Lexical Acquisition with WordNet and the Mikrokosmos Ontology

Tom O'Hara, Kavi Mahesh and Sergei Nirenburg
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003-0001
tomohara,mahesh,sergei@crl.nmsu.edu

Abstract

This paper discusses an approach to augmenting a lexicon for knowledge-based machine translation (KBMT) with information derived from WordNet. The Mikrokosmos project at NMSU's Computing Research Laboratory has concentrated on the creation of the Spanish and Japanese lexicons, so the English lexicon is less developed. We investigated using WordNet as a means to automate portions of the English lexicon development. Several heuristics are used to find the WordNet synonym sets corresponding to the concepts in the Mikrokosmos language-independent ontology. Two of these heuristics exploit the WordNet *is-a* hierarchy: one performs hierarchical matching of both taxonomies, and the other computes similarity based on frequency of defining words and their ancestors in a corpus. The result is a lexicon acquisition tool that produces plausible lexical mappings from English words into the Mikrokosmos ontology. Initial performance results are included, which indicate good accuracy in the mappings.

1 Introduction

1.1 Problem Area

It's an understatement that lexicon acquisition is a costly endeavor. Traditional dictionaries have been developed over the course of decades through the employment of many lexicographers and numerous consultants. Furthermore, the development of semantic lexicons incurs additional cost for the explicit encoding of meaning representations that provide details often omitted in traditional dictionaries, which are written for humans not computers. Atkins (1995) estimates that it would take 100 person-years to properly develop a semantic lexical database comparable in scope to a standard college dictionary.

Lexicons are a key component of machine translation systems (Onyshkevych and Nirenburg, 1994). The Mikrokosmos (μK) project at NMSU's Computing Research Laboratory is developing Spanish, Chinese and Japanese lexicons to support knowledge-based machine translation (KBMT). The following table indicates the amount of effort that was required for developing the initial Spanish lexicon entries from scratch (Viegas et al., 1996):

> 6798 word-sense entries (as of 29 Mar 1996)
> Average of 1.2 meaning per word form
> Acquisition rate: 45 entries/day per person
> Acquisition effort: 4 person years

Like many research centers, we don't have the human resources to construct the entire lexicons manually, so we are investigating several different ways to automate lexicon acquisition. Viegas et al. (1996) discuss one approach at this through the use of lexical rules, such as for generating the morpho-semantic derivatives of Spanish verbs.

A natural solution would be to take advantage of machine readable dictionaries (MRD's), such as Longman's Dictionary of Contemporary English (LDOCE). This approach was popular in the eighties; however, it achieved mixed success. See (Wilks et al., 1996) for a comprehensive survey of MRD research. One of the main problems is that dictionaries aren't explicit about the particular senses of the words used in definitions. In addition, much knowledge of the world is assumed by the dictionary entries. Consider the second LDOCE sense of the noun "whistle" (Procter, 1978):

> whistle, n. ... 2. the high sound made by passing air or steam through a small tube-shaped area, either an instrument, a mouth, or a beak

Obviously, "air" does not mean musical melody, even though this is closer to the topic area of the definition than the gas-mixture sense. Note also that the definition assumes the reader knows what the actual sound is like and just needs to know enough to distinguish it from similar sounds, such as from a flute.

As indicated in (Nirenburg, 1994), relying solely on MRD's will not yield a lexicon sufficient for realistic applications. These need to be fortified with machine-readable versions of other reference sources, such as thesauri, collocation lists, and perhaps even encyclopedias. Furthermore, text corpora from the specific application domain should be used for training and testing. In the Mikrokosmos project, this type of information has been consulted during the development of the language-independent ontology, most of which was created manually (Mahesh and Nirenburg, 1995). Therefore, it is conceivable to automate much of the lexical acquisition by mapping entries from MRD's into concepts from the ontology. As long as a word being defined is a simple refinement of a concept present in the ontology, this mapping forms the basis of the lexical representation, which can then be refined manually if necessary. The work reported here demonstrates the feasibility of this approach by mapping entries from Princeton's WordNet (Miller, 1990) into our ontology.

