From WordNet to a Knowlege Base

Peter Clark, Phil Harrison, Tom Jenkins, John Thompson, Rick Wojcik
Knowledge Systems
Boeing Phantom Works, Seattle, WA 98124-2207
{peter.e.clark;philip.harrison;john.a.thompson;thomas.l.jenkins2;richard.h.wojcik}@boeing.com

Abstract

At Boeing we have been attempting to use WordNet – an online lexical resource – for machine reasoning, using both its taxonomic information and other information (e.g., parts relations). While we get some leverage, it is clear that WordNet is drastically limited in the types of knowledge it contains. In this paper, we will describe our work in this area, some of our attempts to expand on WordNet, and present a vision for what we believe a future WordNet-like resource should look like – a large knowledge base with vastly richer representational structures – and the potential such a resource would offer for machine reasoning.

Introduction

A fundamental task for both human and machine understanding is the ability to take fragments of data and construct a plausible, coherent scene from which that data could have come. In the context of natural language processing, for example, understanding means going beyond the facts explicitly stated in text to infer additional implications. For instance, consider the newswire sentence:

> "China launched a meteorological satellite into orbit Wednesday, the first of five weather guardians to be sent into the skies before 2008."

This suggests, among other things:

- there was a rocket launch
- China owns the satellite
- the satellite is for monitoring weather
- the orbit is around the Earth

While none of these facts is explicitly stated in the text, we are able to plausibly infer them, in part due to the vast and rich background knowledge we have.

Intelligent machines should be able to draw such plausible conclusions also. In this context, machine understanding can be viewed as *creating a situation-specific representation* coherent with both data and background knowledge. Data suggests which background knowledge to use; conversely, background knowledge suggests ways of interpreting data. Individual inferences may only be plausible rather than deductively true, but the overall "picture" that the computer constructs – the *scenario* – should be coherent, that is, plausible with respect to the expectations that the background knowledge encodes.

Fundamental to this process is the need for a large knowledge-base (KB) of background knowledge in the first place. However, there is little that is freely available at the moment, and so we have been exploring the construction of such a knowledge base using WordNet – an online lexical resource – as a starting point, augmented with both manually and automatically acquired facts. In this paper, we describe the work we have been doing and the limitations we have found both with WordNet and our experimental augmentations to it, and we sketch out a picture of what we believe a future, WordNet-like knowledge resource should look like.

Understanding as "Scenario Completion"

Before describing the knowledge base (KB) itself, we first describe our target application, in order to provide a focus and motivation for the target KB.

Our application goal is to perform retrieval ("smart search") of captioned objects, such as captioned images and video clips. A caption provides a short, incomplete textual description of the captioned object – we will refer to this description as the *scenario*. In this application, machine understanding of a caption involves creating a logic-based representation of the scenario from the caption, and then automatically elaborating it (adding plausible implications) to create a "most coherent" elaboration. Search then involves matching a logical representation of a user's query against the elaborated scene. As the elaborated scene includes information not explicitly in the original caption, the system can potentially retrieve relevant captions even if there are no keywords or synonyms in common between the user's query and the original caption. We have built an illustrative prototype that demonstrates several examples of this with three different databases of captioned objects.

For now, captions are authored in simplified (controlled) English, making the natural language processing task feasible, and, similarly, search queries are expressed as controlled language sentences and translated into logic (Clark *et al.* 2005).

