Breadth and Depth of Semantic Lexicons

Introduction - Notes on the Workshop

Evelyne Viegas
Computing Research Laboratory
New Mexico State University
Las Cruces, NM 88003
viegas@crl.nmsu.edu

Introduction

Building semantic lexicons is a very time consuming task. Efficient large-scale acquisition and representation of lexical knowledge will be greatly aided by capturing regularities in the lexicon.
   Two main issues present themselves:

a) treatment of lexical ambiguity and

b) lexical rules as a conceptual tool for controlled proliferation of entries.

Whereas the former has been regarded as a topical issue for quite some time, the latter is only now receiving its due attention. This workshop will concentrate on lexical rules as a regulator of breadth and depth of the lexicons. Lexical rules are known under a variety of names, e.g., Leech's (1981) "semantic transfer rules," "lexical implication rules" of Ostler and Atkins (1991) and others. They are also addressed in the framework of such theories as the generative lexicon of Pustejovsky (1995). Such linguistic frameworks as LFG and HPSG have also used the concept, albeit in a different sense and for a different purpose. At the same time, theoretical accounts of the use of lexical rules (such as, for instance, preemption or blocking) are rather too general and underspecified to support actual processing. The workshop will stress issues connected with the practical application of lexical rules: when to apply the rules, how the rules influence system design, how to reexamine and adjust the theoretically posited rules in view of practical needs and evidence. Another central issue for the workshop will be large-scale acquisition of computational-semantic lexicons for practical applications. We are mainly interested in examining the following trade-offs: the coverage vs.the depth of existing semantic lexicons vs. the effort involved in building them.

The workshop is intended for researchers in computational linguistics, artificial intelligence, psycholinguistics or other fields who have been working in lexical semantics and large-scale lexical knowledge acquisition.

Notes on the Workshop

The workshop was organized around three main themes reflecting the topic addressed in the papers we received:

a large scale acquisition of semantic lexicons using a corpus-based approach

b development of micro-theories addressing various topics and phenomena such as, word sense disambiguation - nominal compounds - deverbal adjectives - inflectional morphology

c position papers discussing the status of lexical rules or focusing on the dynamic aspect of knowledge sources and of cognitive processes

Some (though not necessarily all) specific questions suggested for discussion include:

1 What are the different types of lexical rules which should be considered in the building of computational lexicons (inflectional and derivational morphology, verbal diatheses, regular word-sense shifts, other)

2 When should the rules be applied (run-time, load-time, acquisition, other)

3 How to evaluate the cost-efficiency of the acquisition effort against the utility of the resulting lexicons. How could we characterize an NLP system along the dimensions of size, corpus coverage, and depth.

4 Analyses of appropriate types of inheritance for different lexical rules.

5 The use of lexical underspecification (and contextual word-use grounding) as a partial alternative to lexical rules.

In order to facilitate interaction between participants, a pre-workshop mailing list was established. Papers were also available to all participants a month prior to the event in Santa Cruz at **http://crl.nmsu.edu/lex-rule**. This web site will remain available after the workshop for further discussions.

References

Geoffrey Leech (1991) *Semantics*, Cambridge: Cambridge University Press.
Nicholas Ostler and B. T. S. Atkins (1992) Predictable meaning shift: Some linguistic properties of lexical implication rules. In J. Pustejovsky and S. Bergler (eds), *Lexical Semantics and Knowledge Representation*. Berlin: Springer.
James Pustejovsky (1995) *The Generative Lexicon*. Cambridge, MA: MIT Press.

Acknowledgments