Like a thesaurus, WordNet is structured around groups of synonymous words. In WordNet these groups are called synsets and are taken as indistinguishable in particular contexts. Like a dictionary, it provides definitions and usage examples. However, unlike both, it provides explicit relationships among the synsets (e.g, *is-a* and *has-a*). Thus WordNet also represents an explicit ontology of concepts, provided that word senses are considered as concepts. It is this lexicalized ontology that facilitates mapping the WordNet senses (i.e., synsets) for a English word into the corresponding concept in the Mikrokosmos ontology.

Note that there are several reasons why mapping directly from the WordNet synset to the equivalent in a particular foreign language is undesirable. The main reason is that mapping into a language-neutral ontology facilitates richer text meaning representations that are not tied to the specifics of particular languages. For more details on this and other benefits of KBMT see (Nirenburg et al., 1992).

1.2 Overview of the Solution

Overall, the algorithm is fairly straightforward and incorporates few domain-specific dependencies. To begin with, the mapping algorithm does not perform parsing or pattern matching of the definition entries but instead relies on the conventions used in the ontology development. For instance, the names of concepts use the corresponding English word whenever possible; in cases where no English word provide a suitable label, the name is based on the concept's parent label (e.g., VOLUNTARY-PERCEPTUAL-EVENT, a kind of PERCEPTUAL-EVENT). Therefore, when a word from a synset and a μK concept name are the same or just slight variations, there is a good chance that the same concept is being referred to. The control mechanism thus is a generate-and-test cycle applied to each μK concept: generate words potentially in WordNet synsets and test the senses of these for the best matches with the Mikrokosmos concept.

The algorithm incorporates two main heuristics that exploit the WordNet ontology hierarchy. One specifically assigns the highest weight to the potential synset/concept mapping which has the highest degree of overlap between the respective *is-a* hierarchies. For instance, when mapping to the concept DOG, the canine sense of "dog", which goes through MAMMAL and to ENTITY, is preferred over the scoundrel sense, which goes through PERSON to ENTITY.

The other heuristic uses synset frequencies, estimated from a corpus of Wall Street Journal articles. This is based on a technique for disambiguating noun groups using WordNet by Resnik (1995). For each pair of defining words that share a common ancestor, the support for the match is increased by an amount inversely proportional to the frequency of the ancestor, because unrelated words only have common ancestors at the top level of the hierarchy.

Two other heuristics exploit the ontology but in a more localized manner. One assigns a weight based on the degree of overlap among the children for each concepts. The other does the same thing for the siblings of each. The final heuristic computes the degree of overlap among the words in the definition texts. This was meant mainly as a weak supplement to the others to help discriminate close mappings, but it turned out to be more accurate than the other localized heuristics.

In cases where a fully-automated system might not be considered robust enough for a particular application domain, this approach can easily be adapted to an interactive one. In this case, the main benefit would be that the human lexicographer can be relieved of much of the tedious aspects of the English lexicon development. Specifically, the matching of new words against concepts in the ontology can be done automatically, along with providing definition glosses and examples. Therefore, the lexicographer can concentrate on filling in the details necessary to realize the lexical entry.

2 Background and Examples

In Mikrokosmos, most of the language-neutral information is stored in the ontology. Each concept is represented by a rich frame structure, which allows for numerous links among the concepts. For instance, it includes theme and instrument relations, as well as the more usual *is-a* and set-membership relations; the ontology also contains selectional restrictions on case roles. In contrast, the lexicon contains just the information needed to realize a concept in a given language (Onyshkevych and Nirenburg, 1994). It will contain the usual information about morphology and syntax; but, most of the semantics will be defined in terms of a concept from the ontology. Therefore, the semantics for the lexicon entry might just be a direct mapping to the associated

```
(book
    (book-N1
        (cat n)                         ;; category
            (morph)                     ;; morphology
            (anno                       ;; annotations
            (def "a copy of a written work or composition
                that has been published")
                (ex "I just read a good book on economics")
                (cross-ref))
            (syn)                       ;; syntactic features
            (syn-struc                  ;; syntactic structure
                (1 ((root $var0)
                    (cat n)) ))
            (sem-struc                  ;; semantic structure
                (lex-map                ;; lexical mapping
                    (1 (BOOK)) ))       ;;   to concept book
            (lex-rules)                 ;; lexical rules
            (pragm)                     ;; pragmatics
            (styl)))                    ;; stylistics
```

Figure 1: Lexical representation for "book"

concept. This is illustrated in the lexical entry for "book" in figure 1.

