

Adam Pease, John Li
Articulate Software
[apease|jli]@articulatesoftware.com

Controlled English to Logic Translation

Abstract: The Controlled English to Logic (CELT) system translates a restricted English grammar to expressions in formal logic. The logic statements use terms from a large formal ontology, the Suggested Upper Merged Ontology (SUMO), giving each resulting statement a wealth of deep meaning, similar in kind if not in degree to capturing the meaning a human associates with words in context.

1 INTRODUCTION

It has long been a goal in computer science for users to communicate with computers in human language. Instead of complex menu systems or command languages, people would certainly rather interact with a computer in the same way one would interact with other people. While there have been many advances in creating computer software that handles English, or other languages, understanding is largely limited to search and retrieval. That is, a user enters a phrase, keywords or a question, and the computer attempts to find the best match between the user's input, and some body of stored text. The computer doesn't understand the question, or the content that it is searching. It can't combine information across documents to create a novel answer to a novel question. It can't understand simple logical conclusions that any human could draw from the stored text. If natural language could be translated to logic, the computer would have a form that could be used for inference, and a degree of real understanding.

We have developed a program which takes a restricted form of natural language and automatically translates it to logic. The Controlled English to Logic (CELT) system (Pease&Murray 2003) has a deep understanding of queries and statements. CELT performs syntactic and semantic analysis on English input, and transforms it to logic. CELT allows new information to be added to a knowledge base and be immediately available for queries.

There have been a small number of previous efforts that are similar to our approach, ACE (Fuchs, 1999) being the most notable, and is the work that inspired CELT. However, a key advance in our work is that the output of the language translation system uses terms in an existing large ontology, the Suggested Upper Merged Ontology (SUMO) (Niles&Pease, 2001). Previous efforts have lacked such a connection, and resulted in logical statements in which terms have meaning only to the extent they are used in a series of sentences from the user. By linking statements to SUMO, each term has a wealth of existing definition, much as people have a wealth of meaning and understanding behind each word that is used in communication. One might also question whether this work is truly new, since most of the issues in the semantics of English of concern to CELT have been presented in linguistics research over the years. The challenge is in bringing all this material together in one computational model, and using a single consistent ontology.

The Suggested Upper Merged Ontology (SUMO) is a freely available, formal ontology of about 1000 terms and 4000 definitional statements. It is provided in first order logic, and also translated into

the OWL semantic web language. It is now in its 75th free version; having undergone five years of development, review by a community of hundreds of people, and application in expert reasoning and linguistics. SUMO has been subjected to formal verification with an automated theorem prover. SUMO has been extended with a number of domain ontologies, which are also public, that together number some 20,000 terms and 70,000 axioms. SUMO has also been mapped to the WordNet lexicon (Fellbaum, 1998, Fellbaum, this volume) of over 100,000 noun, verb, adjective and adverb word senses (Niles&Pease, 2003), which not only acts as a check on coverage and completeness, but also provides CELT's lexicon. SUMO and all the associated tools and products are made available at www.ontologyportal.org

Although CELT uses a simplified syntax it does not limit the parts of speech or the number of word senses a word can have, at least to the extent that by using WordNet, CELT's vocabulary grows as WordNet grows. More importantly, CELT is not a domain specific system as with (Allen et al, 1994). It is a completely general language, but one which can be specialized and extended for particular domains along with domain specific vocabulary.

We see an analogy with the PalmPilot and our work. The Apple Newton was an innovative product that attempted to do recognition of unrestricted handwriting, after some training. The Newton was a failure. The problem was simply too hard to be tractable at the time. The system did not correctly interpret handwriting enough of the time to be useful. The PalmPilot took a different approach. It requires handwriting to be one character at a time, in a special alphabet. These restrictions eliminate most of the hard problems in handwriting recognition. A small burden is placed on the user, and in return the user is provided a very useful product. People will change their behavior if the change is relatively small in proportion to the benefit derived. Placing the jobs that are hard for machines and easy for people in the domain of the human user can make an impossible job practical. People have been predicting the arrival of full text understanding ever since the beginning of AI. The realization of this prediction is usually a constant 10 years from the time when the prediction is being made. The time horizon keeps being extended and is unlikely to arrive soon. The best solution may be to simplify the problem in a smart way.

