Combinatory Categorial Grammar (CCG)

Linguistics 614

Based on Steedman & Baldridge (2011)

Spring 2015
Combinatory Categorial Grammar (CCG)

- Motivation for Categorial Grammar
- Categorial Grammar (CG)
  - Categories
  - Functional Application
  - Linguistic use: Raising and Control
- Combinatory Categorial Grammar (CCG)
  - Functional Composition
  - Type-raising
  - Linguistic use: Coordination and Information Structure
Motivation for Categorial Grammar

Why investigate categorial grammar (CG)?

- **CG is a minimal extension to CFGs.**
  - CG has a close relation to compositional semantics & is well-understood from a logical standpoint.

- **Cross-linguistic generalizations can be made easily since the same set of rules always apply.**

- **Flexible constituency allows for coordination of unlikes and treatment of “unbounded” constructions.**
  - This also supports psychologically plausible analyses, since processing can proceed in a left-to-right fashion.
Categorial Grammar: Categories

- Rules of grammar are entirely conditioned on lexical categories
  - i.e., many categories & a small set of applicable rules
- Categories (or **types**) come in two varieties:
  - **Primitive categories**: N, NP, PP, S, etc.
    
    (1) Marcel := NP
    (2) man := N

  → can further be distinguished by features
- **Functions**: a combination of primitive categories, i.e., a function from one category (primitive or function) to another:
  - S/NP, (S/NP)/(S/NP), etc.
Syntactically potent elements such as verbs are associated with a syntactic category that identifies them as functions.

- Functions specify the type and directionality of their arguments and the type of their result.
- $S\backslash NP$ is an intransitive verb because it is looking for an NP (to the left) in order to form an S.

A “result leftmost” notation is used here:

- $\alpha/\beta$ is a rightward-combining functor over a domain $\beta$ into a range $\alpha$.
- $\alpha\backslash\beta$ is the corresponding leftward-combining functor.
- $\alpha$ and $\beta$ may themselves be functional categories.
Example lexical entries

- Intransitive, transitive, & ditransitive verbs:
  
  (3)  
  a. ran := $(S \backslash NP)$  
  b. proved := $(S \backslash NP)/NP$  
  c. gave := $((S \backslash NP)/NP)/NP$

- Sentence and verb modifying adverbs
  
  (4)  
  a. yesterday := $S \backslash S$  
  b. always := $(S \backslash NP)/(S \backslash NP)$
Rules and derivations

- Functor categories can combine with their arguments by the following rules:

(5) Forward application (>)
\[ X / Y \ Y \Rightarrow X \]

(6) Backward application (<)
\[ Y \ X \ Y \Rightarrow X \]

- Derivations are written as shown below. Note the direct correspondence to an upside-down constituency tree.

```
Marcel  proved  completeness
NP      (S\NP)/NP     NP
       S\NP            >
       S\NP            <
```
Other examples

\[
\begin{align*}
\text{Marcel} & \quad \text{gave} \quad \text{John} \\
\text{NP} & \quad ((S\text{NP})/\text{NP})/\text{NP} & \quad \text{NP} & \quad > \\
& \quad (S\text{NP})/\text{NP} & \quad > \\
& \quad S\text{NP} & \quad > \\
& \quad S \\
\text{NP} & \quad always \\
& \quad (S\text{NP})/(S\text{NP}) & \quad > \\
\text{NP} & \quad ((S\text{NP})/\text{NP})/\text{NP} & \quad > \\
& \quad (S\text{NP})/\text{NP} & \quad > \\
& \quad S\text{NP} & \quad > \\
& \quad S\text{NP} & \quad S \\
& \quad S 
\end{align*}
\]
Adjuncts

The method for combining a head with adjuncts is the same as for the combination with complements: functional application.

- Note how in the previous proof tree, *always* was added with the same forward application as we used to add *chocolate*.
- Adjuncts are generally defined as being of type $X/X$ (or $X/X$).

**Exercise:** What is the derivation of *Marcel proved completeness yesterday*?
Lexical categories are augmented with an explicit identification of their semantic interpretation.