Nouns form the bulk of any lexicon and often have direct mappings into the ontology. Therefore, this tool facilitates acquisition significantly by producing many entries that are nearly complete. In contrast, the verbal entries it produces are only partially complete since WordNet doesn't provide detailed information on subcategorizations and selectional restrictions. Two examples will be given to illustrate the problems that arise in the matching process. The first one is what one would think as a trivial case to match, namely a concept for a simple concrete object, a chair. The other illustrates problems with different ontological decompositions, specifically two different views of singing.

2.1 Direct mapping for "chair"

The μK hierarchy for the concept CHAIR is shown in figure 2. WordNet has four synsets or senses for "chair", so each of these are potential candidates for a direct mapping:

- s1. chair: a seat for one person, with a support for the back
- s2. professorship, chair: the position of professor
- s3. president, chairman, chairwoman, chair, chairperson: the officer who presides at the meetings of an organization
- s4. electric chair, chair, death chair, hot seat: an instrument of death by electrocution that resembles a chair

Since synsets 2 and 3 cover the human agent sense of chair, they won't match well with the μK concept. But senses 1 and 4 will produce similar matches since they both are derived from ARTIFACT, as in μK:

```
chair
⇒seating-furniture
    ⇒furniture
        ⇒interior-building-part
            ⇒building-part
                ⇒building-artifact
                :   ⇒artifact
                :       ⇒inanimate
                :           ⇒separable-entity
                :               ⇒physical-object
                :                   ⇒object
                :                       ⇒all
                ⇒place
                    ⇒physical-object
                        ⇒object
                            ⇒all
```

Figure 2: μK concept hierarchy for CHAIR

```
s1: chair
        ⇒seat
            ⇒furniture, ..., article of furniture
                ⇒furnishings
                    ⇒instrumentality, instrumentation
                        ⇒artifact, artefact
                            ⇒object, ..., physical object
                                ⇒entity

s4: electric chair, chair, death chair, hot seat
        ⇒instrument of execution
            ⇒instrument
                ⇒device
                    ⇒instrumentality, instrumentation
                        ⇒artifact, artefact
                            ⇒object, ..., physical object
                                ⇒entity
```

A problem that complicates selecting the appropriate sense is that μK classifies FURNITURE under the generic INTERIOR-BUILDING-PARTS whereas WordNet uses the more specific FURNISHINGS.

2.2 Problematic mapping with "sing"

The verb "sing" illustrates what could go wrong when trying to match from WordNet to μK. There are two main reasons for this problem: The WordNet verb hierarchy is much shallower than μK's event hierarchy; and, the concept of singing has been represented differently. Here is the μK hierarchy for SING:

```
sing: to produce musical sounds with the voice
    ⇒human-voice:
    the sound made through the mouth by
    human beings
        ⇒emit-sound:
        to create wave-like disturbances in the air
```

⇒wave-energy-event:
events in which light, sound, etc.
waves are transmitted or emitted
⇒physical-event:
events involving physical force
⇒event:
any activity, action, happening, or situation
⇒all:
refers to any concept

The salient aspect is singing as emitting waveform energy. In contrast, the WordNet hierarchy of the closest synset for "sing" emphasizes the communicative aspects of singing:

s2: sing:
produce musical tones with the voice
⇒talk, speak, utter, mouth, verbalize:
express in speech
⇒communicate, ..., transmit feelings:
n/a
⇒interact, ..., act towards others:
act together with others
⇒act, move, take measures, ...:
carry out an action

The other senses cover miscellaneous meanings of "sing":

s1. sing, deliver by singing: n/a
s3. whistle, sing: let off steam; as of tea kettles
s4. tattle, talk, blab, ...: divulge information or secrets; spill the beans

Consequently, the hierarchy match will not produce any alignment; and, the similarity match will not be effective since the synset frequency counts are indirectly based on the WordNet hierarchy. But the text match will still be applicable. Plus, since the children & sibling matches are localized, a plausible match can still be generated.