We have found the primary bottleneck in this application

;;; “IF a person is carrying an entity that is inside a room THEN (almost) always the person is in the room.”
$\forall p, c, e, r\ isa(p, Person\_n1) \wedge isa(c, Carry\_v1) \wedge ...agent(c, p) \wedge object(c, e) \wedge is\text{-}inside(e, r) \rightarrow is\text{-}inside(p, r)$

;;; “IF a person is picking an object up THEN (almost) always the person is holding the object.”
$\forall p, u, e, r\ isa(u, Pick\_up_v1) \wedge ...agent(u, p) \wedge object(u, o) \rightarrow \exists h\ isa(h, Hold\_v1) \wedge agent(h, p) \wedge object(h, o)$

;;; “IF an entity is near a 2nd entity AND the 2nd entity contains a 3rd entity THEN usually the 1st entity is near the 3rd entity.”
$\forall x, y, z\ is\text{-}near(x, y) \wedge contains(y, z) \rightarrow is\text{-}near(x, z)$

;;; “ABOUT boxes: usually a box has a lid.”
$\forall b\ isa(b, Box\_n1) \rightarrow \exists l\ isa(l, Lid\_n1) \wedge has\text{-}part(b, l)$

;;; “BEFORE a person gives an object, (almost) always the person possesses the object.”
$\forall p, o, g\ isa(g, Give\_v1) \wedge ...agent(g, p) \wedge object(g, o) \rightarrow precondition(g, posesses(p, o))$

;;; “AFTER a person closes a barrier, (almost) always the barrier is shut.”
$\forall p, c, b\ isa(p, Person\_n1) \wedge ...agent(c, p) \wedge object(c, b) \rightarrow add\text{-}list(c, property(b, Shut\_a1)$

Figure 1: Some example rules in the knowledge base (created using controlled language processing techniques).

to be the lack of background knowledge. As an example, consider the caption and its logical representation from a segment of a video concerning safety procedures:

(1) “A man picks up a large box from a table”
$\exists p, m, b, t\ \ isa(m, Man\_n1) \wedge isa(p, Pick\_up\_v1) \wedge isa(b, Box\_n1) \wedge isa(t, Table\_n1) \wedge agent(p, m) \wedge object(p, b) \wedge size(b, Large\_a1) \wedge origin(p, t)$

From this caption, a person would also realize that, most likely, the man is holding the box; the box was probably on the table; the man is near the table; the man is lifting the box with his hands; the box is now above the table; the man is standing; the scene is probably indoors; etc. In other words, the original caption only gives a partial description of the full scene, and it is our commonsense and domain-specific knowledge of the world that allows us to imagine the “bigger picture”. In this application, we are replicating this process by having the inference engine apply rules to the initial (interpreted) caption in a depth-limited, forward-chaining fashion. The result is a much richer caption description, against which search can be performed. For example, a query for “A person holding something” would match the above caption (1), even though there are no words in common, because the enriched caption description includes the fact that the man is holding the box (inferred from a general rule that picking up something implies holding it), and the system also knows a man is a person, and a table is a thing. Figure 1 gives some of the inference rules from the background knowledge base. We have about 1000 rules in the system, encoded by hand, a drop in the ocean of what would be required for general application of this technology.

The Knowledge Base: Building on WordNet

For the knowledge base itself, we have started with WordNet 2.0 (Miller *et al.* 1993), a broad coverage, highly used, dictionary-like, on-line lexical resource. The core part of WordNet is a large (100,000) ontology of concepts (“synsets”), each denoting a sense of one or more English words. Each concept is denoted by a unique number[1], and defined by the set of words that can refer to it and a textual definition (“gloss”). Concepts are organized into a taxonomy using “more general than” (hypernym) links. In addition, a small number of semantic relations between concepts are listed (parts, causes, entails, ...), currently the “parts” relation being the only one which is extensively specified. WordNet is attractive to use because of its comprehensive coverage, syntactic simplicity, comprehensibility, availability, ease of use, and semantic organization (by concepts, rather than by words – the name “WordNet” is somewhat of a misnomer). It has also become somewhat of a common standard among research groups. However, WordNet also has several well-known limitations, which we will discuss shortly.

Concerning the limited semantic information in WordNet beyond its extensive taxonomic (“isa”) and meronymic (“part-of”) content, we have explored three avenues to add in extra knowledge, as follows:

Hand-coding knowledge. Using language processing techniques, we have been authoring rules in a simplified (controlled) English which are then translated to first-order logic. While we can demonstrate the value of common-sense knowledge in our applications in this way, it is clearly an expensive way to acquire knowledge (although not infeasible).

