguistic forms. In particular, Martin (1988) implemented the MIDAS system which both uses metaphoric word senses to help with language under-

standing, and to extend the lexicon when a new metaphoric use of a word is encountered. For example, the sentence "John have Mary a cold." is presumed to make recourse to a "a cold is a possession" metaphor. We call such a conventionalized metaphor a *core* metaphor, since it seems to serve as the basis for related metaphoric uses. Thus, the sentence "John gave Mary a cold" is presumed to involve the "infecting with a cold is giving the cold" metaphor, which entails the previous "cold is possession" metaphor.

Suppose the system encounters an utterance like "John got the flu from Mary", but is not familiar with this use of the verb "get", nor with the notion of a flu being treated as a possession. Then both the available non-metaphoric sense of "get", along with the metaphors involving diseases and possession, are brought to bear to hypothesize the word sense that might be in play. Hypotheses are generated by two kinds of lexical extension processes: core extension and similarity extension. Understanding "get a cold" given an understanding of "give a cold" involves core extension, as the core metaphor "cold is possession" is extended to the "getting" concept; understanding "get the flu" given an understanding of "get a cold" involves similarity extension, as the generalization about a role in the metaphoric structure must be extended from colds to diseases in general. Understanding "get the flu" given an understanding of "give a cold" involves both kinds of extension.

The MIDAS system has been used in conjunction with UC to extend metaphoric word senses in the computer domain. The following is an example of MIDAS learning a new sense of the word "kill", given that it knows some metaphoric extensions of this sense outside the computer domain.

```
-------------------------------
# How can I kill a process?

No valid interpretations. Attempting
to extend existing metaphor.

Searching for related known metaphors.

Metaphors    found:    Kill-Conversation
Kill-Delete-Line Kill-Sports-Defeat

Selecting metaphor Kill-Conversation
to extend from.

Attempting    a    similarity    extension
inference.

Extending    similar    metaphor    Kill-
Conversation    with    target    concept
Terminate-Conversation.
```

```
Abstracting  Terminate-Conversation  to
ancestor concept

Creating new metaphor:

   Mapping main source concept Killing
to main target concept

      Terminate-Computer-Process

   Mapping source role killer to target
role c-proc-termer.

   Mapping source  role  kill-victim  to
target role c-proc-termed.

Calling UC:

You  can  kill  a  computer  process  by
typing ^c to the shell.

-------------------------------
```

Here MIDAS first retrieves a number of metaphors related to the input; of these, "Kill-Conversation" is chosen as most applicable. A simple similarity extension is attempted, resulting in a proposed "Terminate-Computer-Process" metaphor for interpretation of the input. The interpretation thus provided is passed along to UC, which can answer this question. Meanwhile, the metaphor is incorporated into UC's knowledge base, which allows UC's language generator to use the same terminology in encoding the answer.

MIDAS is discussed in detail in Martin (1988).

3. Why MIDAS Works

We believe that MIDAS works because it is exploiting metaphoric subregularity by a form of analogical reasoning. That is, it finds a metaphorical usage that is closest to the given case according to some conceptual metric; it then exploits the structure of the prior metaphor usage to construct an analogous one for the case at hand, and proposes this new structure as a hypothetical word sense. Note that according to this explanation, metaphor does not play a crucial role in the extension process. Rather, it is the fact that the metaphor is a subregularity rather than the fact that it is a metaphor that makes it amenable to analogical exploitation.

Analogy, of course, has played a prominent role in traditional linguistics. Indeed, rather influential linguists (for example, Paul (1891) and Bloomfield (1933) seemed to attribute all novel language use to analogy. However, today, analogy seems almost entirely relegated to diachronic processsses. A notable exception to this trend

is the work of Skousen (in press), who appears to advocate a view quite similar to our own, although the primary focus of his work is morphological.

