Combining Dictionary-Based and Example-Based Methods for Natural Language Analysis

Stephen D. Richardson
steveri@microsoft.com

Lucy Vanderwende
lucyv@microsoft.com

William Dolan
billdol@microsoft.com

Microsoft Corp.
One Microsoft Way,
Redmond, WA 98052-6399

Abstract

We propose combining dictionary-based and example-based natural language (NL) processing techniques in a framework that we believe will provide substantive enhancements to NL analysis systems. The centerpiece of this framework is a relatively large-scale lexical knowledge base that we have constructed automatically from an online version of Longman's Dictionary of Contemporary English (LDOCE), and that is currently used in our NL analysis system to direct phrasal attachments. After discussing the effective use of example-based processing in hybrid NL systems, we compare recent dictionary-based and example-based work, and identify the aspects of this work that are included in the proposed framework. We then describe the methods employed in automatically creating our lexical knowledge base from LDOCE, and its current and planned use as a large-scale example base in our NL analysis system. This knowledge base is structured as a highly interconnected network of words linked by semantic relations such as *is_a, has_part, location_of, typical_object,* and *is_for.* We claim that within the proposed hybrid framework, it provides a uniquely rich source of information for use during NL analysis.

1. Introduction

We propose combining in a single framework aspects of two methods that have recently been the subject of much research in natural language (NL) processing. The first of these, dictionary-based (DB) processing, makes use of available machine readable dictionaries (MRDs) to create computational lexicons, some of which have been used in such tasks as sense disambiguation and phrasal attachment during NL analysis. The second method, example-based (EB) processing, uses example phrases or sentences taken from real text, represented either as strings or in some more structured form, to resolve ambiguities or determine corresponding translations in various machine translation (MT) systems.

The thesis of this paper is that these two methods are not only compatible, but in fact, that they share a number of common characteristics, and that these characteristics may be combined in such a way as to provide substantial benefit to NL analysis systems. At the heart of both methods is the assumption that natural language is an ideal knowledge representation language, both in terms of expressive power and

overall computational efficiency. This view has been asserted in other work that provided some of the basis for our current project (Jensen, et al. 1992) and is shared by other researchers as well (e.g., Wilks, et al. 1992).

In the past few years, DB research has focused mainly on aspects of NL analysis such as phrasal attachment (e.g., Jensen and Binot 1987, Vanderwende 1990) and word sense disambiguation (e.g., Braden-Harder 1992, Wilks, et al. 1992), while EB efforts in MT have dealt with both analysis and transfer processing (e.g., Okumura, et al. 1992, Jones 1992, Sumita and Iida 1991, Watanabe 1992). There has been some debate in the MT field whether EB methods may be used effectively during analysis, and in the next section, we provide a rationale for their use in this context. Together with their similarity to DB methods, this provides justification for their use in the proposed framework, which focuses on enhancing NL analysis systems. Also, we characterize the complementary nature of EB and rule-based (RB) processing in creating coherent, hybrid NL systems.

In the following section, we review and compare recent DB and EB work, identifying the aspects of this work that are included in our framework. The framework consists of the following four components:

1. A large, lexical knowledge base, created automatically from an online dictionary using DB methods, containing structured semantic relations between words, and accessed by the functions described in points 2, 3, and 4 below.
2. A similarity measurement function, based on the "semantic contexts" of words defined by their relations in the knowledge base, employing EB methods tor determining similarity, and used by the two functions in points 3 and 4 below.
3. A function for disambiguating word senses by matching the contexts of words in text with the semantic contexts of those words in the knowledge base.
4. A function for disambiguating phrasal attachments by representing attachment alternatives in the form of different semantic contexts for words, which are then matched against the semantic contexts of those words in the knowledge base.

In the final sections, we describe the DB methods employed and the results obtained in automatically creating a large-scale lexical knowledge base (the first and central component in the proposed framework) from an online version of Longman's Dictionary of Contemporary English (LDOCE). This knowledge base is structured as a highly interconnected network of words linked by semantic relations such as *is_a, has_part, location_of, typical_object,* and *is_for.* We conclude by briefly discussing the current use of the knowledge base in our NL analysis system and the planned uses, which fit within the proposed framework. Our NL analysis system is intended for eventual integration into various applications, including MT, information retrieval, and authoring tools.

