The Omega Ontology

Andrew Philpot, Eduard Hovy and Patrick Pantel
Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292
{philpot,hovy,pantel}@isi.edu

Abstract

We present the Omega ontology, a large terminological ontology obtained by re-merging WordNet and Mikrokosmos, adding information from various other sources, and subordinating the result to a newly designed feature-oriented upper model. We explain the organizing principles of the representation used for Omega and discuss the methodology used to merge the constituent conceptual hierarchies. We survey a range of auxiliary knowledge sources (including instances, verb frame annotations, and domain-specific sub-ontologies) incorporated into the basic conceptual structure and applications that have benefited from Omega. Omega is available for browsing at http://omega.isi.edu/.

1 Introduction

Omega is a 120,000-node terminological ontology constructed at USC ISI as the synthesis of WordNet 2.0 (Miller 1990; Fellbaum 1998), a lexically oriented network constructed on general cognitive principles, and Mikrokosmos (Mahesh 1996; O'Hara et al. 1998), a conceptual resource originally conceived to support translation, whose result is subordinated under a new upper model, created expressly in order to facilitate the merging of lower models into a functional whole. Omega, like its close predecessor SENSUS (Knight and Luk 1994), can be characterized as a shallow, lexically oriented, term taxonomy—by far the majority of its concepts can be stated in English using a single word. Omega contains no formal concept definitions and only relatively few interconnections (semantic relations) between concepts. By making few commitments to any specific theories of semantics or particular representations, Omega enjoys a malleability that has allowed it to be used in a variety of applications, including question answering and information integration.

2 Constituents of Omega

WordNet, the largest constituent of Omega by size, has a cognitive science orientation, modeling conceptual entities (synonym sets or *synsets*) as the shared meaning of a set of lexical items, but having relatively few kinds of inter-concept relationships. Accordingly, it is richer in the lower part of the network, that houses more concrete concepts named by many lexical items, but unfortunately possesses less structure in the upper part, which contains more general concepts named by fewer or no specific lexical items (e.g., "tangible object", "dispositive material action").

In contrast, Mikrokosmos is a much smaller network, possessing a wider range of conceptual links, but with a lesser focus on lexicalizations of the concepts. Its strength lies in its upper structure and representational expressiveness, not its breadth of coverage.

A major aim in constructing Omega was to leverage the strengths and minimize the weaknesses of the two major constituents: to have a large, lexically rich resource work with a clear comprehensive organization, supporting both inference and lexical access. To support this, we began with a newly designed upper model (Philpot et al. 2003) of about 200 nodes, referred to here as the NUM (New Upper Model). Rooted in a single point, NUM is constructed by successive refinement over a set of mutually exclusive features. Figure 1 illustrates a subset of Omega's upper ontology. Top-level children of the root concept |Summum Genus|[1] are |Object|, |Event|, and |Property|, which are mutually disjoint. Subtypes of |Object| are two sets of mutually disjoint lattice points: {|Tangible-Object|, |Intangible-Object|} and {|Mental-Object|, |Physical-Object|, |Social-Object|}. Children of |Tangible-Object| are the two lattice point sets {|Nonvolitional-Object|, |Volitional-Object|} and {|Biological-Object|, |Nonbiological-Object|}. The leaves of this upper structure are high-level conceptual categories such as |Nonvolitional-Object| and |Intangible-Multi-Participant-Event|. As these are equivalent to conjunctions of (mutually exclusive) features, they provide an excellent place to root lower-level subtrees of concepts. For example, all plant life belongs under |Nonvolitional-Object|, while interpersonal actions without an intrinsic material effect (e.g., "to cooperate") are subtypes of |Intangible-Multi-Participant-Event|.

3 Structure of Omega

Like most ontologies, the heart of Omega is a network of concepts linked by a set of instantiated relationships. Drawing its relationships from both Mikrokosmos and WordNet, and much like that of SENSUS, Omega's concept space is articulated in terms of hierarchical relations such as IS-A, PART-OF, SUBSTANCE-OF, and ELEMENT-OF, as well as lateral ones such as THEME, INSTRUMENT, and PERTAINS-TO. Concepts also possess nonsemantic attributes such as GLOSS.

