1 Project overview

Your project is to delve more deeply into CCG, by writing a small grammar in it. You will select ten or so sentences in a language of your choice (English is fine) and write a grammar in OpenCCG to accurately parse those sentences.

I heavily, heavily encourage you to work in pairs, with at least one person being a student who is comfortable with computational work.

There are a few ways to approach this project:

- Pick sentences from some naturally-occurring text to analyze, i.e., cover constructions occurring in natural language data.
- Cover the basic constructions of a language.
- Cover a particular set of constructions of interest and analyze those.

Whatever language you choose to work on, you will have to provide an organized overview of your data. Furthermore, you will need to examine previous approaches in the literature in order to figure out how to handle specific linguistic data that you find, i.e., the data that is not obvious (to you) how to handle. You are not required to use someone else’s theory, but you should at least be aware of what’s out there, so that your grammar makes sense.

2 Project details

Your project will have two main components:

- A small grammar fragment covering the sentences in your data, implemented in OpenCCG.
- A write-up describing the grammar fragment & the decisions you made (approx. 5–10 pages)

The write-up should roughly be organized as follows:

1. An organized presentation of the data to be covered.
2. A discussion of your OpenCCG implementation (architecture, technical challenges, etc.)
   - Your grammar should: a) result in no system errors, b) parse all the sentences of interest, and c) not parse related sentences which are ungrammatical.
3. A discussion of the formal and linguistic issues which arose in writing the small grammar.
4. Some broader perspective: an extended discussion of a particularly interesting construction, a discussion of whether and how the grammar could be extended to handle a bigger fragment of the language, etc.

3 Timeline

- **Tuesday, April 14:** partners determined
- **Tuesday, April 21:** sentences/topic selected: You should submit a listing of your data and your general approach, so that I can see whether you have a sufficient amount of (interesting) data.
  - I highly recommend proving to yourself by this time that you can use OpenCCG ... which you can do by starting to develop a grammar.
- **Thursday, May 7 (5:00pm):** final grammar and write-up due