Information Extraction. We have explored extracting general knowledge following techniques suggested by Schubert (Schubert 2002). For example, from the sentence “The black cat sat on the mat”, we can infer not only something about a specific black cat; we can also infer that:

Cats can be black.
Cats can sit.
Mats can be sat on.

We have extracted 50 million such statements from 1 GB of text. While we believe there is significant potential here, problems of word sense disambiguation and noise remain challenging.

[1]For legibility we have renamed the WordNet synset numbers (e.g., 101789046, denoting the concept of a domestic cat) with more friendly names (e.g., `Cat_n1`).

Knowledge Collection from Volunteers. Some groups have explored acquiring knowledge from volunteers through Web-based acquisition tools (e.g., OpenMind and Learner). We have made some preliminary attempts to make use of both OpenMind and Learner data in our work. As with information extraction, word sense disambiguation and noise remain the major challenges.

From WordNet to a Large-Scale Knowledge Resource

Each release of WordNet contains more information, sometimes adding new relations, and always populating old ones more richly. For example, in WordNet2.1 the instance/class distinction is made, previously conflated in earlier WordNets. One can think of this as a gradual migration from something like a dictionary towards something like a knowledge base. What might WordNet look like in 10 years time? Or more importantly, what *should* a WordNet-like knowledge base look like in 10 years time. We now present some reflections on answers to these questions.

1. Nouns: Recognizing Systematic Polysemy

A frequent criticism of WordNet is that it has "too many" word senses. One goal for a future WordNet-like KB is to use senses (concepts) which are coarser grained than in WordNet. However, a second goal that we envisage is to organize existing senses, especially for nouns, in a systematic way as we now describe.

While nouns often have multiple, related senses (polysemy), Buitelaar has shown that those senses are often not random, but related in a systematic and predictable fashion (Buitelaar 1998). Consider the noun "school" for instance. The senses include[2]:

- `School_n1` an institution;
- `School_n2` a building;
- `School_n3` the process of being educated;
- `School_n4` staff and students;
- `School_n5` a time period of instruction.

In fact, these senses are closely related to each other, and we can sketch out this "pattern" of relationships, shown in Figure 2. Furthermore, special types of school (e.g., academy, college, university), and instances (e.g., the University of Washington), will similarly have five senses, following this pattern; also other types of organization (e.g., government) follow this pattern. In other words, in many cases polysemy is systematic rather than random. We would like to capture that somehow in a KB, making clear what the pattern of polysemy is (i.e., the senses and their relationships), and avoiding having five parallel concept hierarchies for each of the five senses.

We have not designed how this would be encoded in a KB, but it requires somewhat of a significant shift in thinking. For example, the KB might have a single symbol denoting the (underspecified) notion of a school, not placed in any particular taxonomy, plus the associated pattern of

[2] Ignoring the two homonym senses as in "school of fish" and "school of thought"

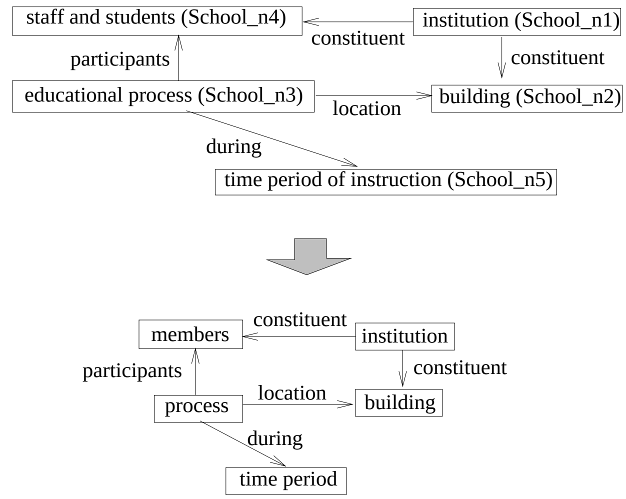

Figure 2: The different senses of "school" are related in a systematic and predictable way, and the pattern of relationships can be abstracted out and made explicit. Ideally, a large-scale KB will contain explicit representations of these patterns, rather than creating multiple parallel taxonomies for school-like concepts which share this pattern of polysemy.