As in SENSUS, paralleling the concept hierarchy, we collect all lexical items into what can be called language-specific 'lexical spaces'. As with concepts, lexical items may have attributes and lateral links to other entities; in particular, each lexical item contains spelling, morphology, and other orthographic information, and is indirectly, via senses, attached to all concepts it names.

In Omega, rather than directly attaching a lexical item to an appropriate concept as is usually done, we created a sense object that sits between the two and is linked to both. This step permitted us to treat more accurately the information in WordNet, Mikrokosmos, PropBank (Kingsbury et al. 2002), and other resources. We group together all senses for all languages into a single 'sense space' (see Figure 2).

Sense space considerably simplifies resource alignment and the creation of concepts. Over time it has become apparent that builders of larger-scale lexico-semantic resources like dictionaries, WordNet, and PropBank, find it most convenient to work with wordsenses rather than concepts. On the other hand, builders of ontologies and knowledge representation schemes prefer to work with concepts. Lexico-semanticists prefer wordsenses because it is easier to illustrate small shades of difference in word usage with examples than to provide (formalizable) differentiae that adequately distinguish concepts from one another. For this reason, also, the granularity of wordsenses tends to be rather finer-grained than that of concepts; one estimate based on several French and English dictionaries estimates that people define two to three times as many senses as concepts (Cooper 2005), using as criterion for sense vs. concept the ability (or not, respectively) of a new sense for a word to be metaphorically generated from the existing pool of senses.

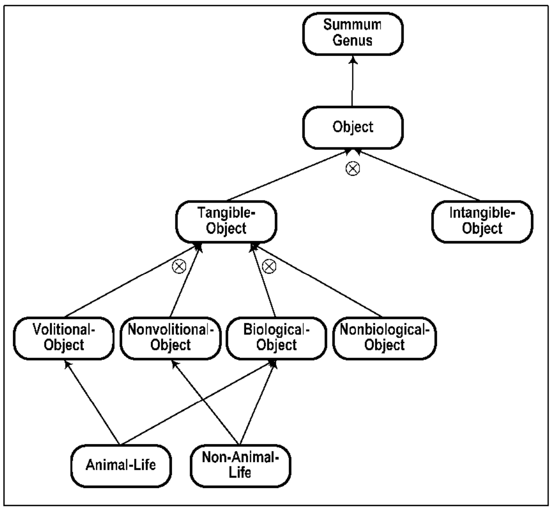

Figure 1. A subset of Omega's upper structure (NUM) with lattice points. Junctions marked with ⊗ indicate points where a superconcept is elaborated by choosing some value from an exhaustive and mutually exclusive feature set.

[1] In this paper, we will notate Omega concepts using |vertical bars|; lexical items will be notated in "quotation marks".

In almost all existing large-scale ontology alignment studies using WordNet, Levin classes, and similar, wordsenses have been viewed as if they were concepts. The mismatches in term granularity and definition style leads to an awkward hybridization if one mixes WordNet-style and Mikrokosmos-style entities, but is resolvable following a process of controlled sense compression going from sense to concept space; see (Hovy 2005).

Sense space also facilitates linkage of words from different languages. In a project to manually annotate the nouns and verbs of texts (in Hindi, Arabic, Korean, French, Spanish, and Japanese, as well as their translations into English) with Omega concepts (Farwell et al. 2004; Reeder et al. 2004), we found it useful to gather the various languages' word senses into a single sense space, where overlaps and differences could be identified, before defining the actual Omega contents. Here the granularity of the concept in question had to be such as to represent the meaning distinctions common across the various translations, while their individual language-idiosyncratic facets of difference could remain in the sense hierarchy.

One can thus think of *ontology space* as the *interlingual* representation symbols (symbols capturing common, or common enough, meaning aspects); of *sense space* as the *multi-lingual* representation symbols (symbols for senses that may or may not co-occur across languages, but that are mapped to meanings no more specific than they denote themselves), and of *lexical space* as the *monolingual* representation symbols (namely, the words of each language).