Analogy has also been widely studied in artificial intelligence and cognitive psychology. The work of Carbonell (1982) and Burstein (1983) is most relevant to our enterprise, as it explores the role of analogy in knowledge acquisition. Similarly, Alterman's (1985, 1988) approach to planning shares some of the same concerns. However, many of the details of Carbonell's and Alterman's proposals are specific to problem solving, and Burstein's work is focused on formulating constraints on the relations to be considered for analogical mapping. Thus, their work does not appear to have an obvious application to our problem. Many of the differences between analogical reasoning for problem solving and language knowledge acquisition are discussed at length in Martin (1988).

Another line of related work is the connectionist approach initiated by Rumelhart and McClelland (1987), and explicitly considered as an alterative to acquisition by analogy by MacWhinney et al. (1989). However, there are numerous reasons we believe an explicitly analogical framework to be superior. The Rumelhart-McClelland model maintains no memory of specific cases, but only a statistical summary of them. Also, the analogy-based model can use its knowledge more flexibly, for example, to infer that a word encountered is the past tense of a known word, a task that an associationist networks could not easily be made to perform. In addition, we interpret as evidence supportive of a position like ours psycholinguistic results such as those of Cutler (1983) and Butterworth (1983), which suggest that words are represented in their full "undecomposed" form, along with some sorts of relations between related words.

3.1. Other Kinds of Lexical Subregularities

If MIDAS works by applying analogical reasoning to exploit metaphoric subregularities, then the question arises as what other kinds of lexical subregularities there might be. One set of candidates is approached in the work of Brugman (1981, 1984) and Norvig and Lakoff (1987). In particular, Norvig and Lakoff (1987) offer six types of links between word senses in what they call *lexical network theory*. However, their theory is concerned only with senses of one word. Also, there appear to be many more links than these. Indeed, we have no reason to believe that the number of such subregularities is bounded in principle.

We present a partial list of some of the subregularities we have encountered. The list below uses a rather informal rule format, and gives a couple of examples of

words to which the rule is applicable. It is hoped that explicating a few examples below will let the reader infer the meanings of some of the others:

(1) function-object-noun -> primary-activity-"determinerless"-noun
("the bed" -> "in bed, go to bed"; "a school -> at school"; "my lunch -> at lunch"; "the conference -> in conference")

(2) noun -> resembling-in-appearance-noun
("tree" -> "(rose) tree"; "tree" -> "(shoe) tree"); "tiger" -> "(stuffed) tiger", "pencil" -> "pencil (of light)")

(3) noun -> having-the-same-function-noun
("bed" -> "bed (of leaves)")

(4) noun -> "involve-concretion"-verb
("a tree" -> "to tree (a cat)"; "a knife" -> "to knife (someone)")

(5) verb -> verb-w-role-splitting
("take a book" -> "take a book to Mary", "John shaved" -> "John shaved Bill")

(6) verb -> profiled-component-verb
("take a book" -> "take a book to the Cape")

(7) verb -> frame-imposition-verb
("take a book" -> "take someone to dinner", "go" -> "go dancing")

(8) activity-verb-t -> concretion-activity-verb-i
("eat an apple" -> "eat [a meal]", "drink a coke" -> "drink [alcohol]", "feed the dog" -> "the dog feeds")

(9) activity-verb-t -> dobj-subj-middle-voice-verb-i
("drive a car" -> "the car drives well")

(10) activity-verb-i -> activity-verb-t-primary-category
("John dreamed" -> "John dreamed a dream"; "John slept" -> "John slept the sleep of the innocent")

(11) activity-verb-i -> do-cause-activity-verb-t(patient as subject)
("John slept" -> "The bed sleeps five")

(12) activity-verb -> activity-of-noun
("to cry" -> "a cry (in the wilderness)"; "to punch" -> "a punch (in the mouth)")

(13) activity-verb <-> product-of-activity-noun
("copy the paper" <-> "a copy of the paper"; xerox, telegram, telegraph)