word senses that apply. Behind the scenes this might expand to multiple senses in multiple taxonomies, but the source storage would be a single notion of "school" plus its associated polysemous pattern. During reasoning, the appropriate sense of the word would be selected at run-time, depending on the context in which it was used.

In fact, this is essentially a KR view of what has been proposed in Generative Lexicon theory (Pustejovsky 1991), namely: replacing (or to be more precise, hiding) multiple senses with a single data structure, and then using the context of use to determine which element of that structure is relevant (i.e., using a "smart" inference engine that can select appropriate senses at run-time based on context). We consider this a significant idea, highly relevant to a manageable and useful construction of a large-scale KB.

2. Verbs: Faceted Verb Representations

WordNet also has a prolifation of senses for verbs, typically mixing a core meaning of the verb with its use in a particular setting. For example, the verb "cut" in WordNet2.1 has 41 senses, including "cut hair", "cut grain", and "cut timber". These senses have arisen primarily from WordNet's linguistic (rather than representational) orientation. For example, there is a sense for the meaning "cut grain" because there happens to be a single English word for this ("harvest"), rather than because the axioms defining this action are fundamentally different than for other senses of "cut". Such proliferation is problematic for KR, because the shared meaning of different senses is lost.

Our desire is to reduce the senses of cut to just core meanings. In the case of "cut", this might be just three: physical cut ("cut paper"); collection cut ("cut from the team"); and terminate (“cut the engine”).

In addition, an axiomatization of (physical) “cut” can refer to various “facets” of the action, including:

- the instrument used (knife, scissors, etc.)
- whether the result is fully or partially separated
- the physical cutting action itself

Rather than creating (say) $3 \times 3 = 9$ separate axiom sets, we would prefer to axiomatize each facet independently, and then describe specific uses of cutting in terms of the facets which apply. In this way, we can create specific representations compositionally from general elements, rather than writing each from scratch, again desirable for reuse and maintenance (Clark & Porter 1997).

3. Basic Properties of Objects and Events

WordNet provides a reasonably rich collection of “part-of” links, showing basic properties of objects and their structure. A future knowledge base would include a number of additional such predicates. A few that we have used or would like to use are shown below, with examples:

part-of: An engine is part of a car.
contains: The cranium contains the brain.
connected-to: An airplane wing is connected to the fuselage.
made-of: A car is made of metal.
property: The door is open. The view is beautiful.
normal-shape: A ball is spherical.
normal-size: A book is 8 inches long.
normal-weight: A book weighs 1 pound.
capable-of: A dentist is capable of extracting teeth.
subevent: boiling water is a subevent of making coffee.
enables: Opening a door enables entry into a room.
instruments-needed: A hammer is needed for nailing.
location: Stoves are located in the kitchen.
causes: Splitting an atom causes release of energy.
used-for: A fireplace is used to burn wood.

Other groups have come up with similar lists (e.g., the Univ Texas slot dictionary[3] – some of our examples come from that list). It is interesting to note how similar in content lists from different institutions are.

4. From Rules to Prototypes

Besides the ontology, what form should axioms take? In our experimental knowledge base, we often ended up encoding many rules for “scene completion” for example (just showing the English equivalent of the logic rules):

- “If a person drives a car then [usually] the car is moving”
- “If a car is moving then [usually] a person is driving the car”
- “If a car is moving then [usually] the car is moving on a road”
- “If a car is moving then [usually] the car is using gasoline”
- “If a car is using gasoline then [usually] the car is moving”
- ...etc...

