Computational Semantics and Pragmatics 2011

Raquel Fernández – ILLC, University of Amsterdam

Homework #1 Due: 22 September 2011

Exercise 1 [let me know in case this assignment conflicts with attending other courses]

Attend the talk by Shalom Lappin on *Probabilistic Semantics for Natural Language* at the ILLC Computational Linguistics Seminar (check the CLS webpage for time and location details at http://www.illc.uva.nl/LaCo/CLS/) and write a report of the talk (min. 1 page, max. 2). The report should be written in proper, coherent English and it should include two parts:

1. an objective summary of the contents of the talk that covers the most important points made by the speaker (broad research area, approach, main claims, results, etc.), in your own words but without including any kind of personal opinions;
2. a critical comment where you do expose your opinion on the work presented. Did you find it interesting? Were the speaker's claims convincing? Was there anything you did not understand?

Exercise 2

The following examples from the data set of the 2007 edition of the Recognising Textual Entailment Challenge show three textual entailment pairs that were labeled as TRUE (✓) by the annotators. What seem to be the causes of the entailment in each case and what background knowledge is required, if any, to recognise that the entailment does indeed hold?

T	H	TE
T_1: In one case five workers at a plating company were asphyxiated by hydrogen cyanide while trying to clean a sludge tank.	H_1: Cyanide fumes killed five workers cleaning one tank.	✓
T_2: Nival was founded in 1996 by Sergey Orlovskiy. In early 2005, the company was bought by Ener1 Group, a Florida-based holdings company, for around US$ 10 million.	H_{2a}: Nival was sold in 2005.	✓
	H_{2b}: Sergey Orlovskiy founded Nival.	✓

Exercise 3

The purpose of this exercise is to get you acquainted with WordNet. You do not need to send me a written answer. We will discuss this in class.

Find out what WordNet is (the Wikipedia entry for WordNet is a good source of information, as is WordNet's own website http://wordnet.princeton.edu). Some of the questions you should find answers to include: What kind of words (part of speech) are included in WordNet? How is WordNet organised: what are *synsets* and what semantic relation are covered? Use WordNet's online interface to search for some examples and navigate through the results obtained for each of them.

Exercise 4

Collect a small corpus of sentences (at most 10) of varying lengths from any newspaper or blog. Using WordNet or any standard dictionary, determine how many senses there are for each of the open-class words (nouns, adjectives, verbs, and adverbs) in each sentence. How many distinct combinations of these senses are there for each sentence? How does this number seem to vary with sentence length?