367

(14) functional-noun –> use-function-verb
("the telephone" –> "telephone John"; machine, motorcycle, telegraph)

(15) object-noun –> central-component-of-object
("a bed" –> "bought a bed [=frame with not mattress]; "an apple" –> "eat an apple [=without the core]"))

Consider the first rule. This rule states that, for some noun whose core meaning is a functional object, there is another sense, also a noun, that occurs without determination, and means the primary activity associated with the first sense. For example, the word "bed" has as a core meaning a functional object used for sleeping. However, the word can also be used in utterances like "go to bed" and "before bed" (but not, say, "*during bed"). In these cases, the noun is determinerless, and means something akin to sleeping. Other examples include "jail", "conference", "school" and virtually all the meal terms, e.g., "lunch", "tea", "dinner". British English allows "in hospital", while American English does not.

The dialect difference underscores the point that this is truly a subregularity: concepts that might be expressed this way are not necessarily expressed this way. Also, we chose this example not because it in itself is a particularly important generalization about English, but precisely because it is not. That is, there appear to be many such facts of limited scope, and each of them may be useful for learning analogous cases.

Consider also rule 4, which relates function nouns to verbs. Examples of this are "tree" as in "The dog treed the cat" and "knife" as in "The murderer knifed his victim". The applicable rule states that the verb means some specific activity involving the core meaning of the noun. I.e., the verbs are treated as a sort of conventionalized denominalization. Note that the activity is presumed to be specific, and that the way in which it must be "concreted" is assumed to be pragmatically determined. Thus, the rule can only tell us that "treeing" involves a tree, but only our world knowledge might suggest to us that it involves cornering; similarly, the rule can tell us that "knifing" involves the use of a knife, but cannot tell us that it means stabbing a person, and not say, just cutting.

As a final illustration, consider rule 5, so-called "role splitting" (this is the same as Norvig and Lakoff's semantic role differentiation link). This rule suggests that, given a verb in which two thematic roles are realized by a single complement may have another sense in which these two complements are realized separately. For example, in "John took a book from Mary", John is both the recipient and the agent. However, in "John took a book to Mary", John is only the agent, and Mary is the recipient. Thus, the sense of "take" involved in

the first sentence, which we suggest corresponds to a core meaning, is the basis for the sense used in the second, in which the roles coinciding in the first are realized separately. A similar prediction might be made from an intransitive verb like "shave", in which agent and patient coincide, to the existence of a transitive verb "shave" in which the patient is realized separately as the direct object. (Of course, the tendency of patients to get realized as direct objects in English should also help motivate this fact, and can presumably also be exploited analogically.)

4. An Analogy-based Model of Lexical Acquisition

We have been attempting to extend MIDAS-style word hypothesizing to be able to propose new word senses by using analogy to exploit these other kinds of lexical subregularities. At this point, our work has been rather preliminary, but we can at least sketch out the basic architecture of our proposal and comment on the problems we have yet to resolve.

(A) Detect unknown word sense. For example, suppose the system encountered the following phrase:

"at breakfast"

Suppose further that the function noun "breakfast" were known to the system, but the determinerless usage were not. In this case, the system would hypothesize that it is lacking a word sense because of a failure to parse the sentence.

(B) Find relevant cases/subregularities. Cues from the input would be used to suggest prior relevant lexical knowledge. In our example, the retrieved cases might include the following:

bed-1/bed-3, class-1/class-4

Here we have numbered word senses so that the first element of each pair designates a sense involving a core meaning, and the latter a determineless-activity type of sense. We may have also already computed and stored relevant subregularities. If so, then these would be retrieved as well.

Relevant issues here are the indexing and retrieval of cases and subregularities. Our assumption is that we can retrieve relevant cases by a conjunction of simple cues, like "noun", "functional meaning", "extended determinerless noun sense", etc., and then rely on the next phase to discriminate further among these.

