Introduction to Natural Language Syntax and Parsing
Lecture 5: Categorial Grammar

Stephen Clark

October 15, 2015

Categorial Grammar (CG) Categorial Grammar is a linguistic theory in the lexicalist tradition, along with other grammar formalisms such as Tree Adjoining Grammar, Lexical Functional Grammar and Head Driven Phrase Structure Grammar, and in contrast to the earlier transformation-based theories of Chomsky. The key idea is that most of the information required for specifying legal syntactic structures within a language resides in the lexicon, and a small number of additional rules specify how to combine such structures.

This lexicon-centered perspective is attractive for a number of reasons. From a theoretical viewpoint, it helps explain how children are able to learn languages: the small number of language-independent combination rules could be innate, and the structures residing in the lexicon which vary across languages can be learnt by exposure to the language. From a practical viewpoint, the rich syntactic structures applying at the word level can be assigned to words using highly efficient sequence labelling methods (often referred to as *supertagging*).

Connection with Semantics The other half of this course deals with compositional semantics, but it's worth pointing out that one of the attractions of (Combinatory) Categorial Grammar — and one of the reasons it is still an active research area in the ACL community — is the close connection to compositional semantics. The tight interface between the syntactic derivations and the underlying semantic (predicate-argument) structure holds the promise of building representations which can be used for "deep" natural language understanding tasks.

Contrast with Phrase-Structure Rules Categorial grammar captures similar information to that contained in traditional phrase-structure rules, but the information is encoded in the *lexical categories* which reside at the leaves of the derivation tree. (Informally, I like to think of the information being "pushed down" a traditional phrase-structure tree onto the complex types at the leaves.)

Categorial grammar is a relatively old linguistic formalism, pre-dating Chomskian linguistics and appearing as early as the 1930s in the work of Polish mathematicians. For the reader interested in theoretical computer science, the notion of grammatical *type* being used here relates to that used in theoretical CS; and the *Combinatory* in Combinatory Categorial Grammar is taken from Curry and Fey's combinatory logic.

Lexical Categories Lexical categories assigned to words represent elementary syntactic structures. The idea is that lexical categories encode the combinatory potential of words to combine with other words, based on their types.

Lexical categories are either atomic or complex. The set of atomic categories is typically small, for example $\{ S, N, NP, PP \}$. Complex categories are built recursively from atomic categories and slashes (forward or backward), where the slash indicates the direction of the argument. The key intuition with a complex category is to think of it as a function. For example, the transitive verb category $(S\backslash NP)/NP$ is to be thought of as a function that requires an NP to the right (its object), an NP to the left (its subject), and that returns a sentence S.

A Simple CG Derivation Since the transitive verb category is a function, it can combine with arguments using function application. The bracketing in the complex category means that it has to combine with (or apply to) the argument NP to its right first — using so-called *forward* application — and then the argument NP to its left — using *backward* application. The intermediate category — $S\backslash NP$ — is the type of a verb phrase in English, and can be thought of as a sentence missing an NP to its left.

Combination Rules in CG Another useful intuition is that, when a complex category applies to its argument, categories effectively "cancel". The example on the previous slide demonstrates this, with the categories in blue cancelling. Earlier work in CG also thought of the combination rules as being akin to multiplication and division; for example, the combination of verb phrase and subject NP is analagous to the arithmetic expression $NP \times \frac{S}{NP} = S$.

Classical Categorial Grammar Early, "classical" variants of categorial grammar only had forward and backward application as the combination rules. In fact, in terms of weak generative capacity, classical categorial grammar is context-free. So why not just use a context-free grammar? The key notion at this stage, and the difference compared to a CFG, is that of lexicalisation − the fact that lexical categories are so rich in terms of the syntactic information they encode (compare the categories we've seen so far with the typical POS tags used for verbs, for example). In the next lecture we'll see Combinatory Categorial Grammar, which has additional combinatory rules and the potential for grammars with greater than context-free power.

Readings for Today's Lecture

- Categorial Grammar, Mark Steedman, 1999. Short encyclopedia entry for MIT Encyclopedia of Cognitive Sciences, R. Wilson and F. Keil (eds.). Available at: http://homepages.inf.ed.ac.uk/steedman/papers.html