In general, there is a complex many-to-many mapping across both gaps. Accordingly, we created a sense space as part of Omega to make explicit the nature of the attachments between concepts and lexical items. Besides providing a clean substrate for expressing sense-sense relationships, such as syntactic or morphological derivation links, sense objects are useful for lexical annotations of verb frames (e.g., as in Fleischman et al. 2003a), where one is focusing on making relationships between concepts as expressed in texts. See Figure 2, where three concept nodes (on the left), denoting different concepts which can be expressed using the term "shovel" are linked to two different lexical items,

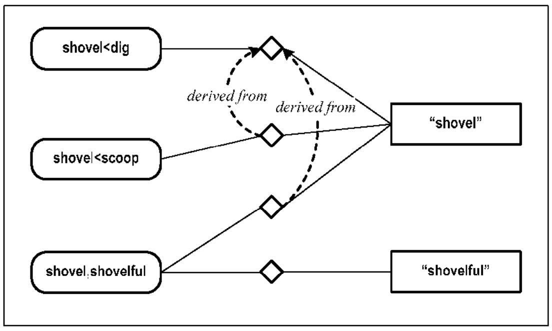

Figure 2. Concept/Sense/Lexical linkage, where senses are unlabeled diamonds, solid lines indicate concept/sense and sense/lexical links, and dashed arrows are sense/sense links.

via four unlabeled sense nodes. Two of these senses, |shovel<scoop| (a garden implement) and |shovel,shovelful| (the amount of material a shovel contains) are derived from a third sense corresponding to the act of shoveling; these sense-sense links are labeled with *derived-from.* Since such links are predicated on the uses of a particular word in context, they belong in the sense space rather than in the concept or lexical item spaces.

In Omega, concepts are interned, i.e., attached to identifying names which are referenced relative to name spaces; additionally, sets of related name spaces (e.g., lexical items expressed in different natural languages) are grouped together into vocabularies. These constructs allow modularity and flexibility in concept and lexical item name orthography.

Names for Omega concepts derived from WordNet synsets are constructed using a local constraint relaxation procedure, described more fully in (Philpot et al. 2003). While use of arbitrary identifiers such as WordNet synset offsets would be possible, having names based on distinguishing characteristics greatly assists users when browsing and reusing the ontology. Human-specified concept names are prohibitively expensive to generate for a network of this size.

First, for each synset, a set of candidate names is generated using attributes of the concept or its neighbors. Some of these generation methods include: word(s) from the associated synset (|rattlesnake|, |port,left|); reference to parent or children concepts (|olive<fruit|, |mob>Mafia|); and suffixing with usage, region or subject domain tags (|tonic(music)|, |class(biology)|). Each name is as-

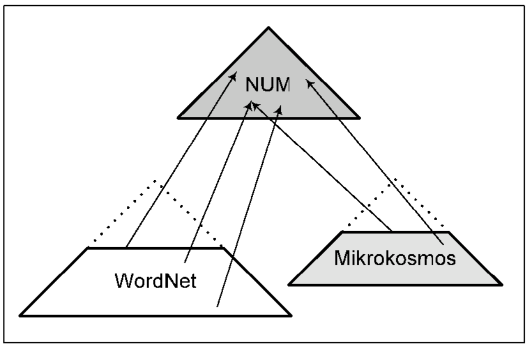

Figure 3. Schematic: Merging WordNet and Mikrokosmos into NUM.

signed a score, based on local metrics such as simplicity and brevity as well as global metrics (primarily avoiding ambiguity). Local constraint cost minimization is used to choose names while approximating the maximum global utility of the assignment.

4 Construction of Omega via Merging

Figure 3 shows our merging framework. We constructed Omega using WordNet (reorganized into concepts (named per the above), senses, and lexical items) and Mikrokosmos (slightly reformulated to satisfy minor orthographic and structural issues); together these are termed the source ontologies. The remainder of this section describes the procedure we used to merge WordNet and Mikrokosmos.