- Rules of functional application are accordingly expanded with an explicit semantics.

\[(7) \quad \text{proved} := (S\setminus NP)/NP : \lambda y \lambda x. \text{prove}'(x, y)\]

\[(8)\]

a. Forward application (\(>)\)
   \[X/Y : f \quad Y : a \Rightarrow X : fa\]

b. Backward application (\(<\))
   \[Y : a \quad X\setminus Y : f \Rightarrow X : fa\]

- Note that the syntactic head is the semantic functor.
Example derivation with semantics

\[
\begin{align*}
\text{Marcel} & \quad \text{proved} \quad \text{completeness} \\
\text{NP: marcel'} & \quad (S \setminus \text{NP}) / \text{NP: } \lambda y \lambda x.\text{prove'}(x, y) \quad \text{NP: completeness'} \\
S \setminus \text{NP: } \lambda x.\text{prove'}(x, \text{completeness'}) & \quad < \\
S : \text{prove'}(\text{marcel'}, \text{completeness'}) & \quad > \\
\end{align*}
\]

**Exercise:** What is the derivation for *Marcel always gave John chocolate*?
Principle of Type Transparency

- The semantic interpretation of all combinatory rules is fully determined by the *Principle of Type Transparency*:
  - **Categories**: All syntactic categories reflect the semantic type of the associated logical form.
  - **Rules**: All syntactic combinatory rules are type-transparent versions of one of a small number of semantic operations over functions including application, composition, & type-raising.

In other words: deriving a sentence in different ways results in the same semantics
Raising and Control

- Raising verb
  
  \[(9) \quad \begin{align*}
  &\text{a. John seemed to be angry.} \\
  &\text{b. seem} := (S\ NP)/(S_{TO}\ NP) : \lambda p \lambda y.\ \text{seem}'(p(y))
  \end{align*}\]

- (Subject-)Control verb
  
  \[(10) \quad \begin{align*}
  &\text{a. John promised to take a bath.} \\
  &\text{b. promise} := \\
  &\quad (S\ NP)/(S_{TO}\ NP) : \lambda p \lambda y.\ promise'(y, p(y))
  \end{align*}\]

- Syntactic functions are identical, but the semantics of control
  - include \(y\) as the interpretation of the subject,
  - which also serves as the subject of the embedded infinitive (the predicate \(p\)).
**Constituent Coordination**

- The coordination *and* can be given the following (schematic) lexical entry:

\[(11) \quad \text{and} := (X/X)/X\]

illustrated by the following constituent coordination example:

<table>
<thead>
<tr>
<th>Marcel</th>
<th>conjectured</th>
<th>and</th>
<th>proved</th>
<th>completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>(S\NP)/NP</td>
<td>(X\X)/X</td>
<td>(S\NP)/NP</td>
<td>((S\NP)/NP)((S\NP)/NP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(S\NP)/NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S\NP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S</td>
</tr>
</tbody>
</table>

*Exercise:* Can we handle Right Node Raising, e.g., *John conjectured and Harry proved completeness?*
Combinatory Categorial Grammar

So far: categorial grammar (CG)

Combinatory Categorial Grammar (CCG) increases the power of CG by adding combinators, e.g., these two combinatory rules:

- Functional composition
- Type-raising
Combinators
Functional Composition (FC)

- To account for coordination of contiguous strings that do not constitute traditional constituents, CCG allows the rule of functional composition:

\[
\begin{align*}
\text{(12)} & \quad \text{a. Forward composition (}>B) \quad X/Y : f \quad Y/Z : g \Rightarrow X/Z : \lambda x. f(gx) \\
& \quad \text{b. Backward composition (<B)} \quad Y\backslash Z : g \quad X\backslash Y : f \Rightarrow X\backslash Z : \lambda x. f(gx)
\end{align*}
\]

Functors can now select for items which are “missing” elements—without needing to change any categories!

- These are sometimes called “harmonic rules” since all the slashes are in the same direction.
- Often, the slashes are \textit{typed}, so that their application can be restricted (to certain so-called modalities).
Composition example

\[
\begin{array}{cccccc}
\text{Marcel} & \text{NP} & \text{conjectured} & (S\backslash NP)/NP & \text{and} & (X\backslash X)/X \\
& & & & \text{might} & (S\backslash NP)/(S\backslash NP) \\
& & & & \text{prove} & (S\backslash NP)/NP \\
& & & & \text{completeness} & \text{NP} \\
\end{array}
\]

\[
\begin{array}{cccccc}
(S\backslash NP)/NP & (X\backslash X)/X & (S\backslash NP)/(S\backslash NP) & (S\backslash NP)/NP & (S\backslash NP)/(S\backslash NP) \\
\quad & \quad & \quad & (S\backslash NP)/NP & \quad \\
\quad & \quad & \quad & \quad & \quad \\
\quad & \quad & \quad & \quad & \quad \\
\quad & \quad & \quad & \quad & \quad \\
\quad & \quad & \quad & \quad & \quad \\
\end{array}
\]

NB: to save space in the future, we can abbreviate $S\backslash NP$ as $VP$, thus making $\text{might}$’s lexical category more intuitive: $VP/VP$

**Exercise:** Without FC, what would the lexical entry for $\text{might}$ need to be?
Without function composition

If we had not used functional composition, we would have had to posit *might*'s lexical entry as:

(13)  \[\text{might} := ((S \backslash NP)/NP)/((S \backslash NP)/NP)\]

- Select a transitive verb, \((S \backslash NP)/NP\)
- Then select the object of that transitive verb, \(NP\)

Consequence: separate rules for every different kind of complement verb (intransitive, ditransitive, sentential complement, etc.)
Exercise: Using function composition

- What are the analyses for *Marcel might prove completeness*?
- Do they correspond to different semantic representations?

