Formal Models of Language: Worksheet 1

Paula Buttery

Easter 2019

The following comprises suggested questions for discussion in Supervision 1 (lectures 1-4). Supervisors understand that it's exam term—be prepared to discuss the ideas if you don't have time to tackle all the questions.

Natural Languages

A few exercises to get you thinking about natural languages...

1. The following natural language sentences are ambiguous. Describe the ambiguities.
   (a) *She fed her cat food.*
   (b) *She saw the man with one eye.*
   (c) *She saw the queen in the garden with the telescope.*

2. The following natural language sentences are difficult to process. Hypothesise what is causing the difficulty.
   (a) *I told the girl the rabbit knew the caterpillar would help her.*
   (b) *The twins the rabbit the girl chased liked laughed.*
   (c) *She shook the bottle containing the potion which had made her grow very tall up.*

   Discuss how we might test your hypotheses.

Formal Languages

Properties of regular and context free languages

1. (reminder question) If $\mathcal{L}_1$ and $\mathcal{L}_2$ are regular languages prove the following are also regular:
   (a) $\mathcal{L}_1 \cup \mathcal{L}_2$
   (b) $\mathcal{L}_1\mathcal{L}_2$
   (c) $\mathcal{L}_1 \cap \mathcal{L}_2$

2. If $\mathcal{L}_1$ is regular and $\mathcal{L}_2$ is context free prove the following is also context free:
   (a) $\mathcal{L}_1 \cap \mathcal{L}_2$

Pumping Lemma for regular and context free languages

1. (reminder question) Use the pumping lemma for regular languages to prove that the following are not regular:
   (a) $\mathcal{L} = \{ab^n cd^n e | n \geq 1\}$
   (b) $\mathcal{L} = \{a^n b^{n+1} | n \geq 1\}$
   (c) $\mathcal{L} = \{ww | w \in \{a, b\}^*\}$

2. Use the pumping lemma for context free languages to prove that the following are not context free:
   (a) $\mathcal{L} = \{a^n b^n c^n | n \geq 1\}$
   (b) $\mathcal{L} = \{a^n b^n c^m | n \leq m\}$

Top-down parsing of context free grammars

1. Write an implementation of the Earley parser that can use the toy grammar from Lecture 3 to parse the sentences below:

   (a) They can fish in rivers.

   (b) They can fish in rivers in December.

   How many parses are there for each sentence?

Don't over-engineer this, you just need to implement the algorithm to build the chart—you don't need write code to print derivation trees etc. The point of this exercise is to help you think through the algorithm. Your supervisor doesn't need to see the code, this is not a tick.

Comparing grammar formalisms

1. Consider the following sentences:

   - *Alice eats cakes.*
   - *The caterpillar gives Alice cakes.*
   - *The cat with a grin disappears.*
   - *Alice paints white roses red.*

   Using the examples in the notes/slides to start you off, complete the following tasks:

   (a) Define a context free grammar that could generate the sentences.

   (b) Draw a dependency parse for the sentences.

   (c) Define a tree adjoining grammar that could generate the sentences.