Natural Language
Practical Classes

Luísa Coheur

2025

P9

Syntax

Image generated by ChatGPT

- **Summary**:
  - Syntax and syntactic analysis
  - NLP tools
- **Operational objectives**:
  - Practice creating grammars and performing syntactic analysis (also with CKY)
  - Play with NLP tools (SpaCy and Stanza)
  - Develop some intuition for relation between syntax and semantics
- **This class needs**: paper, a pen/pencil and computer
- **Class material**: these guidelines, a notebook and a test file

Almost in the eye of the storm

You had barely sat down for breakfast when you saw you had an email from Morcela:

> *Youngster,*
>
> *Refine your syntax skills. Something terrifying is about to happen, but we don't know what.*

Throughout the day, you can't stop thinking about Morcela and the NLP freak. Will Morcela ask for your help again? You check the email several times (and think that you should convince Morcela to use Discord). But nothing happens. You decide that the best thing to do is to study syntax. After all, coincidentally, that's your practical class today. Let's go!

Hands on

1. Consider a grammar with the following nonterminals (all other symbols are terminals):

   $$S,\ NP,\ VP,\ N,\ V,\ Det,\ Conj,\ ConjNP$$

   Also, consider the set of production rules:

   $$\begin{aligned} S &\rightarrow NP\ VP \\ NP &\rightarrow Det\ N \\ NP &\rightarrow NP\ ConjNP \\ NP &\rightarrow \text{“John”} \\ ConjNP &\rightarrow Conj\ NP \\ VP &\rightarrow V\ NP \\ Det &\rightarrow \text{“the”} \\ N &\rightarrow \text{“dog”} \mid \text{“cat”} \mid \text{“garden”} \\ V &\rightarrow \text{“chased”} \\ Conj &\rightarrow \text{“and”} \end{aligned}$$

   Apply the CKY algorithm to:

   - parse the sentence *John chased the dog and the cat.*
   - what can you say about the sentence *John chased the dog*? Does it also belongs to the language generated by that grammar?

2. Run the notebook P9_syntax_tools (Colab).

   Test SpaCy and Stanza with the several sentences. You can also test it with the classic “I saw the man on the hill with the telescope”.

   At the end, check the beautiful HTML produced by SpaCy.

3. Consider the following sentences. Create a simple Context Free Grammar (CFG) that allows their generation. Try to use POS tags related with the syntactic function of each constituent (for instance, NP for noun phrase, and VP for verbal phrase).

   - alex eats soup everyday
   - children like chips
   - cats adore salty soup

4. Show that the sentence “children adore soup” is generated by the grammar by creating a parse tree.

5. Consider the notebook P9_grammars:
   (a) Rewrite grammar1, in the notebook, according with the grammar you just wrote.
   (b) Test your grammar in the notebook (colab) to be sure that it accepts the previous sentences.
   (c) Now move to grammar2. Use the code to generate sentences of the language generated by that grammar. Ah! Ah! Ah! (check why "Ah! Ah! Ah!").

6. Consider grammar2 from the notebook P9_grammars. Check the parse trees.

   Try to associate each one of them to the 5 different meanings of the sentence "I saw the man on the hill with the telescope". This should give you some intuition about the idea that syntax supports semantics.

By the end of the day

No news from Morcela, but you feel that something is in the air, and it is not love. A major literary event with Luís de Camões, Tolkien, Rui Saramago, and José Luís Peixoto was canceled in Figueira da Foz. A concert with the Beatles, Taylor Swift, and Mâneskin, in Coimbra, was also canceled. A football match between Sporting and Manchester City, in Lisbon, was postponed. Even the Sunday dinner at your grandma's was canceled. For a moment, you have some expectations that your tests might be canceled too. But, no, just these major events. However, the authorities give no explanation for it. You cannot stop thinking that it has something to do with Morcela and the NLP freak.