2. The Use of Example-Based Processing

Researchers have recently debated how EB processing may be used most effectively. The question has arisen whether its use should be confined to transfer components of MT systems, or whether it can provide benefit to analysis components as well. Sumita, et al. (1990) state that "...it is not yet clear whether EBMT can/should deal with the whole process of translation." They go on to suggest that a gradual integration of EB methods with existing RB methods in MT systems is preferred and that experimentation will determine the correct balance. Sumita and Iida (1991) suggest pragmatic, but subjective criteria for implementing EB methods, including the level of translation difficulty and whether or not translations to be produced are compositional in nature. While these criteria may comprise reasonable guidelines, they lack any son of principled motivation.

Grishman and Kosaka (1992) also argue for a combination of RB and "empiricist" approaches, including EB, statistics-based (SB), and corpus-based (CB) processing. They suggest that these latter methods "...should be used to acquire the information which is more lexically or domain specific, and for which there is (as yet) no broad theoretical base." However, they seem to reduce the issue to a simple distinction between the analysis/generation components of MT systems (for which, they claim, theories are well-developed) and transfer components of those systems (for which theories are not so developed).

Jones (1992), after examining the arguments for and against hybrid (integrating RB with EB or SB processing) systems, opts for a non-hybrid, pure EB approach. He makes this choice based on his conviction that such methods are superior at handling "the complex issues surrounding human language," and seemingly, because no one else has tried yet to implement a completely non-hybrid MT system.

We claim that, while many of the reasons given above, at least those in favor of an integrated approach, are valid, a more principled rationale for the use of EB methods may be given. In essence it is that *examples specify contexts, contexts specify meaning, and therefore, EB methods are best suited to meaning-oriented, or semantic, processing, wherever it occurs.* The fact that examples specify contexts is obvious, but the point that contexts specify meaning is worth at least a bit of discussion, since we claim it in the strong sense, rejecting the general use of selectional features, lexical decomposition, and related methods which attempt to cast in concrete the fuzzy and flexible boundaries that exist in natural systems of lexical semantics. Others have confirmed their belief in the principle of context as meaning: Wilks (1972) states, "... except in those special cases when people do actually draw attention to the external world in connexion with a written or spoken statement, 'meaning' is always *other words*"; Sadler (1989) indicates that word matching in the DLT translation system "is based on the simple idea that meaning is context"; and Fillmore and Atkins (1991) define "word meaning" in terms of lengthy "when" definitions, which are nothing more than extended contexts.

This simple criterion for the use of EB methods is significant because it is consistent with our central assumption that natural language is ideal for knowledge representation. Not surprisingly, it also matches well with the uses to which DB methods have been applied, namely disambiguation of word senses and phrasal attachments. It is unlikely anyone would disagree that word sense disambiguation falls into the category of semantic processing. However, some may question the semantic nature of phrasal attachments. In response, we point out that in linguistic systems, structure and content often complement each other differently at different levels, and what is considered content at one level may be represented by structure at another. We believe this to be the case with the semantic content represented by phrasal attachments.

Furthermore, we position RB processing as complementary to EB processing, in the same way that structure is complementary to content. Where structural relationships may be clearly defined, or relatively small finite sets with fairly static boundaries may be established, the generalizing power of RB processing has proven itself to be highly effective. Our past experience has shown this to be especially true in the development of useful syntactic grammars (Jensen, et al. 1992). EB methods, on the other hand, excel at dealing with the vast multitude of subtle, fluid, contextual distinctions inherent in semantic processing. We therefore advocate the development of so-called "hybrid" NL systems, in which RB and EB[1] methods cooperate to form a coherent, powerful approach.

---

[1] We do not intend completely to exclude SB and CB methods from consideration here. They may also be used in a complementary fashion, and in fact, exhibit some of the same characteristics as EB methods in that they derive linguistic knowledge directly from large amounts of NL text, although they represent that knowledge in a non-linguistic form. General statistical techniques may also be applied at various levels in NL systems, but that is not within the scope of this paper.

3. A Comparison of Dictionary-Based and Example-Based Methods

We now examine the characteristics of recently developed DB and EB methods and compare them with aspects of components in our framework.