With the NUM in place, the upper models of the source ontologies were first removed and the remaining concepts were linked into the leaf nodes of the NUM. Identification and removal of the upper ontology of the WordNet portion was trivial because it had been previously linked to the Penman upper model (Bateman et al. 1989) that we used in SENSUS and our previous version of Omega. For Mikrokosmos, we considered the top four levels to be the upper model in general. The remnants of the source ontologies then formed a set of isolated sub-lattices of related concepts. The root of each of these was either merged with or made a child of a node (typically a leaf node) of NUM, by manual inspection using the glosses, the local lexical item names, and the feature definition of the NUM concept (e.g., lattice points in Figure 1). The result of this process was one single top region (the NUM) below which hang strands of concepts once linked into the upper ontologies of Mikrokosmos and WordNet. These strands are linked together at the leaf nodes of the NUM and form two "curtains" that hang below, as yet unconnected. On one curtain typically there exists a strand of Mikrokosmos concepts and on the other side exists a strand of WordNet concepts. In the next two phases we sewed up the curtain by first merging the leaf of one side of the curtain into the other, forming a "concept bubble." Then, the bubble was flattened out by merging the interior elements of one side of the curtain into the other. See Figure 4.

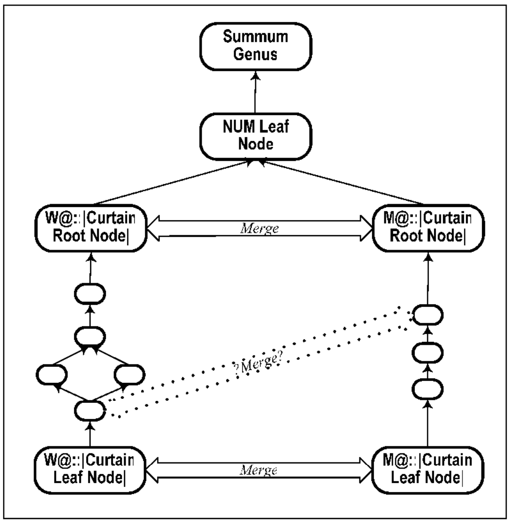

Figure 4. Merging Ontology "Curtains". M@ indicates concepts derived from Mikrokosmos whereas W@ indicates concepts derived from WordNet. Concept nodes interior to the curtains are unlabeled.

At this point, we merged leaves from one curtain (typically the Mikrokosmos curtain) into (possibly leaf) nodes from the other curtain. Then, the resulting bubbles were sewn together. All combinations of pairs of concepts, one from each side of a given bubble, were compared using a learned classifier based on a few hand-aligned examples. The combination which provided the largest number of consistent high-quality matches was presented to a human to accept, reject or edit. The relatively small number of rejected bubbles was retained unmerged. For more details on the merging process, see (Philpot et al. 2003).

5 Omega's Auxiliary Knowledge Sources

The Omega ontology, constructed as outlined above, contains about 120,000 concepts, 156,000 English-language lexical items, 28,000 Spanish-language lexical items, and 270,000 senses. As such, it has shown great utility for research and applications such as information integration and translation. Beyond these, in this section we provide summary descriptions of the various tools we have developed and other knowledge resources we have linked to the ontology core.

5.1 Tools

A web-based browser called Mammoth is available at http://omega.isi.edu/. Mammoth allows interactive visualization of both the current and research versions of Omega, as well as the older ontology SENSUS, also developed at ISI. A command-line interface to Omega, suitable for calling by client programs, is available. Omega is currently implemented using the persistent storage mechanism of the PowerLoom description logic (Chalupsky et al. 2003), which also exposes a relational database view of the concepts, senses, and lexical items.

5.2 Frames

As part of a semantic annotation experiment, we have used Omega as the substrate for merging various available data collections which define the semantic frame structures having the predicates used for annotation: FrameNet, PropBank, and LCS database. Each has a different view and a different coverage, so we integrate all information into the Omega ontology. We currently assign frame information only to verb senses and align frame roles among frames.

FrameNet (Baker et al. 1998) defines semantic frames involving various participants, properties, and other conceptual roles, and for each frame, corresponding words are associated. FrameNet II (as of January 2004) defines 487 distinct frames and 6,743 predicate lexicons (2,300 verbs). In Omega, these frames are represented as a set of 73,000 links between sense objects and interned frame pseudo-concepts.

PropBank (Kingsbury et al. 2002) defines predicate-argument structures on a per-predicate basis, and the core elements of each predicate are simply numbered. PropBank (as of February 2004) covers 3,323 predicate verbs and 4,659 framesets, for a total of 40,000 links in Omega.

LCS (Lexical Conceptual Structures) database (Dorr and Habash 2001) contains hand-tagged structures organized based on Levin's English verb classes and alternations (Levin 1993); it contains 4,452 verbs in 492 classes and corresponds to 73,000 links in Omega.

