FNLP Tutorial 4

1 The CKY Algorithm

1. Determine the potential parses of the sentence "The women fish with bait" according to the PCFG below, using the CKY algorithm (see Lecture 15). Number the symbols you put in the chart in the order they would be computed. For each chart item, assume that candidate grammar rules are searched in the order they are listed below. Use these numbers to construct backpointers in the chart. For example, when constructing the 6th symbol NP out of a determiner with symbol number 2 and a noun with symbol number 3, you could indicate that by putting '6: NP (2, 3)'. (This is not a widely used custom, but a useful one when doing CKY on paper for the context of this tutorial.)

2. What are the probabilities of the parse trees?

S → NP VP (1)
NP → DT N (0.5)
NP → N N (0.1)
NP → N (0.4)
VP → V PP (1)
PP → P NP (1)
DT → the (1)
P → with (1)
V → fish (1)
N → women (0.5) | fish (0.3) | bait (0.2)

2 Dependency Parsing

Consider the following sentence: "Red squirrels and mice eat tree nuts."

1. This sentence contains an ambiguity: point out that ambiguity by giving paraphrases of the different meanings. Which interpretation is the most salient, and why?

2. Assume that when creating dependency trees we treat conjunction as described here (https://universaldependencies.org/u/dep/conj.html). The assumption that is crucial to the usage of dependency trees (relations between words are asymmetric and binary) is violated when we construct a dependency tree for a sentence with a conjunction: explain why.

3. Provide the dependency tree for the sentence. You can use the following labels to mark the arcs: amod, compound, conj, cc, nsubj, dobj, root. J&M, Figure 14.2, 3rd edition provides explanations per label.

4. Using this way of representing conjunction, could we represent both interpretations discussed in (1) above?

5. If you could change the way conjunction is represented in any way you would like, you would be able to represent the least salient interpretation of the sentence: provide that dependency tree.

3 Compositional Semantics

Consider the following sentence: "Some student ate every apple with a fork."

1. Explain why this sentence is ambiguous.
2. Provide paraphrases that express the different interpretations that are possible due to the ambiguity.
3. Give the FOL representation of those interpretations. Use Davidsonian semantics as introduced in Lecture 21 (March 11^{th}).
4. Finally, provide the lambda calculus derivation of "A student eats a green apple" using the grammar below. Adjectives were not included in the lecture. Which of the following captures the role of "green" best? Use that representation in your derivation.
   (a) $\lambda x.green(x)$
   (b) $\lambda P.\lambda x.(green(x) \wedge P(x))$
   (c) $\lambda P.\lambda Q.\lambda x.(P(x) \wedge green(x) \wedge Q(x))$
5. Adjectives were not included in the lecture because they are not always straightforward. If we treated all adjectives analogously to "green", we could make invalid inferences. Can you think of two adjectives where this would be the case?

S	$\rightarrow$	NP VP	*NP.Sem(VP.Sem)*
NP	$\rightarrow$	Det N′	*Det.Sem(N.Sem)*
VP	$\rightarrow$	Vt NP	*Vt.Sem(NP.Sem)*
N′	$\rightarrow$	A N′	*A.Sem(N′.Sem)*
N′	$\rightarrow$	N	*N.Sem*
N	$\rightarrow$	student	$\lambda x.student(x)$
N	$\rightarrow$	apple	$\lambda x.apple(x)$
Vt	$\rightarrow$	eat	$\lambda R.\lambda z.R(\lambda y.eat(z, y))$
Det	$\rightarrow$	a	$\lambda P.\lambda Q.\exists x.(P(x) \wedge Q(x))$
A	$\rightarrow$	green	???

4 Discourse relations

Read Section 22.1.2 from J&M, 3rd edition. This subsection discusses the Penn Discourse Treebank, and how the corpus provides annotations for text spans that are related according to certain discourse connectives. The main purpose of the corpus is to identify arguments that share a discourse relation. The largest two classes identified are:

- Explicit relations: those connected with expressions from well-defined syntactic classes, among which because, when, since, although, and, or, however, otherwise, then, as a result, for example.
- Implicit relations: those where the relationship can be described by one of the connectives that is considered "explicit", without the connective being present in the sentence.

For the fragments below, (1) identify two arguments that have a discourse relationship, and (2) characterise it as explicit or implicit, while (3) specifying the type and subtype according to Figure 22.3 from Section 22.1.2 from J&M, 3rd edition. If no such relationship can be determined, indicate that. If you do think there is a relationship, but that it cannot be characterised with the discourse connectives, explain your reasoning. For examples of two consecutive sentences, the discourse relationship should be between those sentences (not within one of the two).

1. Ms. Bartlett's previous work, which earned her an international reputation in the non-horticultural art world, often took gardens as its nominal subject. Mayhap this metaphorical connection made the BPC Fine Arts Committee think she had a literal green thumb.
2. Jacobs is an international engineering and construction concern. Total capital investment at the site could be as much as $400 million, according to Intel.
3. The U.S. wants the removal of what it perceives as barriers to investment; Japan denies there are real barriers.
4. Hale Milgrim, 41 years old, senior vice president, marketing at Elecktra Entertainment Inc., was named president of Capitol Records Inc., a unit of this entertainment concern. Mr. Milgrim succeeds David Berman, who resigned last month.
5. It's harder to sell stocks when the sell programs come in because some market makers don't want to take the orders.
6. Marketers themselves are partly to blame: They've increased spending for coupons and other short-term promotions at the expense of image-building advertising. What's more, a flood of new products has given consumers a dizzying choice of brands, many of which are virtually carbon copies of one other.