3 Implementation

The Onto-WordNet Mapper works by performing a breadth-first traversal of the μK concept space, attempting to find matches for each concept node with a synset from WordNet. The end result is a list of potential mappings sorted by a match score derived by weighting the scores from the individual heuristics. An empty list indicates that no suitable matches were determined. This mapping process is detailed in figure 3, which also shows the default weights used prior to the optimization discussed later. The following sections describe each of the five matching heuristics. Note that a separate component, not described here, is used to produce the lexicon entry from the best match, provided the score is above a certain threshold.

For each Mikrokosmos concept:

1. Get candidate synset
   (a) Try to find a word in WordNet with the same spelling (e.g. REAL-ESTATE vs. "real estate").
   (b) Try to find a word matching a prefix or suffix of the concept (e.g., PEPPER-VEGETABLE vs. "pepper").
2. Perform structure match of the synset and concept hierarchies. For each word and concept pair:
   (a) Check for exact match of the word and concept.
   (b) Check for partial match of the word and concept (as above).
   (c) Check predefined equivalences.
   (d) Evaluate each match by computing the percent of matched nodes on the best-matching branches for each (scaled by length).
3. Perform concept-similarity match using corpus-derived probabilities
   (a) Get words occurring in the definition glosses for synset & concept.
   (b) Compute pairwise similarity by finding ancestor with the highest information content (the most-informative-subsumer).
   (c) Evaluate the match by the degree of support the synset gets from all of the most-informative-subsumers that are applicable.
4. Perform intersection matches for the following:
   (a) the sibling synsets & concepts.
   (b) the children synsets & concepts.
   (c) the definition gloss words from the synset & concept
5. Compute total match score by a weighted sum: $.25 * hier + .25 * sim + .2 * child + .2 * sibl + .1 * text$.

Figure 3: Onto-WordNet Mapping Algorithm

3.1 Hierarchical Match

The hierarchy match (see figure 4) computes a score for the similarity of the two concept hierarchies. Since WordNet gives several words per synset, the matching at each step uses the maximum scores of the alternatives. The matching proceeds node by node up the hierarchies. If a given node doesn't match, it is skipped, but it still is included the total number of nodes.

As given here, this algorithm is quite inefficient since similar subproblems are generated repeatedly. In the actual solution, the results from previous matches are cached, making the solution comparable to one based on dynamic programming. Note that this problem is related to the "Longest Commatch-hierarchies(wn, onto)

1. If both lists empty, the score and node-count is 0
2. If either hierarchy is empty, the score is likewise 0. Determine node-count from the nodes in the other hierarchy.
3. Compute the similarity of the WordNet synset and μK concept names. If the result is above a preset threshold (0.75), add it to the score, and tally in score of recursive match of the parents:
   match-hierarchies(rest(wn), rest(onto))
4. Otherwise, take the maximum of the scores from the recursive matches in which the WordNet node and/or the μK node is skipped:
   max(match-hierarchies(rest(wn), onto),
   match-hierarchies(wn, rest(onto)),
   match-hierarchies(rest(wn), rest(onto)))
   This is done for each possible pairing of the hierarchy paths, in case either concept has more than one parent.

Figure 4: Hierarchy Match Heuristic

mon Subsequence" (LCS) problem, which is solved via dynamic programming in (Cormen et al., 1990). Their recursive formulation for the problem follows:

Let $l[i, j]$ be the length of an LCS of the sequences X_i and Y_j:

$$l[i,j] = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ l[i-1, j-1] + 1 & \text{if } x_i = y_i \\ max(l[i, j-1], l[i-1, j]) & \text{if } x_i \neq y_i \end{cases}$$

3.2 Similarity Match

The idea of the similarity heuristic is to use the information content of the common ancestor concepts (or subsumers) for words in the definition texts. This is based on the technique Resnik (1995) uses for disambiguating noun groups. The frequency of a synset is based on the frequency of the word plus the frequencies of all its descendant synsets. Therefore, the top-level synsets have the highest frequencies and thus the highest estimated probability of occurrence. For each pair of nouns from the text of both definitions (one each from WordNet and μK), the *most-informative-subsumer* is determined by finding the common ancestor with the highest information content, which is inversely related to frequency. The information content of this ancestor determines the pairwise similarity score. The candidate synset that receives the most support from the pairwise similarity scores is then preferred. These calculations are detailed in figures 5 and 6.