Additionally, we have reformulated the simple verb frame schemata that are provided within WordNet itself into a similar format (35 frames, 63,000 links).

The frame alignment, and once that is accomplished, the alignment of individual roles within each pair of frames, was first produced automatically, by an algorithm that considered 13 features of frames and the ontology organization, and then manually checked by two humans. More details are provided in (Kwon and Hovy 2005).

5.3 Instances and Mined Knowledge

Omega's implementation contains infrastructure for representing and managing large numbers of concept instances (including a database-backed persistent storage mechanism). Instance sets which have been linked into Omega include named entities mined at ISI (470,000 from Fleischman et al. 2003b; 764,000 from Pantel et al. 2004) and noun-noun compounds (from Pantel: 36,000 terms). Additionally, two geographic gazetteers (GNIS from USGS: 1.9 million points of interest; GNS from NGA: 5.4 million points of interest) have been fully linked into Omega, including part/whole relations, a feature typology, and lexical items for all known place names.

5.4 Concept Annotations

Leveraging Omega's deep relationship to WordNet, we have incorporated other WordNet-based corpora, including the Semcor corpus[2], WordNet Topic Signatures (Agirre et al. 2004) (197 million links), and WordNet Subject Domains (Magnini et al. 2000) (200 concepts, 86,000 links). We have also begun looking at incorporating Extended WordNet (XWN) (Moldovan et al. 2001).

[2] The Semcor corpus is available from http://www.cs.unt.edu/~rada/downloads.html.

5.5 Domain Models and Domain-specific Extensions

In support of various applications, we have automatically linked the results of domain specific ontologies into Omega (Klavans et al. 2002; Hovy et al. 2003a).

6 Applications

Omega has been used to support information integration across databases. In (Hovy et al. 2001), the conceptual hierarchy was extended with a domain specific model describing aspects of energy time series. These aspects expressed the metadata conceptualizations of several information sources containing tens of thousands of different time series, which were thereby linked to the appropriate Omega concepts. Users could browse the ontology to find the time series of interest and computer systems accessing of the time series data could use the feature descriptions of domain model concepts to plan and execute multi-source queries.

A related use of Omega was in supporting a related multi-lingual question answering application called AskCal (Philpot et al. 2002). The user's natural language question was parsed using Omega's lexical items; the question type and other aspects of the parse were dynamically constrained using ontological relationships existing among already understood fragments of the sentence.

Omega has been used as a term repository in two projects that manually construct shallow 'literal' semantic representations for text. The IL-Annot project (Farwell et al. 2004; Reeder et al. 2004), containing six partners, had humans annotate text translated from six languages into English; the symbols for nouns, verbs, and adjectives were taken by specialized annotation interface directly from Omega. A similar, ongoing, project, OntoBank, is collaborating with PropBank and other partners to perform the same type of annotation at a very large scale (Hovy et al. 2003b).Omega has also served more passively as the substrate for integrating hierarchical information harvested from online glossaries (Klavans et al. 2002).

7 Discussion and Future Work

Creating an ontology requires repeated decisions about concept creation and placement. Different decision criteria and methodologies give rise to legitimate but different ontologies (see (Hovy 2005) for a discussion of the five major methodologies in use). Omega's growth is grounded in our desire to avoid committing to any specific semantic theory or representation. That way, we avoid falling into methodology-derived black holes and besides can support more users who can tease out the parts of Omega that suit their tasks. The future of Omega lies then in merging together more ontologies, including upper models such as Dolce (Gangemi et al. 2002) and SUMO (Niles et al. 2001), as well as automatically harvesting and integrating instances, entailments, and other knowledge from the Web, domain documents, video, speech, and other media.

In an effort to automatically update and grow ontologies, many researchers including the authors have proposed algorithms for harvesting shallow semantic resources such as term lists, conceptualizations and semantic relations from text corpora and the Web. However, few have succeeded in automatically incorporating this knowledge into a formal ontology.

The need for machine-assisted ontology construction is stronger than ever. It is increasingly clear that humans cannot manually structure the available knowledge at the same pace as it becomes available. However, addressing the general problem of automatic ontology growing is daunting and over-ambitious.