(C) Chose the most pertinent case or subregularity. Again, by analogy to MIDAS, some distance metric is used to pick the best datum to analogize from. In this

case, perhaps the correct choice would be the following:

class-1/class-4

One motivation for this selection is that "class" is compatible with "at", as is the case in point.

Finding the right metric is the primary issue here. The MIDAS metric is a simple sum of two factors: (i) the length of the core-relationship from the input source to the source of the candidate metaphor, and (ii) hierarchical distance between the two concepts. Both factors are measured by the number of links in the representation that must be traversed to get from one concept to the other. The hierarchical distance factor of the MIDAS metric seems directly relevant to other cases. However, there is no obvious counterpart to the core-relationship component. One possible reason for this is that metaphoric extensions are more complex than most other kinds; if so, then the MIDAS metric may still be applicable to the other subregularities, which are just simpler special cases.

(D) Analogize to a new meaning. Given the best case or subregularity, the system will attempt to hypothesize a new word sense. For example, in the case at hand, we would like a representation for the meaning in quotes to be produced.

class-1/class-4 ::
breakfast-1/"period of eating breakfast"

In the case of MIDAS, the metaphoric structure of previous examples was assumed to be available. Then, once a best match was established, it is relatively straightforward to generalize or extend this structure to apply to the new input. The same would be true in the general case, provided that the relation between stored polysemous word senses is readily available.

(E) Determine the extent of generalization. Supposing that a single new word sense can be successfully proposed, the question arises as to whether just this particular word sense is all the system can hypothesize, or whether some "local productivity" is possible. For example, if this is the first meal term the system has seen as having a determinerless activity sense, we suspect that only the single sense should be generated. However, if it is the second such meal term, then the first one would have been the likely basis for the analogy, and a generalization to meal terms in general may be attempted.

(F) Record a new entry. The new sense needs to be stored in the lexicon, and indexed for further reference. This task may interact closely with (E), although generalizing to unattested cases and computing explicit subregularities are logically independent.

There are many additional problems to be addressed beyond the ones alluded to above. In particular, there is the issue of the role of world knowledge in the proposed process. In the example above, the system must know that the activity of eating is the primary one associated with breakfast. A more dramatic example is the role of world knowledge in hypothesizing the meaning of "treed" in expressions like "the dog treed the cat", assuming that the system is acquainted with the noun "tree". All an analogical reasoning mechanism can do is suggest that some specific activity associated with trees is involved; the application of world knowledge would have to do the rest.

5. Other Directions of Investigation

We have also been investigating exploiting subregularities in "intelligent dictionary reading". This project involves an additional idea, namely, that one could best use a dictionary to gain lexical knowledge by bringing to bear on it a full natural language processing capability. One problem we have encountered is that dictionaries are full of inaccuracies about the meaning of words. For example, even relatively good dictionaries have poor entries for the likes of determinerless nouns like "bed". E.g., Webster's New World (Second Edition) simply lists "bedtime" as a sense of "bed"; Longman's Dictionary of Contemporary English (New Edition) uses "in bed" as an example of the ordinary noun "bed", then explicitly lists the phrase "time for bed" as meaning "time to go to sleep", and gives a few other determinerless usages, leaving it to the reader to infer a generalization.* However, a dictionary reader with knowledge of the subregularity mentioned above might be able to correct such deficiencies, and come up with a better meaning that the one the dictionary supplies. Thus, we plan to explore augmenting our intelligent dictionary reader with the ability to use subregularities to compensate for inadequate dictionary entries.

We are also attempting to apply the same approach to acquiring the semantics of constructions. In particular, we are investigating verb-particular combinations and conventionalized noun phrases (e.g., nominal compounds). We are also looking at constructions like the ditransitive (i.e., dative alternation), which seem also to display a kind of polysemy. Specifically, Goldberg (1989, 1990) has argued that much of the data on this construction can be accounted for in terms of subclasses that are conventionally associated with the construction itself, rather than with lexical rules and transformations as proposed, for example, by Gropen et al. (1989). If so, then the techniques for the acquisition of polysemous

---
*Longman's also defines "make the bed" as "make it ready for sleeping in". We have no idea how to cope with such errors, but they do underscore the problem.

