BM1 Advanced Natural Language Processing
University of Potsdam, Winter 2016/17

Assignment 6: Compositional Semantic Construction

Tatjana Scheffler, Ph.D.

Due: Friday, February 3, 10:00 a.m.

In this final assignment, we are asking you to write a grammar that can compute semantic representations for English sentences. Download Alexander Koller's semcon tool from `https://bitbucket.org/tclup/semcon` and the GeoQuery corpus (geoqueries880) from `https://www.cs.utexas.edu/users/ml/nldata/geoquery.html`. It's a Python2 project that uses NLTK.

Write a semcon grammar that is capable of parsing at least twenty different sentences from the GeoQuery corpus. "Different" means that they should differ in more than just named entities (i.e., "name the rivers in Oregon" and "name the rivers in California" do not count as different). In your submission, tell us the sentences that your grammar can do, and sketch the design principles underlying your grammar.

Observe that you can use the `parse` function to compute a semantic representation using the logic classes from NLTK, and that you can use `query` to run your semantic representation as a query against the GeoQuery database and have the results displayed. Test your grammar on the complete geoqueries880 corpus. Report the proportion of sentences in the corpus for which your grammar finds an analysis, i.e. can parse the sentence and returns a semantic representation on which `query` yields an answer (not necessarily the correct one).

Submission Submit your grammar and a document describing your design principles, which sentences are covered, and your evaluation results.

Extra credit. Come up with automatic or semi-automatic methods for deriving or improving your grammar from the corpus or other resources. Implement more evaluation measures, e.g. what percentages of the semantic representations match the gold standard exactly, or yield the correct results as a database query. Figure out a way to deal with the ambiguity in your grammar when you do the evaluation.

Submit your solutions and code via email to `johannsmeier@uni-potsdam.de`