We have tested some algorithms for automatically harvesting semantic knowledge, such as new term lists, concepts, similarity relations, subclass/superclass relations, and several fine-grained verb semantic relations, and have encountered several challenging issues when deploying these resources in natural language applications since the knowledge is not integrated into any formal knowledge representation.

In response, we have developed a general computational approach for representing a lexical ontology, such as Omega, that enables the automatic integration of certain kinds of shallow semantic resources into the ontology (Pantel 2005). The approach assigns syntactic features to each node in an ontology and then attaches shallow semantic resources by matching on these features. We term *ontologizing* a lexical-semantic resource as the task of sense disambiguating the resource. This problem is different but not orthogonal to word-sense disambiguation. If we could disambiguate large collections of text with high accuracy, then current methods for building lexical-semantic resources could easily be applied to ontologize them by treating each word's senses as separate words. Our method does not require the disambiguation of text. Instead, it relies on the principle of distributional similarity, which links the semantics of words with their syntactic behavior, and the observation that polysemous words that are similar in one sense tend to be dissimilar in their other senses.

8 Conclusions

In this paper, we introduced Omega, ISI's 120,000-node terminological ontology. Omega was constructed by merging WordNet 2.0 and Mikrokosmos into a new upper model, created expressly in order to facilitate the merging of lower models into a functional whole. Several auxiliary knowledge sources (such as FrameNet, PropBank, automatically mined knowledge and concept annotations) have also been integrated.

Omega contains no formal concept definitions and only relatively few interconnections between concepts. By making few commitments to any specific theories of semantics or particular representations, Omega enjoys a malleability that has allowed it to be used in a variety of applications, including question answering and information integration.

The future of Omega lies in harvesting and integrating more knowledge sources such as existing ontologies, like Dolce and SUMO, and new concepts and relations mined from media such as textual documents and the web.

References

Agirre, E. and Lopez de Lacalle Lekuona, O. 2004. Publicly available topic signatures for all WordNet nominal senses. *Proceedings of the 4th International Conference on Languages Resources and Evaluations (LREC).* Lisbon, Portugal.

Baker, C.; Fillmore, C.; and Lowe, J. 1998. The Berkeley FrameNet project. *Proceedings of COLING-ACL.* Montreal, Canada.

Bateman, J.A., Kasper, R.T., Moore, J.D., and Whitney, R.A. 1989. A General Organization of Knowledge for Natural Language Processing: The Penman Upper Model. Unpublished research report, USC/Information Sciences Institute, Marina del Rey, CA.

Chalupsky, H. and Russ, T.A. 2003. The PowerLoom Knowledge Representation & Reasoning System. http://www.isi.edu/isd/LOOM/PowerLoom/.

Cooper, M.C. 2005. A Mathematical Model of Historical Semantics and the Grouping of Word Meanings into Concepts. *Computational Linguistics* 31(2).

Dorr, B. and Habash, N. 2001. Lexical Conceptual Structure Lexicons. In Calzolari et al. ISLE-IST-1999-10647-WP2-WP3, *Survey of Major Approaches Towards Bilingual/Multilingual Lexicons.*

Farwell, D., Helmreich, S., Reed, F., Dorr, B., and Habash, N., Hovy, E., Levin, L., Miller, K., Mitamura, T., Rambow, O., and Siddharthan, A. 2004. Interlingual Annotation of Multilingual Text Corpora. In *Proceedings of the Workshop on Frontiers in Corpus Annotation, NAACL/HLT' 04*, Boston, MA.

Fellbaum, C. (ed.) 1998. *WordNet: An On-line Lexical Database and Some of its Applications.* MIT Press, Cambridge, MA.

Fleischman, M., Kwon, N., and Hovy, E.H. 2003a. Maximum Entropy Models for FrameNet Classification. In *Proceedings of EMNLP-03,* Sapporo, Japan.

Fleischman, M., Hovy, E.H. and Echihabi, A. 2003b. Offline strategies for online question answering: Answering questions before they are asked. *Proceedings of ACL-03.* Sapporo, Japan.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. and Schneider, L. 2002. Sweetening Ontologies with DOLCE. *Proceedings of EKAW.* Siguenza, Spain.

