ACS Syntax and Semantics of Natural Language

Lecture 2: CCG – Beyond Function Application

Stephen Clark

Natural Language and Information Processing (NLIP) Group

sc609@cam.ac.uk

The Problem of Long-Range Dependencies

- A central problem for a theory of grammar:
  - "elements of sentences which belong together at the level of semantics or interpretation may be separated by unboundedly much intervening material" (Steedman)
- Obvious example in English is the relative clause construction:
  - *a woman whom Warren likes*
  - *a woman whom Dexter thinks that Warren likes*
  - . . .

The Relative Clause Construction

- Relative clause construction:
  - *a woman whom Warren likes*

$$\frac{a\ woman}{NP} \quad \frac{whom}{?} \quad \frac{Warren}{NP} \quad \frac{likes}{(S\backslash NP)/NP}$$

- *whom Warren likes* should be $NP\backslash NP$
- so *whom* should be $(NP\backslash NP)/X$ for some X to be determined

“Non-Constituents” in Combinatory Categorial Grammar

$$\frac{a\ woman}{NP} \quad \frac{whom}{(NP\backslash NP)/X} \quad \frac{Warren}{NP} \quad \frac{likes}{(S\backslash NP)/NP}$$

- Could *Warren likes* be a constituent?
- The coordination test for constituency suggests so:
  - *Warren likes but Dexter detests contemporary dance*
- So what is its type?
  - how about S/NP?
  - in which case the type of *whom* is $(NP\backslash NP)/(S/NP)$

Deriving “Non-Constituents”

$$\frac{a\ woman}{NP} \quad \frac{whom}{(NP\backslash NP)/(S/NP)} \quad \frac{Warren}{NP} \quad \frac{likes}{(S\backslash NP)/NP} \quad \frac{}{\text{NOT ALLOWED}}<$$

- Can’t combine *Warren* and *likes* using application rules
- Need two new rules: type-raising and composition

Type-Raising

$$\frac{a\ woman}{NP} \quad \frac{whom}{(NP\backslash NP)/(S/NP)} \quad \frac{\dfrac{Warren}{NP}}{\overrightarrow{S/(S\backslash NP)}^{\mathbf{T}}} \quad \frac{likes}{(S\backslash NP)/NP}$$

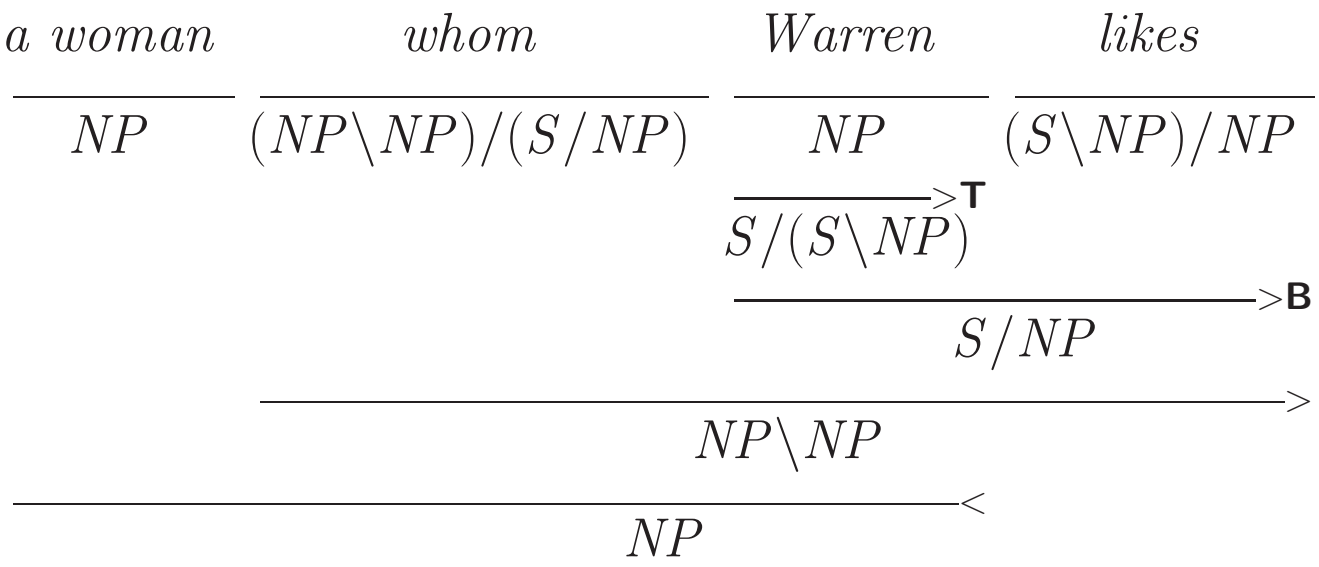
- Subject NP becomes a functional category
- In general: $NP \Rightarrow T/(T\backslash NP)$
  - T is a variable; in practice, for both linguistic and practical parsing reasons, we'd want to limit T to a particular set of types
- Other categories can be type-raised, too, and we can have backward, as opposed to forward, type-raising

