Egghead: My cousin plays second bass for Chicago.
Jughead: No kidding! My cousin too! Does your cousin play for the Cubs?

Note: In the assigned reading, \( \eta \) reduction on p. 171 is wrong. It should be

\[
(\lambda v. Ev) \xrightarrow{\eta} E
\]

Their version is just a special case of \( \beta \)-conversion.

The Principle of Compositionality

Suppose we have a phrase marker (a.k.a. a “parse tree”). We want to know its internal representation, that is, a formula of predicate calculus that captures its meaning. It’s not clear what sense of “capture” we want here, but at least we’d like to distinguish one possible meaning of a sentence from another.\(^1\) It would be nice if every sentence had a single meaning, but many sentences are of course ambiguous. The term reading is commonly used to mean one meaning a sentence could be construed to have in some context. Which raises an issue we should clarify before it gets us into trouble.

Consider the sentence, “I heard she used to be a good friend of Alice, until Alice ran off with her boyfriend.” (1) Syntax tells us that “she” cannot be Alice (subdepartment: “binding theory”). So what does “she” refer to? We can’t tell without knowing the situation in which (1) was uttered. (Let’s pretend we hear it on the street, although issues aren’t dramatically different if you read it in a novel.) The situation sets up a discourse context in

\(^1\)The most extravagant sense would be that the predicate-calculus formula we want is literally the one used by the human brain when it hears the sentence. (But see §§ pragmatics, below.) This makes sense if there is a language of thought, the language the brain “thinks in,” as Jerry Fodor argued in his 1975 of that title. A serious objection is that we evolved from simpler animals; it seems implausible that animals no smarter than cats or rats have a predicate-calculus engine. Fodor dealt with the objection by toughing it out, and claiming that many mammals have a language of thought; if they had better inference engines, they’d be as smart as we are. Having a language of thought does not imply the possession of a language for communication.
which (so the speaker tries to ensure) a particular woman is salient, and will be understood as the referent of “she.” And who is “I”? Again, the situation tells us, although here it’s simpler: “I” refers to the speaker, the person who uttered the sentence. And what about “her”? We naturally hear it as referring to the same woman as “she,” although it’s possible (by straining) to interpret it as referring to Alice or even a third party.

Which of these possible variables sets up different readings of (1)? This issue has been hashed out by linguists and philosophers over the last fifty years or so, and the consensus is that the variation due to the circumstances of utterance of a sentence, including salient objects, who’s speaking, who the intended hearers are, what the time and place of utterance are, and so forth, should be separated out from semantics. The idea is that the true output of semantics is a function from circumstance of utterance to internal representation. Dealing with those circumstances is someone else’s job, in the department of pragmatics.\(^2\)

It’s less clear whose job it is to disambiguate the word “her.” There’s an overwhelming tendency to tie “she” and “her” together, because, through the use of a powerful language of thought, or in some other way, we favor readings that lead to easily imagined scenarios. But it’s conceivable that the discourse supplies the referent.\(^3\)

An important mechanism for expressing readings is the idea of coindexation. The usual notation uses small integers attached as subscripts to noun phrases. For some reason, linguists often don’t use actual numbers, but write \(i, j\), etc. The idea is that if two noun phrases have the same subscript, we’re in a reading where they refer to the same object. If a subscript is of the form \(\neq i\), it means that the two noun phrases do not refer to the same object (a noncoindexation constraint). So we might write

1. I heard she\(_1\) used to be a good friend of Alice\(_2,\neq 1\) until Alice\(_2\) ran off with her\(_1\) boyfriend.
2. I heard she\(_1\) used to be a good friend of Alice\(_2,\neq 1\) until Alice\(_2\) ran off with her\(_2\) boyfriend.
3. I heard she\(_1\) used to be a good friend of Alice\(_2,\neq 1\) until Alice\(_2\) ran off with her\(_3,\neq 1,2\) boyfriend.

\(^2\)From a Greek word, the same one underlying “practice” and “pragmatism” [[I’m pretty sure]]. A less confusing word could have been chosen; it suggests that semantics is to pragmatics as theory is to practice, a very misleading analogy.
\(^3\)[Cf. “he” and “his” in *Wolf Hall*, by Hilary Mantel.]

2
Obviously, in the subscript “2, ≠ 1” the intention is not to say 2 ≠ 1, but that the referent of 2 ≠ the referent of 1. Anyway, in reading 3, the referent of “her” is supplied by context. That is, the unlikeliness is supplied by semantics; the referent of “her” if one explores the option in spite of its unlikeliness is supplied by pragmatics. In the other readings, the referent of “her” is the same as the referent of “she” or “Alice.” Again, semantics should tell us that 1 is the most likely choice.

With this important distinction between semantics and pragmatics cleared away, we can focus on semantics. We’d like the principle of compositionality to be true:

Each possible meaning of a non-leaf tree is a function of (a) the meanings of the subtrees and (b) what label the tree has. By “label” we mean something like “NP” or “VP,” but it will get more complicated later.

The meaning of a leaf is supplied by a lexicon, at least for open-class words, words of a sort we learn all our lives, like nouns, verbs, and adjectives. Closed-class words are the “logical symbols” of natural language, determiners (“the,” “every”), prepositions (“of,” “in”), and conjunctions (“but,” “although”); like logical symbols, their meanings are often supplied by semantic rules, but they can occur in the lexicon, in various roles.

The representation of phrase markers in CSFP, in sections [[]] is woefully inadequate. [[] Open-class words such as nouns are represented by case classes, meaning to add a new one requires recompiling the program. This representation is fixed in sections [[]], . . .] Whether compositionality is precisely true, whether it needs to be refined to be true, or whether it’s just wrong, clearly something like it is necessary to explain how we can understand sentences we’ve never heard before. We know the meaning of each word, so by induction we can infer the meaning of the entire tree.