This technique requires an estimation for the frequencies of the WordNet synsets. Unless the corpus has been annotated to indicate the WordNet synset for each word, there is no direct way to determine the synset frequencies. However, these can be estimated by taking the frequency of the words in all descendant synsets (i.e., all the words the synset subsumes).

$$Words_c = \{Word_i | Word_i \text{ has sense subsumed by } c\}$$

$$Freq_c = \sum_{w \in Words_c} Count_w$$

$$\hat{P}_c = \frac{Freq_c}{N}$$

$$sim(w_i, w_j) = \max_{c \in subsumers(w_i, w_j)} [-log\ \hat{P}_c]$$

Figure 5: Calculation of similarity (Resnik, 1995)

match-similarity(synset, concept)
  for each pair of nouns w_i, w_j from the definitions
    $sim_{i,j}$ = calc-sim(w_i, w_j)
    $mis_{i,j}$ = most-informative-subsumer(w_i, w_j)
    if $mis_{i,j} \in$ subsumers(synset) then
      increase synset-support by $sim_{i,j}$
    increase normalization by $sim_{i,j}$
  score is synset-support scaled by normalization

Figure 6: Similarity-based heuristic

3.3 Miscellaneous matching heuristics

The remaining heuristics are similar in that they each are based on degree of overlap in word-based matching. For instance, in the match-siblings heuristic, the siblings sets for the candidate synset and μK concept are compared by determining the size of the intersection relative to the size of the μK set. The match-children and match-definition-text heuristics are similar. Figure 7 shows the general form of these intersection-based matching heuristics. This uses an equivalence test modified to account for a few morphological variations; the test also accounts for partial matches with the components of a concept name (similar to first step in figure 3).

```
match-word-lists (wn-list, onto-list)
    wn-list = normalize(wn-list)
    onto-list = normalize(onto-list)
    overlap = intersection(wn-list, onto-list, similar-form)
    score = length(overlap) / (1 + length(onto-list))

similar-form(word1, word2)
    return (word-similarity(word1, word2) >= 0.25)
```

Figure 7: General form of intersection heuristics

4 Evaluation

To evaluate the performance of the mapper, two sets of 100 random concepts were mapped by hand into the corresponding WordNet synset (or marked as not-applicable). The first set was selected from the entire set of concepts mapped, whereas the second set was selected just from the cases with more than one plausible mapping (e.g., the corresponding WordNet entry has more than one synset). The results of this test shows that it handles 77% of the ambiguous cases and 94% of the cases overall, excluding cases corresponding to lexical gaps in WordNet (see table 1). This shows an improvement of more than 15% over the lower bound, which was estimated from the proportion of correct mappings using sense 1. Note that these tests were performed after development was completed on the system.

Type	Correct	Lower	Accuracy
ambiguous	70/91	60.2	*76.9*
unrestricted	59/64	90.7	92.2

Table 1: Evaluation of mapper performance

The remainder of this section presents results on analyzing how often the individual heuristics contribute to the correct result. The most important finding is that the hierarchy and text match account for most of the results. Furthermore, when all heuristics are used together, the similarity heuristic has a minor contribution to the result, although it is second when heuristics apply individually.

Table 2 contains the results for each heuristic evaluated individually against the manual mapping of the ambiguous cases. Note that the overall score shows the accuracy using the default weights for comparison purposes.

As a rough estimate for optimizing the weighting scheme, regression analysis was performed on the score produced by each heuristic to the result of the manual mapping for the ambiguous cases. This accounts for the interactions among the heuristics. There are 343 data points, because the score for each sense is included, not just those for the current sense. See table 3. Although the correlation coefficient is only 0.41, the regression suggests that the hierarchy match and the text match are the most significant heuristics. When using these revised weights, the accuracy increases to 81.3%.