[3]http://www.cs.utexas.edu/users/mfkb/

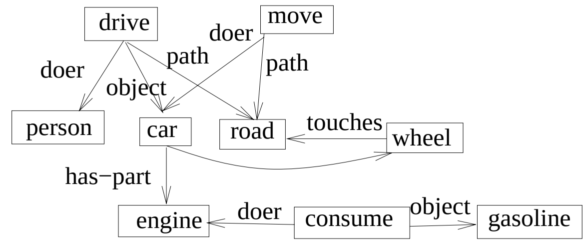

Figure 3: Driving a car can be represented as a single prototype rather than a large set of axioms. Above, nodes denote existentially quantified objects and arcs denote binary predicates.

In fact, when these rules were authored, there was a single scene in mind (of some typical car driving down some typical road), and the aim of the rules was to help complete this scene, i.e., recognize and elaborate it. Unfortunately, this enumeration approach results in a very large number of rules, which are both time-consuming to author and difficult to maintain.

A better way of encoding such knowledge would be to represent such mental scenes explicitly, as a general “typical instance” or prototype of that scene. In this case, the axioms would assert that there exists a car (of some unspecified manufacturer), a road (of some unspecified kind), a person, the person is inside the car, the person is driving the car, etc. A sketch of such a prototype is shown in Figure 3. The resulting reasoning engine for scene completion would then match new data against various prototypes to decide which ones applied. Some features of the new data, e.g., a person doing something, will be common to many prototypes, and thus be only weakly indicative that the prototype applies to that data (has low cue validity). Others will be specific to just a single prototype, providing stronger evidence that that prototype is applicable. We envisage a large-scale knowledge base with many such prototypes. Prototypes essentially encode models of “the way the world might be”, and then reasoning involves search for plausible scene elaborations most consistent with those constraints. Such prototypes may be acquirable automatically or semi-automatically using text extraction techniques, e.g., some modified version of Schubert’s general knowledge extractor from text (Schubert 2002).

FrameNet (Baker, Fillmore, & Lowe 1998) provides something of a start towards this. For any given verb, FrameNet identifies the underlying scene (“Frame”) it is describing; lists and labels the key objects in that scene (“Frame Elements”); and lists the different syntactic patterns of verb usage (alternations) and the different ways the verb’s arguments map onto frame elements. What is missing from FrameNet frames are the axioms describing the scene itself, i.e., the *relationships* among the scene objects. One can view FrameNet’s frames as a specification (only) of a scene, waiting for an axiomatic implementation of the scene to be created.

Information	Materials
Transfer of Possession	Actions & Capabilities
Agents & Organizations	Causality
Changes of State	Containers
Groups	Movement
Part-Whole	Paths and Traversals
Quantities	Structures & Construction
Devices	Spatial
Transformations	Transportation
Agreements	Time and Dates
Substance Composition	Propositional Attitudes
Language	Mathematics
Perception	Financial
Geography	Social Model
Biology/Physiology	Emotions

Table 1: Some top-level theories required for the KB.

5. Core Theories

In addition to the potentially thousands of "mundane" facts about the world, there is also a need for a number of core, general theories, providing a number of important top-level concepts and axioms for reasoning. A back of the envelope list of potential targets is shown in Figure 1. We have made a tentative start on some of these.

As an example, consider the Transfer of Possession model. The process for outlining a model of Transfer of Possession was to study the models in both Cyc and WordNet and to try to organize the ideas into a clean model. It became evident that we would have to tease apart the notion of physical possession from that of legal ownership, even though many of the related concepts in WordNet leave that distinction ambiguous.

For physical possession we identified at least three degrees of physical control: (1) physical possession and control, on one's person; (2) physical control, but not on one's person, such as locked up at home; (3) no physical control.

For legal ownership we also identified three degrees of control: (1) full legal ownership and right of use; (2) temporary right of use (something rented or borrowed); (3) no type of legal possession.