In the area of DB methods for word sense disambiguation, Lesk (1987) shows that by measuring the overlap between words in dictionary definitions and words in the context of a particular word to be disambiguated, a correct sense can be selected with a fair degree of accuracy for a small sample. In the "statistical lexical disambiguation method" described by Wilks, et al. (1992), a similar measurement of overlap is extended to take into account words that are "semantically related" to the words in the dictionary definitions. This relatedness factor is based on statistical co-occurrence processing across all of the words contained in all definitions in the dictionary. Matching of contexts to definitions in the Wilks scheme is performed by vector similarity measurements, which are similar to those used by Sato (1991) in his EB matching procedure. In the work by both Lesk and Wilks, the words in a dictionary definition (and possibly related words) may be thought of in EB terms as forming example contexts which are then matched against contexts in new text to perform sense disambiguation. While this matching does not make any use of semantic information other than that implicitly represented in co-occurrence data, the vector similarity measurements used by Wilks have been shown to be quite useful in information retrieval systems. The methods used in these measurements and the context matching based on them are applicable to the second and third components of the proposed framework.

Veronis and Ide (1990) augment this approach in another fashion, creating explicit links between content words in dictionary definitions and the entries for those words themselves, thereby creating a neural-network-like structure throughout the dictionary. The links provide a similar function to the relatedness factor in the Wilks system. Nodes representing words from the textual context of a word to be disambiguated are "activated" in the network, and the activation spreads forward through the links until the network reaches a stable state in which nodes representing the correct senses (definitions) have the highest activation level. In this work, the dictionary as an example base has an explicit structure, like the lexical knowledge base we propose for our first component, although the relationships represented by the links in this structure are not labeled. The connectionist matching strategy is similar to that which has been proposed by McLean (1992) for EB machine translation, however, connectionist methods have not been included in the framework.

Braden-Harder (1992) takes a somewhat different approach, making use of much of the explicitly coded information in LDOCE (e.g., grammatical codes and subject codes) as well as using a NL parser to extract genus terms from definitions and verbal arguments from example sentences. This information is then combined in a vector and matched (using techniques similar to those of Wilks and Sato mentioned above) against information gleaned from parsing the text surrounding the word to be disambiguated. The information in the vectors in this approach may be considered to constitute example contexts, and it is stored as it is generated in an updated form of the dictionary used by the parser. The "lexicon provider" method in Wilks, et al. (1992) also fills "frames" with information extracted from LDOCE to create sub-domain specific lexicons for use in parsing. It additionally uses a parser designed specifically for LDOCE to extract the genus term from the definitions. Wilks proposes the use of this parser to extract semantic relations such as *instrument* and *purpose* from the definitions as well. In the cases of both Braden-Harder and Wilks, the resulting enhanced dictionary entries provide the kind of deeply-processed, semantic information that both Sato (1991) and Sadler (1989) claim to be most desirable for inclusion in an example base. This is also the kind of information we desire for inclusion in our lexical knowledge base.

Vector-based matching by Braden-Harder is again applicable to second and third components of the framework.

The work by Jensen and Binot (1987) was the first of its kind in applying DB methods to the problem of directing phrasal attachment during parsing. They exploited the same NL parser that they were attempting to enhance in order to analyze definitions from Webster's Seventh New Collegiate Dictionary and extract information used to determine semantic relations such as *part_of* and *instrument*. They then used this information together with a set of heuristic rules to rank the likelihood of alternate prepositional phrase attachments. These rules may be thought of as defining a matching procedure, but the relationship to current EB matching schemes is somewhat weaker than with other DB work described above. Vanderwende (1990) extended this work to the determination of participial phrase attachments, following which Montemagni and Vanderwende (1992) significantly increased the number of semantic relations that were being extracted from the definitions. The list now included such relations as *subject_of, object_of*, *is_for, made_of, location_of,* and *means*. These relations were a natural extension to the set used by Jensen and Binot, and some of them had also been proposed by Wilks, but the realization of their extraction from 4,000 noun definitions resulted from using a broad-coverage NL parser and applying sophisticated structurally-based patterns to the parsed definitions. The use of this or a similar NL parser is essential to being able to extract information in the future from other dictionaries, reference sources such as encyclopedias, and eventually, free text. Although these relations were only generated dynamically and never stored in a form that could be called an example base, they nevertheless constitute the level of rich semantic information we seek for our lexical knowledge base. The use of the heuristic rules described for disambiguating phrasal attachments may be considered functionally as a limited version of what is desired for the fourth component of the framework..