Heuristic	Correct	Accuracy
Hierarchy	66	0.725
Siblings	45	0.495
Text	42	0.462
Similarity	58	0.637
Children	12	0.132
Overall	70	0.769

Table 2: Individual accuracy results (91 cases)

Heuristic	Coefficient	StdError	Weight
Hierarchy	0.815	0.134	0.275
Siblings	0.237	0.069	0.080
Text	0.976	0.176	0.329
Similarity	0.332	0.077	0.112
Children	0.605	0.178	0.204

Table 3: Regression on results (n=343; R^2=0.41)

An alternative method for determining these weights used an optimization search, which accounts for nonlinear relationships. This method produced the weights shown in table 4. This shows that only the hierarchy and text heuristics contribute significantly to the result. When these are applied to the ambiguous sample, the accuracy becomes 83.5%. Note that the results given earlier uses the lower figure, because this represents the evaluation before training the weights on the sample.

Heuristic	Default Weight	Optimized Weight
Hierarchy	0.25	0.40
Siblings	0.20	0.10
Text	0.10	0.30
Similarity	0.25	0.10
Children	0.20	0.10

Table 4: Optimization search for weights

These results are preliminary: larger test sets would be required before conclusions can be drawn. However, it seems clear that a statistical approach is not likely to serve as a complete solution for this problem. Instead, a combination of symbolic and statistical approaches seems appropriate, with an emphasis on the former.

5 Relation to other work

Work of this nature has been more common in matching entries in multilingual dictionaries (e.g., (Rigau and Agirre, 1995)) than in lexical acquisition. This section will concentrate on work augmenting lexical information by ontological mappings.

Knight and Luk (1994) describe an approach to establish correspondences between Longman's Dictionary of Contemporary English (LDOCE) and WordNet entries. A definition match compares overlap of the LDOCE definition text to those of both the WordNet entry and its hypernym along with the words from closely-related synsets. Their hierarchy match uses the implicit hierarchy within LDOCE defined from the genus terms of the definitions, incorporating work done at NMSU (Bruce and Guthrie, 1991) that identifies and disambiguates the head nouns in the definition texts. The hierarchy is used to guide the determination of nontrivial matches by providing local context in which senses can be considered unambiguous by filtering out the other senses not applicable to either subhierarchy. It also allows for matching the parents of words from an existing match. Note that this mapping is facilitated by the target and source domains being the same: namely, English words. Therefore, the problem of assessing correspondence is minimized.

Chang and Chen (1996) describe an algorithm for augmenting LDOCE with information from Longman's Lexicon of Contemporary English (LLOCE). LLOCE is basically a thesaurus, with word lists arranged under 14 subjects and 129 topics. These topic identifiers are used as a coarse form of sense division. The matching algorithm works by computing a similarity score for the degree of overlap in the list of words for each LDOCE sense compared to the list of words from the LLOCE topics that contain the headword (expanded to include cross-references).

Other work is less directly related. Lehmann (1995) describes a methodology for semantic integration that matches classes based on the overlap in the inclusions of typical class members. For this to be effective, these instances must have been consistently applied in both ontologies. O'Sullivan et al. (1995) describe work on doing the reverse process we do. Specifically, they augment WordNet by linking in entries from an ontology describing word processing. However, their approach requires manual linking.

6 Conclusion

Combining traditional symbolic heuristics with a statistical approach yields an effective method for augmenting lexical acquisition. This report illustrated how this facilitated the mapping of WordNet synsets into a KBMT ontology. The symbolic approach included heuristics for structure matching and intersection-based comparisons. The statistical approach added a similarity test based on synset frequency estimated from a Wall Street Journal corpus. The result is a lexicon acquisition system that produces accurate mappings. This system has been used within the Mikrokosmos project to produce a basic lexicon of over 2000 entries, which were manually validated to ensure correctness. Additional mappings will be possible when the ontology is extended to other domains, since it now emphasizes business transactions. To allow for broader coverage, future work will address producing mappings that include refinements of the concepts from the ontology.

Although this work concentrated on nouns, the techniques can be extended to include other types of words. Furthermore, it can be generalized to handle ontology merging, in particular, the problem of merging classification systems. Lehmann (1995) points out that there are several practical ontologies suitable for merging to be used with a variety of intelligent applications, such as the Electronic Data Interchange (EDI) standard for descriptions of business transactions (ANSI, 1994). The idea is to take advantage of the time-consuming classification work already done.