The word reading means a way of understanding a sentence. A given string of words often has a spectacular number of alternative readings, and it’s our job to explain how they are possible. Sometimes the semantic ambiguity is due to syntactic ambiguity, when a sentence has more than one “parse.” For example, consider the sentence, “I’ve got three presents I just bought for you to wrap.”

—Maybe that’s why linguists use i and j instead of 1 and 2.
—And that’s the last mention of probability in this lecture.
It has at least the following syntactic structures:

1. \((s \ I’ve \ got \ (NP \ three \ presents \ (s \ I \ just \ bought \ for \ you \ s) \ NP) \) \\
\((s \ to \ wrap \ s) \ s)\)

2. \((s \ I’ve \ got \ (NP \ three \ presents \ \\
(s \ I \ just \ bought \ for \ you \ to \ wrap \ s) NP) s)\)

3. \((s \ I’ve \ got \ (NP \ three \ presents \ (s \ I \ just \ bought \ s) \ NP) \) \\
(for \ you \ to \ wrap) \ s)\)

The third structure generates the reading you might get the first time you see the sentence, although in spoken speech structure 1 would, in English, probably be signaled by a drop in pitch\(^6\) between “you” and “to.” Structure 2 is a bit bizarre; but it’s still viable. Our goal for a while is to think about all the possible readings of a sentence. Structure 2 can’t be rejected on syntactic grounds, because changing a few words while leaving the structure intact makes it perfectly reasonable:

\((s \ I’ve \ got \ (NP \ three \ bonbons \ \\
(s \ I \ just \ bought \ for \ you \ to \ eat \ s) NP) s)\)

In other cases we don’t seem to need a difference in words or syntax to get different readings. Consider a story beginning,

“It’s so easy to get a gun these days. I know a cellist in a string quartet who shot her first violinist last year.”

It could be followed by the following alternative continuations, each of which suggests a different reading for the key clause beginning “. . . who shot . . .,” repeated at the beginning of each candidate continuation:

- “. . . who shot her first violinist last year. But I think she may have shot some violists before that.”
- or “. . . who shot her first violinist last year. They had been having an affair, until she found out he was cheating on her.”

\(^6[My \ ear \ isn’t \ that \ good; \ but \ it \ seems \ like \ a \ difference \ in \ pitch.]]\)
or?”... who shot her first violinist last year. That violinist had stuck it out for a long time.”

In all of these the structure of the clause we’re interested in is

\[(S \ (NP\ who\ NP) \ (VP\ shot\ (NP\ (Det\ her\ Det)\ first\ (N\ violinist\ N)\ NP)\ VP)\ S)\]

The second reading depends on knowing that string quartets typically consist of a first violin, second violin, viola, and cello. Possessive pronouns have quite stretchable meanings, just like possessive nouns and the pronoun “of.” “X’s Y” means “the Y associated with X in a highly context-dependent way.” The phrase “her first violinist” means “the person who plays first violin most of the time in her group”;\(^7\) so “X’s Y” can mean “the person who usually plays role Y with respect to X,” as in “her auto mechanic.”\(^8\)

The first reading of the mystery clause is even more interesting (over and above the fact that it’s the funniest). We can gloss it as “She killed a violinist last year and she had never killed a violinist before that.” There’s a general construction whose semantics works like this:

“NP\(_1\) verb\(_T\) possess\(_1\) ordinal NP\(_2\)”

\[
\exists x (L(verb\(_T\))(S(NP\(_1\)), S(NP\(_2\))))
\land \{ y : Past(L(verb\(_T\))(S(NP\(_1\)), S(NP\(_2\)))) \}
= L(ordinal) - 1
\]

(where verb\(_T\) is a transitive verb, possess is a possessive, and ordinal is a word like "first" or "tenth")

In this messy\(^9\) formula, L(word) is the meaning of the word as stored in the lexicon. S(tree) is the meaning of a subtree. The two subscripted 1s indicate that NP\(_1\) and possess\(_1\) are coindexed.

---

\(^7\)Technically, the morphology — the internal structure and derivation — of “first violinist” is therefore “[first violin]ist.”

\(^8\)It’s a crucial fact about language that we’re always learning it. Even if you didn’t know anything about string quartets, knowing that “X’s Y” is often understood in terms of typical roles, and having just heard that X plays cello in a string quartet, you may form the hypothesis that “first violinist” is a role in a string quartet. Of course, the fact that there are other readings, and that you don’t really know which role we’re talking about, would probably keep you from jumping to this conclusion wholeheartedly.

\(^9\)And yet seemingly never really complex enough
I used the word “construction” advisedly, because many linguists (following the lead of the late Charles Fillmore) study the many constructions, often called “idioms,” that natural language is riddled with. A construction is a snippet of grammar that corresponds to a meaning in a way you might not be able to figure out from the “normal” meanings of the words and the syntax involved.

Another construction, in English, is “to verb one’s way through NP,” as in

1. “The fullback drilled his way through the defensive line.”
2. “I daydreamed way through the court’s ruling.”
3. “I daydreamed my way through the ocean.”
4. “I daydreamed my way through the door.”
5. “I daydreamed my way through the wall.”
6. “I bulldozed my way through the wall.”
7. “I daydreamed my way through my spleen.”
8. “He swam his way through his right to remain silent.”

Obviously, the “NPi verb possi way through NPj” construction is less robust than “NPi possi ordinal NPj≠i”; it’s applicable in fewer situations where it occurs. How is all this knowledge about meanings organized?