To date, the kind of information present in the example bases of documented EB systems has ranged from unprocessed strings (many of the examples in Furuse and Iida 1992, CTM examples in Sato 1991), to simple syntactic patterns (Sumita and Iida 1991, Tsutsumi 1992, MBT1 examples in Sato 1991), to deeper, semantic structures (Sadler 1989, Takeda, et al. 1992). Invariably, the deeper the processing involved to produce the examples, the more difficult it is to obtain them. Sadler (1989) was able to populate the LKB in the DLT system with 76,000 examples by relying on the completely regular semantic case markings of Esperanto, something which truly natural languages do not exhibit. Takeda, et al. 1992) used a parser together with interactive human verification to obtain 30,000 examples from a limited domain dictionary. Most of the example bases described in the literature are much smaller in size, from a few hundred to a few thousand examples, and while some have been constructed by semi-automatic means, most have employed some degree of manual crafting. We suggest that the lexical knowledge base described in our framework, containing semantic relations extracted from definitions using a parser and sophisticated structural patterns, would be a significant step forward in providing a quantitatively and qualitatively better example base, at least tor the purpose of NL analysis.

A number of algorithms are used in EB systems to match incoming text with examples in the example base. Sumita and Iida (1991), Furuse and Iida (1992), and Tsutsumi (1992) make use of existing online synonym (and in the Tsutsumi case, taxonym) dictionaries. A similarity score is computed between words based on their proximity in the synonym/taxonym hierarchies. Sato (1991) uses vector similarity measurements between the words in his example base to determine similarity relationships which can then be applied to incoming text. Sadler (1989) also uses a computation that measures the overlap in his example base, but it uses a proximity function based on simple counts instead of vectors. Watanabe (1992) employs both a structure matching function (based on his used of tree-structure representations) and a node (word) matching function, which depends on the use of an unspecified semantic hierarchy (we assume similar to synonym hierarchies) and syntactic features in the case that the hierarchy is deficient. The similarity functions of Sato and Sadler are applicable to the second component in our framework, and since a good deal of the information in synonym hierarchies is embodied in the semantic relations we

propose for our lexical knowledge base, any similarity function based on those relations will automatically benefit from synonymy information without having to consider it separately.

4. The Creation of a Lexical Knowledge Base

We now describe our use of DB methods to automatically create a lexical knowledge base from LDOCE, as proposed for the first component of our framework. Our approach builds on the work of Jensen and Binot (1987) and Montemagni and Vanderwende (1992) and extends it both in quantity and quality of information extracted.

The first step in extracting information from LDOCE entries involves parsing them using our broad-coverage grammar of English. The resulting parse structures are then subjected to a set of heuristic rules whose goal is to identify syntactic and lexical patterns which are consistently associated with some specific semantic relation, such as *instrument* or *location.* Consider, for example, the text of the following two definitions from LDOCE:

> **authority** (n, 7): a person, book, etc., mentioned as the place where one found certain information
> **storehouse** (n, 1): a place or person full of information

In each of these definitions, a *location* relation holds between the headword (in boldface) and the word "information," despite the fact that this relation is expressed differently in each case. The patterns that make it possible to identify the underlying semantic similarity in these superficially different definitions can be roughly paraphrased as:

- if there is a relative clause and the relativizer is in the set {*where, in which, on which},* then create a *location* relation using the verb of the relative clause and its arguments

- if the genus term is in the set {*place area space …}* and there is a PP containing the preposition *of,* then create a *location* relation using the noun of the PP, along with any of its modifiers.

Applying these patterns to the parsed definitions of "authority" and "storehouse" yields, in part, the fact that each is the *location* of "information". As with the Montemagni and Vanderwende work, it is these patterns together with the use of a proficient NL parser that enable the automatic extraction of vast numbers of semantic relations across the entire dictionary. Once extracted, these relations are explicitly added to the words from whose definitions they were obtained, thus creating a network of labeled links between words in the dictionary. Shown below are the semantic relations added to this sense entry for "authority."