Hovy, E.H., Philpot, A., Ambite, J.L., Arens, Y., Klavans, J., Bourne, W., and Saroz, D. 2001. Data Acquisition and Integration in the DGRC's Energy Data Collection Project. *Proceedings of* *the NSF's dg.o 2001 conference*. Los Angeles, CA.

Hovy, E.H., Philpot, A., Klavans, J.L., Germann, U., Davis, P., and Popper, S. 2003a. Extending Metadata Definitions by Automatically Extracting and Organizing Glossary Definitions. *Proceedings of the NSF's dg.o 2003 conference*. Boston, MA.

Hovy, E.H., Marcus, M., and Weischedel, R. 2003b. OntoBank. Presentation at Darpa PI Meeting. Arden House, Harriman, NY.

Hovy, E.H. 2005. Methodologies for the Reliable Construction of Ontological Knowledge. In F. Dau, M.-L. Mugnier, and G. Stumme (eds), *Conceptual Structures: Common Semantics for Sharing Knowledge. Proceedings of the 13th Annual International Conference on Conceptual Structures (ICCS 2005)*. Kassel, Germany. Springer Lecture Notes in AI 3596, pp 91–106.

Kingsbury, P, Palmer, M., and Marcus, M. 2002. Adding semantic annotation to the Penn TreeBank. *Proceedings of HLT-2002 conference*. San Diego, CA.

Klavans, J.L., Davis, P.T. and Popper, S. 2002. Building Large Ontologies Using Web-Crawling and Glossary Analysis Techniques. *Proceedings of the NSF's dg.o 2002*. Los Angeles, CA.

Knight, K. and Luk, S.K. 1994. Building a Large-Scale Knowledge Base for Machine Translation. *Proceedings of AAAI-94* (Seattle, WA, 1994)

Kwon, N. and E.H. Hovy. 2005. Integrating Semantic Frames from Multiple Sources. Submitted.

Levin, B. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago, IL.

Mahesh, K. 1996. Ontology Development for Machine Translation: Ideology and Methodology. New Mexico State University CRL report MCCS-96-292.

Magnini, B. and Cavaglià, G. 2000. Integrating Subject Field Codes into WordNet, in Gavrilidou, M., Crayannis, G., Markantonatu, S., Piperidis, S. and Stainhaouer, G. (eds.) *Proceedings of Second International Conference on Language Resources and Evaluation (LREC-2000)*. Athens, Greece.

Miller, G. 1990. WordNet: An online lexical database. *International Journal of Lexicography*, 3(4).

Moldovan, D., and Rus, V. 2001. Explaining Answers with Extended WordNet. *Proceedings of ACL*. Toulouse, France.

Niles, I., and Pease, A. 2001. Towards a Standard Upper Ontology. *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001),* Ogunquit, Maine.

O'Hara, T, Mahesh, K. and Nirenburg, S. 1998. Lexical acquisition with WordNet and the Mikrokosmos Ontology. *Proceedings of the COLING/ACL Workshop on Usage of WordNet in Natural Language Processing Systems*. Montreal, Canada.

Pantel, P. 2005. Inducing Ontological Co-occurrence Vectors. *Proceedings of Association for Computational Linguistics conference (ACL)*. Ann Arbor, MI.

Pantel, P., Ravichandran, D., and Hovy, E.H. 2004. Towards Terascale Knowledge Acquisition. *Proceedings of Conference on Computational Linguistics (COLING-04)*. Geneva, Switzerland.

Philpot, A., Ambite, J.L., and Hovy, E.H. 2002. DGRC AskCal: Natural Language Question Answering for Energy Time Series. *Proceedings of the NSF's dg.o 2002 conference*. Los Angeles, CA.

Philpot, A., Hovy, E.H., and Fleischman, M. 2003. Semi-Automatic Construction of a General-Purpose Ontology. *Proceedings of the International Lisp Conference 2003*. New York, NY.

Reeder, F., B. Dorr, D. Farwell, N. Habash, S. Helmreich, E.H. Hovy, L. Levin, T. Mitamura, K. Miller, O. Rambow, A. Siddharthan. 2004. Interlingua Development and Testing through Semantic Annotation of Multilingual Text Corpora. Interlingual Annotation for MT Development. *Proceedings of AMTA conference*. Washington, DC.