By combining these two dimensions, we cover the important cases of ownership and/or physical possession. For example, if you borrowed someone's book and it's locked in your house, you have physical control level 2 and ownership level 2.

To accommodate these distinctions, we had to reorganize the WordNet ontology of having, owning, giving, getting, and so forth, and we had to insert some new word senses, such as a new sense of "possess" that is strictly physical possession and does not imply any legal right to the object. Then we could sketch out axioms of how certain actions, such as buying, renting, borrowing, and stealing, can change the state of possession in the legal and/or physical dimensions.

6. "Fixing" WordNet

While WordNet is admirably broad, it is missing many key conceptual connections that we have found we need, and it has other problems too. As we are pushing WordNet for tasks that it was never designed for, these should not be viewed as failures of WordNet, but as design criteria for a new WordNet-like KB.

1. Need for the instance/class distinction. Earlier versions used the "hypernym" link to mean both "instance-of" and "is-subclass-of". Fortunately, this distinction has now been made in WordNet2.1.
2. The verb/nominalization dichotomy. From a KR point of view, a verb (e.g., "run") and its nominalization (e.g., "the run") denote the same concept (a running event). However, in WordNet these are considered linguistically different, and as a result most event concepts appear twice (once as a verb sense, once as a noun sense). Worse, WordNet's verb senses and nominalization senses were not authored systematically and so often do not correlate 1-to-1 (typically some nominalizations are missing); in addition, there are no mappings showing which nominalization sense corresponds to which verb sense.

   In our application, we have dealt with this by ignoring nominalizations and working solely with verb senses (e.g., "react" and "reaction" both denote the concept `React_v1`). Philippe Martin has followed a similar course, essentially discarding the verbal part of WordNet in favor of nominalizations in his tidied-up WordNet (Martin 2003). In any large-scale WordNet-style KB, this dichotomy should not exist.
3. Similarly, WordNet does not include key relationships between senses corresponding to morphological variants of a word, e.g., verb/adjective relations ("store"/"stored"), and verb/noun relations ("store"/"storage"). Again, these relationships are critical for commonsense reasoning.
4. In many cases WordNet's hypernym tree is either incorrect (from a set-theoretic, inheritance point of view) or at least dubious (Kaplan & Schubert 2001). For example, in WordNet `Bring_v1` (take something with oneself somewhere) is a type (hyponym) of `Carry_v1` (move while supporting), while conceptually we would consider this to be the other way round (i.e., carry is a type of bring, as one can bring things along by other means, e.g., dragging, wheeling, accompanying, etc.).
5. WordNet's sense definitions ("glosses") are often unclear or vague, making it hard to know exactly what was intended by them. In some cases there is (almost) obvious duplication of senses.
6. Despite WordNet's size, we sometimes have found senses missing. For example, WordNet has only one sense of "animal", while in English there are (at least) two, one which includes people ("people are mammals are animals"), and one which does not ("vets treat animals").
7. WordNet is missing a top-level ontology. Many of the core theories manipulate abstract and non-linguistic concepts (e.g., spatially extended entities), not present in WordNet but that would need to be added in a future KB.

7. Representing Typicality/Uncertainty

Perhaps the biggest challenge from a formal logic viewpoint is that rules/facts in any large-scale common sense KB will only be typical, rather than always true. Or, viewed another way, the semantics of rules are not a simple translation to first-order logic, but more complex. Currently, our hand-written inference rules are flagged with a qualititive degree of typicality (never/sometimes/usually/always), but these are not taken into account in inferencing (beyond the crude approach of ignoring never/sometimes rules). New inferencing methods will be required in any future endeavor to handle the inevitable uncertainty in knowledge that will exist.

8. Syntactic Simplicity

Apart from its broad coverage, a second key property responsible for WordNet's success is its syntactic simplicity. This makes it highly accessible and manipulable by groups interested in using it. Any future knowledge-based equivalent should similarly follow this lesson.