> **authority** (n,7): a person, book, etc., mentioned as the place where one found certain information

> | **authority** | *hypernym:* | person |
> | **authority** | *hypernym:* | book |
> | **authority** | *location:* | find |
> | | | *typical_object:* information |

In this example, *hypernym* indicates an *is_a* relationship and the *location* relation has as its value the verb "find" as well as another relation, *typical_object,* whose value is "information." A paraphrase of this set

of relations is "an authority *is a* person, an authority *is a* book, and an authority is a *location* where someone finds an *object,* typically "information."

There are some limitations in this method as described thus far, however. Definitions often fail to express even basic facts about word meanings, facts which we would obviously want to include in a knowledge base to be used for effective EB processing. A typical case in LDOCE is the word "flower", whose definition is perhaps noteworthy more for the information it omits than for what it provides:

flower: the part of a plant, often beautiful and coloured, that produces seeds or fruit

Missing from this definition is any detailed description of the physical structure of flowers, information about what kinds of plants have flowers, and so on. Even the important fact that flowers prototypically have a pleasant scent goes unmentioned. We might, of course, try to increase our stock of information about this word's meaning by exploring the definitions of words used in *its* definition ("plant", "beautiful", etc.), in a way that is similar to the forward spreading activation in the networks of Veronis and Ide mentioned in section 3.0. In fact, such exploration is facilitated in our lexical knowledge base by the explicit relation links we have added for each definition. In this case, however, such a strategy is not especially productive, yielding general information about plants but no specific details about flowers.

As a solution to this problem, we observe that a great deal of additional information about a given word's meaning is typically stored not in the entry for that word itself, but rather in the entries for *other* words that mention that word. For instance, it is relatively unusual to find the words which describe the parts of some object in the lexical entry for that object; instead, the relationship between the words for these parts and the larger object is defined only in the lexical entries describing the components themselves. Consider again the case of *flower,* whose LDOCE entry provides relatively little information about what a flower is. A simple search through LDOCE for noun entries which mention *flower* in their definitions, however, will allow us to arrive at a much more detailed picture of its meaning. For instance, a number of words in LDOCE (listed below) describe flower components (1) , and others establish such facts as what time of year flowers bloom and are plentiful (2), that they prototypically have a pleasant smell (3), that bees collect nectar from them (4), that they can be put in a vase (5), that they are sold from a shop by a florist (6), and that they are rolled up until they open (7). It is further possible to compile an exhaustive list of flowers and flowering plants, a few of which are given in (8) (LDOCE contains scores of such entries).

1. **corolla** "the part of a *flower* formed by the petals , usu. brightly coloured to attract insects"
   petal "any of the (usu. coloured) leaflike divisions of a *flower"*
   stalk "a long narrow part of a plant supporting one or more leaves, fruits, or *flowers;* stem"
   style "the rodlike part inside a *flower* which supports the stigma at the top"

2. **spring**　"the season between winter and summer in which leaves and *flowers* appear"
   summer　"the season between spring and autumn when the sun is hot and there are many *flowers"*

3. **attar**　"a pleasant-smelling oil obtained from *flowers,* esp. roses"
   fragrant "having a sweet or pleasant smell (esp. of *flowers)"*
   perfume $_1$ "a sweet or pleasant smell, as *of flowers"*
   perfume$_2$　"(any of the many kinds of) sweet-smelling liquid, often made from *flowers,* for use esp. on the face, wrists, and upper part of the body of a woman"
   sweet　"having a light pleasant smell, like many garden flowers"

4. **nectar**   "the sweet liquid collected by bees from *flowers*"

5. **vase**   "a container, usu. shaped like a deep pot with a rather narrow opening at the top and usu. made of glass or baked clay, used either to put *flowers* in or as an ornament"

6. **florist**   "a person who keeps a shop for selling *flowers*"

7. **bud**   "a young tightly rolled-up *flower* (or leaf) before it opens"

8. **aconite**   "any of various plants usually having blue or bluish *flowers* and poisonous qualities"
   alyssum "a type of low-growing plant with yellow or white *flowers*"
   anemone "a plant that produces many red, white, or blue *flowers*"
   asphodel "a plant with white, yellow, or pink *flowers*"
   aster   "a garden *flower* with a bright yellow centre"
   azalea "a type of bush with bright usu. strong-smelling *flowers*"

Based on our observation that a wealth of information about a particular word may be contained in the definitions of (and therefore in the semantic relations associated with) words that mention the word in question, we have further augmented our lexical knowledge base to include explicit "backlinks," which provide access to that seemingly hidden information. We believe that such links have the potential to substantially improve the effectiveness of the matching functions proposed in our framework. This is accomplished by dramatically increasing the relational context for a given word in the knowledge base and therefore increasing the likelihood that other words may be successfully matched against that word. We acknowledge, however, that the majority of these backlinks are currently associated with the words in the limited LDOCE defining vocabulary. Although this restricts the coverage of these expanded contexts, we intend to overcome this problem by integrating another full-sized dictionary with an unrestricted defining vocabulary into our system.