CELT uses a parsing approach that relies on a controlled English input. This means that the user asks queries in a specified grammatical format. This subset of English grammar is still quite extensive and expressive. The advantage of the controlled English is that when the grammar and interpretation rules are restricted, then every sentence in the grammar has a unique interpretation. This eliminates the problems of ambiguity with other approaches that would result in retrieving non-appropriate answers.

Another way to look at language research is that most work has focused on handling all language at a very shallow level of understanding, and advances now work in the direction of gradually increasing the level of understanding while maintaining coverage of all possible utterances. CELT takes the opposite approach of starting from complete understanding, at least to the extent possible in formal logic, of a very restricted subset of English. Research on CELT has focused on increasing the range of understandable sentences, while maintaining the requirement to have complete understanding of the semantics of the sentences handled.

CELT's controlled English is meant to provide a syntax with a lexical semantics, which means that the sense of a sentence depends only upon the sentence, how the words are put together so as to select one out of many possible meanings for each, and not the context within which it is spoken or written, for example, a paragraph, document, conversation, or the gestalt of a perceived situation. This converts the English syntax into that of a formal as opposed to natural language. This approach appropriately partitions application tasks into components which would be hard to handle by machine (extracting the intended sense of a sentence) and hard to handle by a human (quickly and efficiently interpreting extensive text, performing semantically rich Internet searches, etc). The conversion from a fragment of

a natural language (English) to a formal language (controlled English) puts the the converted language fragment in the same class with the mathematical vehicle used for inferencing with the extracted text (formal logic).

2 WORDNET MAPPINGS

We have mapped all of the over 100,000 word senses in WordNet to SUMO, one at a time, by hand. Our original work is described in detail in a previous publication (Niles&Pease 2003). Since the original version, we have modified the mappings to keep up to date with the latest versions of WordNet, which resulted in a port to WordNet 2.0 and then 3.0. We have also revised the mappings to point to more specific terms in the mid-level and domain ontologies that extend SUMO.

Briefly, WordNet is organized as a list of “synsets” or synonym sets. Synsets each have a dictionary definition and a set of links to other synsets. We assigned each synset a link to a particular SUMO term, or in a few case, links to several terms. WordNet is much larger than SUMO so many synsets link to SUMO terms that are more general. For example “to flee”, in the sense of to “run away quickly” is linked to SUMO's **Running**, since there is no more specific term available. The word “valley” however has only one sense in WordNet and it is linked to the equivalent SUMO term of **valley**. Note that these links are not based on the name of the SUMO term, but rather on the meanings of the formal term and the linguistic term. For example, the informal synset “seven seas” links to the SUMO term **Ocean**.

Having a lexicon and an ontology also enforces a certain discipline of clearly separating linguistic and ontological information. Many ontologies, especially those defined in computer languages that are largely limited to taxonomic information, rely heavily on human intuition about natural language term names for their definition. SUMO term names are not strictly part of the definition of any term. Term meanings come solely from the formal axioms on the term. Term names could be replaced with arbitrary unique symbols and still have the same logically defined meanings. By keeping the lexicon separate from the ontology, each can follow principles of organization specific to their goals as independent works. Many lexicons for languages other than English have been created (Vossen&Fellbaum 2002) and many of those have been mutually linked. As a result, SUMO has links to many languages in addition to English, further keeping clear the distinction between formal terms and their correspondence to human languages.