In particular, complex logical axioms can often be summarized in a concise form. In recent years, Cyc has made extensive use of what they call "predicate macros", in which a single ground assertion, e.g.,

```
forallExists(Animal, Head, has-part)
```

expands internally to a full axiom

$$\forall x\ isa(x, Animal) \rightarrow \\ \exists y\ isa(y, Head) \wedge has\text{-}part(x, y)$$

Storing knowledge in macro form like this has many advantages:

- it reduces commitment to a particular axiom syntax, thus improving portability to other environments
- it is easier to read, author, and maintain
- inference engines may take advantage of these forms and work with them directly, removing the need to expand them to their full form and improving efficiency.

A WordNet-like KB should similarly follow this approach.

Concluding Recommendations

In summary, we would like to see progress made on a number of fronts, in an effort to evolve WordNet from a linguistic resource into a useful common-sense knowledge base:

1. When a noun has a group of related word senses (as we saw with "school"), we would like to see them organized according to some set of structured patterns.
2. When a verb has a group of related word senses (as we saw with "cut"), we would like to see them organized into core meanings with various facets.
3. The concepts in WordNet could be richly connected with many useful common-sense relations, such as those that Open Mind has been trying to populate.
4. Typical real-world situations (such as driving a car) could be represented by prototypes that relate all the objects and events involved in each typical situation.
5. A set of upper-level core theories could be developed that would inherit useful deductive axioms to all the lower-level concepts in a WordNet knowledge base.
6. The WordNet ontology should be cleaned up to make it logically correct, and it should all connect with an upper ontology. The glosses should be clarified, and relations should be inserted between related noun, verb, and adjective forms. The word senses should be compared with other dictionaries and all the missing word senses should be added.
7. A knowledge base derived from WordNet should include facilities for representing typical, fuzzy, uncertain, or default knowledge, and mechanisms to reason with this sort of knowledge.
8. Complex axioms could be represented with simple macro predicates that keep the knowledge base syntactically simple to store.

We believe that coordinated efforts that leverage Web volunteers and text mining methods can help to move this effort along at a reasonable pace and without excessive manual labor. We remain optimistic and excited that large-scale, general purpose KBs will become available to the community in the near future.

References

Baker, C. F.; Fillmore, C. J.; and Lowe, J. B. 1998. The Berkeley FrameNet project. In Boitet, C., and Whitelock, P., eds., *Proc 36th ACL Conf.*, 86–90. CA: Kaufmann.

Buitelaar, P. 1998. CoreLex: An ontology of systematic polysemous classes. In Guarino, N., ed., *Proc of the Int Conf on Formal Ontology in Information Systems - FOIS'98 (Frontiers in AI and Applications v46)*, 221–235. Amsterdam: IOS Press.

Clark, P., and Porter, B. 1997. Building concept representations from reusable components. In *AAAI-97*, 369–376. CA: AAAI.

Clark, P.; Harrison, P.; Jenkins, T.; Thompson, J.; and Wojcik, R. 2005. Acquiring and using world knowledge using a restricted subset of english. In *18th International FLAIRS Conference (FLAIRS'05)*.

Kaplan, A., and Schubert, L. 2001. Measuring and improving the quality of world knowledge extracted from WordNet. Technical Report 751, University of Rochester.

Martin, P. 2003. Correction and extension of wordnet 1.7. In *11th International Conference on Conceptual Structures (ICCS'03), to appear*.

Miller, G. A.; Beckwith, R.; Fellbaum, C.; Gross, D.; and Miller, K. 1993. *Five Papers on WordNet*. NJ: Princeton Univ. (http://www.cogsci.princeton.edu/∼wn/).

Pustejovsky, J. 1991. The generative lexicon. *Computational Linguistics* 17(4):409–441.

Schubert, L. 2002. Can we derive general world knowledge from texts? In *Proc. HLT-02*, 84–87.