Forward Composition

$$\dfrac{\dfrac{a\ woman}{NP}\quad\dfrac{whom}{(NP\backslash NP)/(S/NP)}\quad\dfrac{\dfrac{Warren}{NP}}{\dfrac{}{S/(S\backslash NP)}\,{>}\mathbf{T}}\quad\dfrac{likes}{(S\backslash NP)/NP}}{\dfrac{}{S/NP}\,{>}\mathbf{B}}$$

- Composition allows us to “get inside” a functional category
- In general: $X/Y\ \ Y/Z \Rightarrow X/Z$

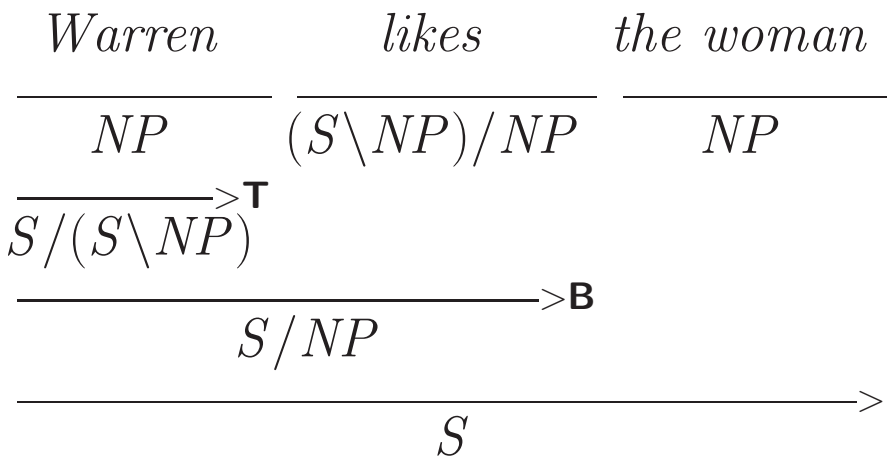
The Complete Derivation for the Object Relative Clause



Type-Raising and Composition

- Used for other syntactic constructions as well, e.g. “non-constituent” coordination (see JHU tutorial slides for the right-node raising example)
- There are other forms of type-raising and composition rules (more on this later)
- *Combinatory* Categorial Grammar is so-called because of the correspondence between CCG’s rules and some of the rules in Combinatory Logic (Curry and Feys)

“Spurious” Ambiguity



- Type-raising and composition can be used to analyse simple sentences with no long-range dependencies
- A different derivation results, *but the interpretation is the same* (hence so-called “spurious ambiguity”)
- In practice we deal with the extra ambiguity by treating it like all other, non-spurious ambiguity (and we still get a highly efficient parser)