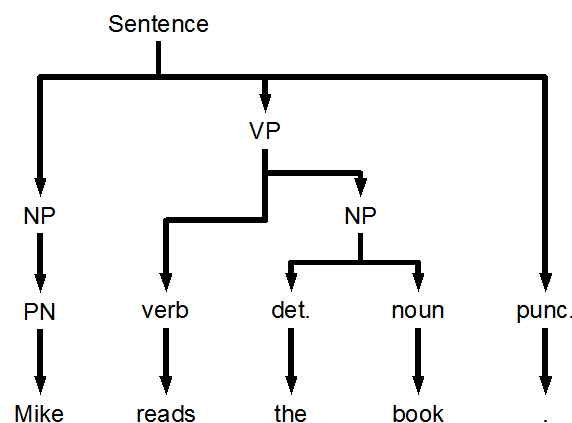


Figure 1: Simple parse tree

3 SIMPLE PARSING AND INTERPRETATION

CELT first parses a sentence to determine the parts of speech for each word. It creates a parse “tree”, such as in Figure 1, that groups these words. For the sentence “Mike reads the book.” there is a proper noun “Mike”, then a verb, determiner, another noun and then a period. The pair of words “the book” forms a noun phrase, noted as “NP”. The phrase “reads the book” is a verb phrase, or “VP”, and so on.

CELT uses the order of the words, and a dictionary that labels words as nouns, verbs and so forth, in order to determine the parts of speech. The dictionary used is

WordNet, augmented with a list of common proper names.

Once CELT has determined the parts of speech of each word, it must also determine the particular sense of each word that is intended. For example, is “reads” the sense of reading text in a book, or the sense of to understand, as in “I read you loud and clear.” Only by determining the proper sense can CELT find the proper related SUMO term to translate to, which would then either be **Reading** or **Communication**, respectively. This is discussed further in section 3.1 below.

Once the CELT parser has determined parts of speech and word senses, it can begin to bring each linguistic element together into a logical expression. SUMO takes a Davidsonian (Davidson, 1980) approach to describing events, in which verbs denoting actions are reified as instances, and the various elements of the event are related to the instance. In this simple example, the subject of the sentence, “Mike” is therefore the agent of the action, and so is brought together with an instance of a “read” event by the SUMO relation **agent**. The object of the sentence is related to the event with the SUMO relation **patient**. A different set of relationships are employed for different linguistic constructions, such as stative verbs, or quantified sentences, and several of these are described further in section 4 below. The value of the resulting logical form is that it can be subjected to deductive inference with a standard first order theorem prover, and yielding new knowledge deduced from known facts and input from the CELT user. This can be contrasted with “statistical” based language understanding approaches that yield approximate and non-logical results for which deductions, if performed, are not truth-preserving.

To elaborate, the SUMO term reading has a number of logically specified truths that are consistent as long as one employs the standard semantics of first order logic (actually SUMO employs some higher-order logical expressions, but let us leave that detail for another paper). One can follow deduction even a hundred steps long and the result will not yield a contradiction. In contrast, although WordNet contains much valuable information, it simply is not intended nor able to support deductions of arbitrary length. For example, a specific sense of “to read” (synset 200626428) does entail the immediate parent (hypernym) synset of “verbalize” (synset 200941990 etc) and its parent “communicate” but that does not necessarily entail that one “interact(s)” (the next parent in turn, which is problematic since one can verbalize to one's self, at least it appears so from the WordNet definition of that synset). While locally consistent, it is not a logical product. Further, nothing in a linguistic product such as WordNet allows one to conclude as does SUMO that **Reading** necessarily involves a **Text** (with a very specific, logical, and formally specified definition of what a **Text** is), that a **Text containsInformation**, and any of the near infinite number of deductions that are possible in turn from each of those small conclusions. This discussion is not intended in any way to criticize WordNet, which is an essential part of CELT, but rather to emphasize the different functions that WordNet and SUMO support in CELT's processing.