5. Current Results for the Creation Process

Rounding numbers of entries to the nearest thousand, of the 75,000 definitions in LDOCE, we currently analyze the 33,000 single word noun definitions and the 12,000 single word verb definitions (45,000 definitions total) in a process that takes about 11 hours on our 486/66 PCs. The exclusion of the 15,000 phrasal entries and other entries for adjectives and adverbs is temporary while we focus on refining the basic methods used by our system. During this processing, we extract some 25 different types of semantic relations, including, for example, *location_of, has_part, is_for, hypernym (is_a), (typical_)subject, (typical_)object,* and *instrument.* The total number of relations currently extracted is over 94,000 (not including certain sub-relations), and these relations are added to the attribute-value structures representing each sense entry in the online version of the dictionary used by our NL parser.

We have hand-checked a random sample of 250 semantic relations across the dictionary and found their overall accuracy to be 78%. Using common statistical techniques we estimate that this rate is representative of the entire dictionary (all 94,000 relations) with a margin of error of +/- 5%. Of note is that just about half of the relations in the sample were of the type *hypernym (is_a),* which were accurate 87% of the time. While this may be seen as inflating the overall accuracy rate, it is counteracted by the currently dismal accuracy of the *part_of* relation (only 15%). Removing the *part_of* numbers from the tallies raised the accuracy of relations other than *hypernym* from 68% to 78%. Our immediate plan for improving the overall accuracy of the relations identified by our system involves computing *hypernym* relations in a first pass through the dictionary, and then using that information to aid in the identification of other relations (such as *part_of)* in a second pass. We also have specific plans to improve our parser

and other aspects of our structural pattern matching, and eventually to disambiguate the word values of relations so that they point to specific senses of words.

We feel that the lexical knowledge base we have created represents a potentially unique and substantial contribution to DB and EB processing. It reflects all of the goals we established for the first component in our proposed framework, containing richly structured semantic information for tens of thousands of words, and from an EB perspective, it represents one of the largest and most deeply processed example bases ever produced for a natural language by automatic means.

6   Using the Lexical Knowledge Base

The lexical knowledge base we created has been integrated into our NL analysis system, although its use is currently limited to determining the correct attachment of prepositional and other phrases through a set of heuristic rules not unlike those described in Jensen and Binot (1987) and Vanderwende (1990). For example, the system correctly attaches the phrase "through natural language" to the verb "communicate" in the following sentence:

> Ultimately, we want to be able to communicate with computers through natural language.

This attachment is achieved through a rule that looks for matches between *typical_objects* of verbs for which "language" is an *instrument* and *typical_objects* of the verb in the sentence. In this case, there are relations in the knowledge base indicating that "language" is an *instrument* for "expressing" "feelings" and that one can also "communicate" "feelings."

These heuristic rules effectively map attachment alternatives to specific semantic relations between the words involved, as described for the fourth component of our framework, and then match those relations against relations in the knowledge base. Although certain kinds of fuzziness are explicitly allowed for during matching, such as in the use of *typical_objects* described above, what is missing from the current implementation is the more general and powerful matching function proposed in the framework. Our present set of rules has already produced some surprisingly good results, but there are still many gaps that will only be filled when we implement that matching function. Of course, experimentation with word sense disambiguation will be made possible by the function's implementation as well.

7.   Conclusion

We have proposed combining DB and EB methods in a framework that offers significant enhancements to NL analysis systems in the areas of phrasal attachment and sense disambiguation. At the center of this framework is a large-scale lexical knowledge base containing over 94,000 semantic relations, which we have created automatically from an online dictionary. We believe that this knowledge base, together with the matching functions proposed in our framework, will provide the benefits anticipated.