369

lexical items should prove equally applicable to the acquisition of knowledge about such constructions. We are attempting to determine whether this is the case.

6. References

Alterman, Richard. Adaptive Planning: Refitting Old Plans To New Situations. In the *Proceedings of The Seventh Annual Conference of the Cognitive Science Society*, 1985.

Alterman, Richard. Adaptive Planning. In *Cognitive Science*, vol. 12, pp. 393-421, 1988.

Bloomfield, L. *Language*. New York: Holt, Rinehart & Winston, 1933.

Brugman, Claudia. The Story of *Over*. University of California, Berkeley M.A. thesis, unpublished. Distributed by the Indiana University Linguistics Club. 1981.

Brugman, Claudia. The *Very* Idea: A Case-Study in Polysemy and Cross-Lexical Generalization. In *Papers from the Twentieth Regional Meeting of the Chicago Linguistics Society*. pp. 21-38. 1984.

Burstein, Mark H. Concept Formation by Incremental Analogical Reasoning and Debugging. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach*, vol. II. Tioga Press, Palo Alto, California, 1982.

Butterworth, B. Lexical representation. In B. Butterworth (ed.), *Language Production* , vol. 2. Academic Press, New York, 1983.

Carbonell, Jaime. Learning by analogy: Formulating and Generalizing Plans from Past Experience. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (eds.) *Machine Learning: An Artificial Intelligence Approach*. Tioga Press, Palo Alto, California, 1982.

Cutler, A. Lexical complexity and sentence processing. In G. B. Flores d'Arcais & and R. J. Jarvella (eds.), *The Process of Language Understanding*, pp. 43-79. Wiley, New York, 1983.

Goldberg, Adele. A Unified Account of the Semantics of the Ditransitive Construction. *BLS 15*, 1989.

Goldberg, Adele. The Inherent Semantics of Argument Structure: The Case of the English Ditransitive Construction. Unpublished manuscript, 1990.

Granger, R. H. FOUL-UP: A Program that figures out the meanings of words from context. In the *Proceedings of the Fifth International Joint Conference on Artificial Intelligence. Cambridge, MA. 1977.*

MacWhinney, B. *Competition and Lexical Categorization. In R. Corrigan, F. Eckman and M. Noonan, Linguistic Categorization.* John Benjamins Publishing Company, Amsterdam/Philadelphia, 1989.

Martin, James. Knowledge Acquisition through Natural Language Dialogue. In the *Proceedings of the 2nd Conference on Artificial Intelligence Applications.* Miami, Florida. December, 1985.

Martin, James. A Computational Theory of Metaphor. Berkeley Computer Science Technical Report no. UCB/CSD 88/465. November 1988.

Norvig, Peter. Building a large lexicon with lexical network theory. In the *Proceedings of the IJCAI Workshop on Lexical Acquisition.* August 1989.

Norvig, Peter and Lakoff, George. Taking: A Study in Lexical Network Theory. In the *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society.* Berkeley, CA. February 1987.

Paul, H. *Principles of the History of Language.* Longmans, Green, London, 1891.

Rumelhart, D., & McClelland, J. Learning the past tenses of English verbs: Implicit rules of parallel distributed processes? In B. MacWhinney (ed.), *Mechanisms of Language Acquisition.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1987.

Selfridge, M. Computer Modeling of Comprehension Development. In W. G. Lehnert & M. H. Ringle, *Strategies for Natural Language Processing.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1982.

Skousen, R. *Analogical Models of Language.* Kluwer, Dordrecht, (in press).

Wilensky, R., Mayfield, J., Chin, D., Luria, M., Martin, J. and Wu, D. The Berkeley UNIX Consultant Project. *Computational Linguistics* 14-4, December 1988.