3.1 Word Sense Disambiguation

Our method for word sense disambiguation (WSD) is based on a large data set of English sentences that have been annotated by hand to specify the precise sense of each word. The data set used is the Brown Corpus (Kucera&Francis, 1967), and the markup of that corpus is called SemCor (Landis et al, 1998). This work is not new, and in fact is rather poor compared to the state of the art. Its only virtues are that the sources used are free, and the approach is simple. We plan on adopting a more sophisticated approach in the future. We began by processing SemCor to generate a map in which each key is a particular word sense and the value is a list of words and frequencies that indicate how many times the given sense occurred in a sentence with each of a set of words. One weakness of SemCor is that it is not as large as many modern corpora used for this purpose, and key words that discriminate a

particular sense may only occur a few times together with that sense in the Brown Corpus. There are also many senses in WordNet that do not occur at all in that corpus. We used two methods to improve the statistical significance of the entries. In support of our first method, many WordNet senses are very “fine-grained”, which is to say that some senses indicate very small differences in word sense that may not even be listed in most dictionaries. There are often cases where very similar word senses may map to the same SUMO concept. When that occurs, we have “collapsed” the sense entries, and added the co-occurrence data together. The second method used to improve performance over a simple use of SemCor relates to a recent effort by Princeton to disambiguate manually all the words in English definitions of senses in WordNet. We are processing these sentences just as with SemCor to add these statistics to our overall set. A key improvement resulting from this new work is that we will have at least some co-occurrence statistics on all the word senses in WordNet.

An additional method we have implemented is to use the entire history of a CELT dialog in WSD. In a very short sentence such as “Bob runs to the store.” it is very hard to get the sense of the highly polysemous word “run” even with a large manually disambiguated corpus. But, combined with previous sentences in a dialog such as “Bob is training for a marathon.” and “Bob bought new sneakers.” getting the correct result is much more likely.

4 ISSUES IN TRANSLATION

4.1 Case roles and word order

In English, word order is Subject-Verb-Object (except in the case of passive voice, which CELT does not handle). The roles of Subject, Object (and, when present, Indirect Object) have a

```
(exists (?M ?B ?R)
  (and
    (instance ?R Reading)
    (agent ?R Mike-1)
    (instance Mike-1 Human)
    (attribute Mike-1 Male)
    (instance ?B Book)
    (patient ?R ?B)))
```

Figure 2: Case role example

correspondence to SUMO's set of CaseRole(s), which define different kinds of participation in events. Continuing with the preceding example, “Mike” is the subject and “book” is the direct object. In SUMO, the **agent** relation brings together the performer of an action, and the action itself. The **patient** relation specifies a participant in the action. The output from CELT also specifies the types of the participants (Figure 2).

4.2 Statives

The type of verb in a sentence has a critical role in interpretation of meaning. Most verbs indicate occurrence of an event or process. “Mike reads the book.” indicates that a **Reading** event is taking place. Other verbs however connote relationships. “Mike owns the book.” indicates that there is the relationship of ownership between Mike and the book. There is no “owning” event that takes place. Such verbs are called “statives”.

To complicate things, some stative verbs have an appropriate non-Process interpretation. For example "Dickens writes Oliver Twist." should generate (**authors Dickens OliverTwist**). However, once that stative is augmented with a time or a place, then we have two things that are being said, the timeless fact, and the event. For example "Dickens writes Oliver Twist in 1837." makes both the statement of authorship above, and the additional statement about a **Process** occurring in a particular year ().

```
(and
  (authors Dickens OliverTwist)
  (exists (?EV)
    (and
      (instance ?EV Writing)
      (agent ?EV Dickens)
      (equals (YearFn 1837) (WhenFn ?EV))
      (result ?EV OliverTwist))))
```

Figure 3: Stative example

```
(and
  (attribute Bob-1 Male)
  (instance Bob-1 Human)
  (attribute Bob-1 Musician))
```

