HW #6

Goals

Through this assignment you will:

- Explore the logic and lambda calculus based representations of natural language semantics.
- Analyze the interaction between semantic attachments and syntactic structure for semantic analysis.
- Gain some further familiarity with NLTK.
- Augment grammars with semantic attachments to produce sentential analyses.

Background

Please review the class slides and readings in the textbook on meaning representation and semantic analysis.

Creating Target Semantic Representations

Based on the examples in the text and class slides, **manually** create target representations in first-order logic for each of the sentences in the test sentence file (see the files section below). These forms constitute the semantic representations that your grammar and parsing process below **should** derive.

NOTE: Your representations should employ the Neo-Davidsonian event-style semantics presented in class. In this format, events (primarily, verbs) typically have the form
exists e.(eventing(e) & eventarg1(e,x1) & eventarg2(e,x2) …)

The format will parallel that of the automatically generated output. For each sentence you analyze, please write two lines:

- The sentence itself
- The first-order logic representation of the sentence

If you want to use "fancy" symbols, please use a PDF file. You may also choose to use the NLTK-style formatting instead; that can be stored in a .txt file if you wish.

NOTE: Some of these sentences may be ambiguous, especially in relation to quantifier scope. Please provide only a single semantic representation for each.

Building a Grammar with Semantic Attachments

Based on the materials above and the format for implementing semantic attachments in NLTK Ch. 10 (Links to an external site.) , create a set of context-free grammar rules augmented with semantic attachments that are adequate to analyze a small set of English natural language sentences and produce appropriate semantic analyses.

Your grammar should be able to produce an appropriate semantic analysis for each sentence in the test sentence file (see the files section below). It should be written using the NLTK **.fcfg** format that can be read in by nltk.data.load().

Sample grammars made be found in NLTK book Chapter 10, the online grammars in **/corpora/nltk/nltk-data/grammars/**, and in the examples referenced above. **NOTE:** The NLTK grammars are provided largely to help with formatting the semantic attachments. Remember that the semantic analysis must employ the Neo-Davidsonian event-style semantics in the class slides. Many of the NLTK grammars do not.

Semantic analysis

Create a program to perform semantic analysis on the test sentences based on your grammar and output the results. Specifically, your program should:

- Load your grammar with semantic attachments.
- Read in test sentences to analyze.
- For each sentence
  - Write the sentence to an output file.
  - Parse the sentence and compositionally create a semantic analysis for it. Use the nltk.parse.FeatureChartParser class (or similar) to perform this process.
  - Extract the semantic representation of the sentence.
  - Write that semantic representation to the output file.

Note: You only need to print a single representation if the sentence is ambiguous.

Programming

Create a program called **hw6_semantics.sh** which performs the parsing and compositional semantic analysis described above invoked as:
hw6_semantics.sh <input_grammar_filename> <input_sentences_filename> <output_filename> where,

- **<input_grammar_filename>** is the name of the file holding the grammar with FOL semantic attachments that you created to implement the rule-to-rule style compositional semantic analysis.
- **<input_sentences_filename>** is the name of the file holding the sentences to parse and perform semantic analysis on.

- **<output_filename>** is the name of the file to write the results of your automatic semantic analysis.

Files

Please adhere to the naming conventions below.

Test and Example Data Files

All files related to this assignment may be found in /dropbox/20-21/571W/hw6/, as below:

- **sentences.txt**: These are the sentences to analyze, first by creating target representation manually and then automatically.
- **semantics_example_grammar.fcfg**: This file provides examples of the semantic attachment format for NLTK and shows the desired Neo-Davidsonian event-style representation.
- **semantics_example_sentences.txt**: Example sentences analyzed by the grammar with semantic attachments above and corresponding to the example output.
- **semantics_example_output.txt**: This file provides an example of the desired output format of semantic analysis.

Submission files

- **hw.tar.gz**: Tarball containing the following:
  - **hw6_target_semantics.{txt|pdf}**: This file should hold your manually created target semantic representations for all of the test sentences.
  - **hw6_semantic_grammar.fcfg**: This file should store the grammar you develop for the “rule-to-rule” semantic analysis model and should cover the test sentences. It should be in the NLTK .fcfg format.
  - **hw6_semantics.sh**: Primary program file to call your input grammar and parse the provided sentences.
  - **hw6_output.txt**: The output file with the results of automatic semantic analysis of the input sentences in **sentences.txt** with your **hw6_semantic_grammar.fcfg**.
- **readme.{txt|pdf}**: Write-up file
  - This file should describe and discuss your work on this assignment. Include problems you came across and how (or if) you were able to solve them, any insights, special features, and what you learned. Give examples if possible. If you were not able to complete parts of the project, discuss what you tried and/or what did not work. This will allow you to receive maximum credit for partial work.