8.   Acknowledgements

9. References

Braden-Harder, L. 1992. "Sense Disambiguation using an Online Dictionary." In *Natural Language Processing: the PLNLP Approach,* ed. K. Jensen, G. Heidorn, and S. Richardson. Boston: Kluwer Academic Publishers.

Fillmore, C. and S. Atkins. 1991. "Word Meaning: Starting where the MRDs Stop." Unpublished, invited talk at the *29th Annual Meeting of the Association of Computational Linguistics.*

Furuse, O. and H. Iida. 1992. "An Example-Based Method for Transfer-Driven Machine Translation." *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 139-150.

Grishman, R. and M. Kosaka. 1992. "Combining Rationalist and Empiricist Approaches to Machine Translation." *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 263-274.

Jensen, K., and J.-L. Binot. 1987. "Disambiguating Prepositional Phrase Attachments by using On-line Dictionary Definitions." *Computational Linguistics,* vol 13, no 3-4, pp 251-260.

Jensen, K., G. Heidorn, and S. Richardson. 1992. "Introduction to the PLNLP Approach." In *Natural Language Processing: the PLNLP Approach,* ed. K. Jensen, G. Heidorn, and S. Richardson. Boston: Kluwer Academic Publishers.

Jones, D. 1992. "Non-Hybrid Example-Based Machine Translation Architectures." *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 163-171.

Lesk, M. 1987. "Automatic Sense Disambiguation using Machine Readable Dictionaries: How to Tell a Pine Cone from an Ice Cream Cone." *Proceedings of the 1986 ACM SIGDOC conference,* pp 24-26.

Maruyama, H. and H. Watanabe. 1992. "Tree Cover Search Algorithm for Example-Based Translation." *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 173-184.

McLean, I. 1992. "Example-Based Machine Translation using Connectionist Matching." *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 35-43.

Montemagni, S., and L. Vanderwende. 1992. "Structural Patterns vs. String Patterns for Extracting Semantic Information from Dictionaries." *Proceedings of COLING92,* pp 546-552.

Okumura, A., K. Muraki, and K. Yamabana. 1992. "A Pattern-Learning Based, Hybrid Model for the Syntactic Analysis of Structural Relationships among Japanese Clauses." *Proceedings of the Fourth International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 45-54.

Sadler, V. 1989. *Working with Analogical Semantics: Disambiguation Techniques in DLT.* Dordrecht: Foris Publications.

Sato, S. 1991. "Example-Based Translation Approach." *Proceedings of the International Workshop on Fundamental Research for the Future Generation of Natural Language Processing,* ATR Interpreting Research Laboratories.

Sumita, E., H. Iida, and H. Kohyama. 1990. "Translating with Examples: A New Approach to Machine Translation." *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation,* pp 203-212.

Sumita, E. and H. Iida. 1991. "Experiments and Prospects of Example-Based Machine Translation. Pro*ceedings of the 29th Annual Meeting of the Association of Computational Linguistics,* pp 185-192.

Takeda, K., N. Uramoto, T. Nasukawa, and T. Tsutsumi. 1992. "Shalt 2--a Symmetric Machine Translation System with Conceptual Transfer." *Proceedings of COLING92,* pp 1034-1038.

Tsutsumi, T. 1992. "Word Sense Disambiguation by Examples." In *Natural Language Processing: the PLNLP Approach,* ed. K. Jensen, G. Heidorn. and S. Richardson. Boston: Kluwer Academic Publishers.

Vanderwende, L. 1990. "Using an On-line Dictionary to Disambiguate Verbal Phrase Attachment." *Proceedings of the 2nd IBM conference on NLP,* pp 347-359.

Veronis, J., and N.M. Ide. 1990. "Word Sense Disambiguation with Very Large Neural Networks Extracted from Machine Readable Dictionaries." *Proceedings of COLING90,* pp 289-295.

Watanabe, H. 1992. "A Similarity-driven Transfer System." *Proceedings of COLING92,* pp 770-776.

Wilks, Y. 1972. *Grammar, Meaning, and the Machine Analysis of Language,* pp 86. London: Routledge & Kegan Paul.

Wilks, Y., D. Fass, C. Guo, J. McDonald, T. Plate, and B. Slator. 1992. "Providing Machine Tractable Dictionary Tools." In *Semantics and the Lexicon,* ed. J. Pustejovsky. Boston: Kluwer Academic Publishers.