Figure 4: Stative example

```
(exists
  (?CLNrats ?event)
  (and
    (attribute Robert-1 Male)
    (forall
      (?I)
      (=>
        (member ?I ?CLNrats)
        (instance ?I Rat)))
    (instance Robert-1 Human)
    (instance ?CLNrats Collection)
    (member-count ?CLNrats 5)
    (agent ?event Robert-1)
    (instance ?event Killing)
    (patient ?event ?CLNrats)))
```

Figure 5: Counting example

```
(exists (?boat)
  (and
    (attribute Robert-1 Male)
    (instance Robert-1 Human)
    (instance ?boat Watercraft)
    (orientation Robert-1 ?boat On)))
```

Figure 6: Preposition example

```
(exists (?party ?monday)
  (and
    (instance ?party SocialParty)
    (instance ?monday Monday)
    (during ?party ?monday)))
```

Figure 7: Another preposition example

4.3 Attributes

In the SUMO-WordNet mappings we map most nouns to classes. However, some nouns can be better mapped to attributes that also imply class membership. A noun mapped to a **Position**, should be a **Human** that has the indicated **Position** as an attribute. For example, given the mapping from "pianist" to **Musician**, the sentence "Bob is a pianist."

results in Figure 4

4.4 Counting

CELT handles simple expressions about singulars, plurals, mass nouns and countability. "Bob kills 5 rats." generates Figure 5

4.5 Copula expressions

CELT handles the many different meanings of the verb "to be", which is what linguists call the "copula". "Bob is a pianist." is relatively straightforward, as shown above in the section on attributes. However, copula expressions can also be general statements about classes, as in "An apple is a fruit.", which generates a subclass expression as follows:

```
(subclass Apple Fruit)
```

CELT provides the same translation for "Apples are fruits." since the two are semantically equivalent.

4.6 Prepositions

Prepositions can have significantly different interpretations such as in "Bob is on the boat." (Figure 6). In contrast, the sentence "The party is on Monday." refers to a time, rather than a location, and therefore is translated as in Figure 7. A list of many of the different impacts of argument type on the translations of prepositions is given in Table 1

Preposition	Class	SUMO relation
at,in,on	location	location
at,in,on	time	time
for	person	destination
for, through	time	duration
with	person	agent
with	object	instrument
across	path	traverses
within,into	object	properlyFills
from	object	origin
from	time	BeginFn
through	object	traverses
until	time	EndFn
after	time	greaterThan
before	time	lessThan

Table 1: Prepositions, class membership and relations

CELT type	SUMO class
time	TimeMeasure, Process
person	Human, OccupationalRole, SocialRole
object	all others
mass	Substance, Food, and special words such as money/furniture/data/life/beauty/truth/crime/law/education
count	all others

Figure 8: Finding CELT types with SUMO class membership

```
(exists
  (?event ?hay ?horse)
  (and
    (instance ?horse Horse)
    (instance ?event Eating)
    (instance ?hay Hay)
    (patient ?event ?hay)
    (agent ?event ?horse)))
```

Figure 9: Universal quantification example

```
(forall (?boy)
  (exists (?fudge)
    (and
      (instance ?boy Man)
      (instance ?fudge Food)
      (enjoys ?boy ?fudge))))
```

Figure 10: Universal quantification example

In order to determine the type of elements in the sentence we use SUMO class membership as shown in Figure 8.