Acknowledgements

The WordNet support code uses the CommonLisp interface developed by Mark Nahabedian of MIT. At CRL, many people provided valuable input for this work, including Jim Cowie, Mark Davis, Nick Ourusoff, Arnim Ruelas, Evelyne Viegas, and Janyce Wiebe.

References

ANSI, *X12 Standard*, Subrelease 003041, ANSI, Alexandria, VA, February 1994.

Atkins, B. (1995), "The Dynamic Database", in *Machine Tractable Dictionaries: Design and Construction*, C. Guo, ed., Norwood, NJ: Ablex Publishing Corporation, pp. 131-143.

Bruce, R., and L. Guthrie (1991), "Building a Noun Taxonomy from a Machine Readable Dictionary", Technical Report MCCS-91-207, Computing Research Laboratory, NMSU.

Chang, J., and J. Chen (1996), "Acquisition of Computational-Semantic Lexicons from Machine Readable Lexical Resources", in *Proc. ACL'96 Workshop on the Breadth and Depth of Semantic Lexicons*, June 28, 1995, Santa Cruz, CA.

Cormen, T., C. Leiserson, and R. Rivest (1990), *Introduction to Algorithms*, Cambridge, MA: MIT Press.

Knight, K., and S. Luk (1994), "Building a Large-Scale Knowledge Base for Machine Translation", in *Proc. Twelfth National Conference on Artificial Intelligence*, August 1-4, 1994, Seattle, Washington: American Association for Artificial Intelligence.

Lehmann, F. (1995), "Combining Ontologies, Thesauri, and Standards", in *Proc. Workshop on Basic Ontological Issues in Knowledge Sharing*, International Joint Conference on Artificial Intelligence (IJCAI-95), Aug. 19-20, 1995. Montreal, Canada.

Mahesh, K., and S. Nirenburg (1995), "A Situated Ontology for Practical NLP", in *Proc. Workshop on Basic Ontological Issues in Knowledge Sharing*, International Joint Conference on Artificial Intelligence (IJCAI-95), Aug. 19-20, 1995. Montreal, Canada.

Nirenburg S., J. Carbonell, M. Tomita and K. Goodman (1992), *Machine Translation: A Knowledge-based Approach*, San Mateo, CA: Morgan Kaufmann.

Nirenburg, S. (1994), "Lexicon Acquisitive for NLP: A Consumer Report", in *Computational Approaches to the Lexicon*, B. Atkins & A. Zampoli, eds., Oxford: Oxford University Press, pp. 313-347.

Miller, G., (1990), "WordNet: An on-line lexical database", *International Journal of Lexicography* 3(4).

Onyshkevych, B., and S. Nirenburg (1994), "The Lexicon in the Scheme of KBMT Things", Technical Report MCCS-94-277, Computing Research Laboratory, NMSU.

O'Sullivan, D., A. McElligott, R. Sutcliffe (1995), "Augmenting the Princeton WordNet with a Domain Specific Ontology", in *Proc. Workshop on Basic Ontological Issues in Knowledge Sharing*, International Joint Conference on Artificial Intelligence (IJCAI-95), Aug. 19-20, 1995. Montreal, Canada.

Procter, P. (1978), ed., *Longman Dictionary of Contemporary English*, Harlow, Essex: Longman Group.

Resnik, P. (1995), "Disambiguating Noun Groupings with Respect to WordNet Senses", in *Proc. Third Workshop on Very Large Corpora*, June, 1995.

Rigau, G. and E. Agirre (1995), "Disambiguating Bilingual Nominal Entries against WordNet", in *Proc. Computational Lexicon Workshop at the European Summer School in Logic, Language and Information*, Barcelona, pp. 71-82.

Viegas, E., B. Onyshkevych, V. Raskin, and S. Nirenburg (1996), "From Submit to Submitted via Submission: On Lexical Rules in Large-Scale Lexicon Acquisition", in *Proc. 31st Annual Meeting of the Association for Computational Linguistics*, Santa Cruz, CA.

Wilks, Y., B. Slator, and L. Guthrie (1996), *Electric Words*, Cambridge, MA: MIT Press.