4.7 Quantification

Quantification statements are generally those which use words like “every”, “all” or “some”. For example, “Every boy likes fudge.” results in the output shown in Figure 10, and “Some horses eat hay.” results in Figure 9

4.8 Possessives

Possessive statements are of the form “X’s Y”, or, less fluently, “The Y of X.”. The type of the arguments changes the form of the relationship between the entities. “Tom’s father is rich.”, “Bob’s

```
(exists
 (?cookies ?event1 ?event2 ?man ?store)
 (and
  (forall
   (?cookie)
   (=>
    (member ?cookie ?cookies)
    (instance ?cookie Food)))
  (instance ?cookies Group)
  (instance ?event1
   LandTransportation)
  (instance ?event2 Buying)
  (attribute ?man Male)
  (instance ?man Man)
  (instance ?store RetailStore)
  (patient ?event2 ?cookies)
  (agent ?event1 ?man)
  (destination ?event1 ?store)
  (agent ?event2 ?man)))
```

Figure 11: Anaphor example

```
(exists
 (?cat ?house)
 (and
  (attribute Mary-1 Female)
  (attribute Robert-1 Male)
  (instance Mary-1 Human)
  (instance Robert-1 Human)
  (enjoys Mary-1 ?cat)
  (instance ?cat Feline)
  (possesses Robert-1 ?cat)
  (instance ?house House)
  (possesses Robert-1 ?house)))
```

Figure 13: Anaphoric reference for possessives

```
(exists
 (?father)
 (and
  (attribute Tom-1 Male)
  (instance Tom-1 Human)
  (attribute ?father Rich)
  (father ?father Tom-1)))

(exists
 (?nose)
 (and
  (attribute Robert-1 Male)
  (instance Robert-1 Human)
  (attribute ?nose
   SubjectiveAssessmentAttribute)
  (instance ?nose Nose)
  (part ?nose Robert-1)))

(exists
 (?car)
 (and
  (attribute Mary-1 Female)
  (instance Mary-1 Human)
  (attribute ?car Fast)
  (instance ?car Automobile)
  (possesses Mary-1 ?car)))
```

Figure 12: Possessive examples

nose is big.”, “Mary’s car is fast.” See their respective CELT translations in Figure 12. The second example shows a case of where no equivalent term has been created in SUMO for the adjective “big”. This also raises a general issue that some words in natural language are sufficiently vague that it is not possible to state a formal logical meaning for them, at least not

without a more sophisticated approach than relating them to a single formal term.

4.9 Anaphor

CELT handles simple pronoun references, including those across multiple sentences. The user can chose to have CELT process any number of sentences at one time into a single (possibly large) logical expression. It also handles some references on the basis of class descriptions. The example “The man drives to the store. He buys cookies.” yields the translation shown in Figure 11. CELT can also handle possessive pronoun references, such as “Bob has a house. Mary likes his cat.” (Figure 13).


```

(exists
 (?event1 ?event2 ?newspaper ?photo)
 (and
  (attribute Mary-1 Female)
  (attribute Robert-1 Male)
  (instance Mary-1 Human)
  (instance Robert-1 Human)
  (agent ?event1 Mary-1)
  (agent ?event1 Robert-1)
  (instance ?event1 Cutting)
  (agent ?event2 Mary-1)
  (agent ?event2 Robert-1)
  (instance ?event2 Attaching)
  (instance ?newspaper Newspaper)
  (instance ?photo Photograph)
  (patient ?event1 ?newspaper)
  (patient ?event1 ?photo)
  (patient ?event2 ?newspaper)
  (patient ?event2 ?photo)))

```

Figure 14: Conjunction

4.10 Conjunction and Disjunction

Although CELT can handle simple conjunctions, more work is needed. While handling conjunction of predicates, CELT can handle shared subjects and objects. For example, the sentence “Bob and Mary cut and paste the photo and the newspaper.” is translated in Figure 14. Here the assumption that every participant participates in every activity is not necessarily true but CELT takes a simplistic assumption on this issue.

4.11 Negation

Negation is a challenging issue for interpretation. While the logic expressions of two sentences ‘Bob wrote the book.’ and ‘Bob wrote a book.’ are similar, their negative counterparts, ‘Bob did not

write the book.’ (Figure 15) and ‘Bob did not write a book.’ (Figure 16) have quite different logic interpretations. The former sentence assumes context of reference to a particular book, stating that Bob did not write it. The latter states that there exists no book that Bob wrote.

```

(exists (?book)
 (and
  (attribute Robert-1 Male)
  (instance Robert-1 Human)
  (not
   (authors Robert-1 ?book))
  (instance ?book Book)))

```

Figure 15: Negation of a definite reference

```

(not
 (exists (?book)
  (and
   (attribute Robert-1 Male)
   (instance Robert-1 Human)
   (authors Robert-1 ?book)
   (instance ?book Book))))

```

Figure 16: Negation of an indefinite reference

5. CELT COMPONENTS

We use morphological processing rules, derived from the "Morphy" code of WordNet to transform other verb tenses and plural verbs into the various tenses and numbers required. Discourse Representation Theory (DRT) (Kamp & Reyle 1993) handle context to resolve anaphoric references, implications, and conjunctions. CELT is implemented in SWI-Prolog and its grammatical rules are expressed in a Definite Clause Grammar (DCG). The DCG formalism is extended with the feature grammar extension of GULP 3.1 (Covington 1993). Feature grammars specify features such as case, number, and gender. Thus CELT's grammar rules form a unification grammar.

6 ACKNOWLEDGEMENTS

The authors would like to thank the US Army and DARPA for funding the development of CELT, and Army CECOM for funding development of SUMO and the SUMO-WordNet mappings. We would also like to thank the anonymous reviewers of this chapter for their helpful comments, some of which were directly included in the text to improve explanation.

REFERENCES

- Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. H., Kato, T., Light, M., Martin, N. G., Miller, B. W., Poesio, M., and Traum, D. R. 1994 The TRAINS Project: a Case Study in Defining a Conversational Planning Agent. Technical Report. UMI Order Number: TN94-3., University of Rochester.
- Covington, M. (1993) Natural Language Processing for Prolog Programmers. Prentice Hall.
- Davidson, D. (1980) "The Logical Form of Action Sentences." Essays on Actions and Events. New York, NY: Oxford University Press, pp. 105-22. ISBN: 0198246374.
- Fellbaum, C. (ed.) WordNet: An Electronic Lexical Database. MIT Press, 1998.
- Fuchs, N., U. Schwertel, R. Schwitter. (1999). Attempto Controlled English (ACE) Language Manual, Version 3.0, Technical Report 99.03, Department of Computer Science, University of Zurich, August 1999.
- Kamp, H., Reyle, U. (1993). From Discourse to Logic. Kluwer Academic Publishers.
- Kucera and Francis, W.N. (1967). Computational Analysis of Present-Day American English. Providence: Brown University Press.
- Landes S., Leacock C., and Teng, R.I. (1998) "Building semantic concordances". In Fellbaum, C. (ed.) (1998) WordNet: An Electronic Lexical Database. Cambridge (Mass.): The MIT Press.
- Niles, I & Pease A., (2001). "Towards A Standard Upper Ontology." In Proceedings of Formal Ontology in Information Systems (FOIS 2001), October 17-19, Ogunquit, Maine, USA, pp 2-9. See also www.ontologyportal.org
- Niles, I and Pease, A. (2001) "Origins of the IEEE Standard Upper Ontology," Working Notes of the IJCAI-2001 Workshop on the IEEE Standard Upper Ontology, Seattle, pp37-42.
- Niles, I., and Pease, A. (2003) Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology, Proceedings of the IEEE International Conference on Information and Knowledge Engineering, pp 412-416.
- Pease, A., and Murray, W., (2003). An English to Logic Translator for Ontology-based Knowledge Representation Languages. In Proceedings of the 2003 IEEE International Conference on Natural Language Processing and Knowledge Engineering, Beijing, China, pp 777-783.
- Vossen, P., and Fellbaum, C., (editors) (2002) Proceedings of the First International WordNet Conference - GWC 2002, Mysore, India, January 19-25, 2002, Central Institute of Indian Languages