Q: Is function composition too powerful?

A1: Dowty (1988) shows that, in terms of word order, no new orders are allowed by function composition:

Application:  
\[
\frac{A/B}{B/C} \quad \frac{C}{B} > \frac{A}{A/B} 
\]

Composition:  
\[
\frac{A/B}{B/C} > \frac{B}{A/C} \quad \frac{C}{A} > \frac{B}{\frac{A}{A/B}} 
\]

Other word orderings simply won’t work for either case.

A2: Modalities can be used to limit the power of function composition where needed.
CCG includes type-raising rules, which turn arguments into functions over functions-over-such-arguments.

(14) Forward type-raising (\(>T\))

\[ X : a \Rightarrow T/(T\setminus X) : \lambda f. fa \]

(15) Backward type-raising (\(<T\))

\[ X : a \Rightarrow T\setminus(T/X) : \lambda f. fa \]

- \(X\) ranges over argument categories (e.g., NP and PP).
- The rules are order-preserving, e.g., (14) can turn an NP into a rightward-looking function over leftward functions, preserving the linear order of subjects and predicates.
Type-raising example

With type-raising, we have more than one option for derivation, where the semantics work out to be the same:

\[
\begin{array}{c}
\text{Marcel} \\
\text{NP: marcel'} \\
\hline
\text{ran} \\
\text{S\NP: run'} \\
\hline
\text{S: run'marcel'}
\end{array}
\]

\[
\begin{array}{c}
\text{Marcel} \\
\text{NP: marcel'} \\
\hline
\text{ran} \\
\text{S\NP: run'} \\
\hline
\text{S: run'marcel'}
\end{array}
\]

- Allowing any category to type-raise allows for an infinite number of derivations
  - One has to be sure to restrict it while parsing.
- Type-raising creates sometimes spurious structural ambiguity
  - But it is extremely useful . . .
Non-standard surface structures

Exercise: Work out a derivation for right node raising:

(16) John conjectured and Harry proved completeness

Hint: type-raise the subjects!
Non-standard surface structures

- Complement-taking verbs like *think*, VP/S, can compose with fragments like *Marcel proved*, S/NP, which accounts for right-node raising (17), and also provides the basis for an analysis of unbounded dependencies (18).

  (17)  
  
  [John disproved]$_{S/NP}$ and [you think that Marcel proved]$_{S/NP}$ completeness.

  (18)  
  
  the result that [you think that Marcel proved]$_{S/NP}$

- **Key:** Strings such as *you think that Marcel proved* are taken to be surface constituents of type S/NP.
- i.e., constituents can be missing material
Some phenomena

Right Node Raising

By using type-raising and functional composition, we can account for such phenomena as right node raising:

\[
\begin{align*}
\text{Marcel} & \quad \frac{\text{NP}}{\text{S}/(\text{S}\backslash \text{NP})} \quad >^T \\
\text{proved} & \quad \frac{(\text{S}\backslash \text{NP})/\text{NP}}{\text{S}/\text{NP}} \quad >^B \\
\text{and} & \quad \frac{(X\backslash X)/X}{\text{I}/\text{NP}} \quad >^T \\
\text{I} & \quad \frac{(\text{S}\backslash \text{NP})/\text{NP}}{\text{S}/\text{NP}} \quad >^B \\
\text{disproved} & \quad \frac{(\text{S}\backslash \text{NP})/\text{NP}}{(\text{S}/\text{NP})/(\text{S}/\text{NP})} \quad > \\
\text{completeness} & \quad \frac{(\text{S}/\text{NP})/(\text{S}/\text{NP})}{\text{S}/\text{NP}} \quad < \\
& \quad \frac{\text{S}/\text{NP}}{\text{S}} \quad >
\end{align*}
\]
Likewise, argument cluster coordination is licensed:

\[
\begin{align*}
give & \quad \text{Walt} \quad \text{the salt} \quad \text{and} \quad \text{Malcolm} \quad \text{the talcum} \\
\text{DTV} & \quad \text{VP/DTV} \quad \text{VP/DTV} \quad \text{VP/DTV} \quad \text{VP/DTV} \quad \text{VP/DTV} \\
\text{TV} & \quad \text{TV/DTV} \quad \text{TV/DTV} \quad \text{TV/DTV} \quad \text{TV/DTV} \quad \text{TV/DTV} \\
\text{VP} & \quad \text{VP} \quad \text{VP} \quad \text{VP} \quad \text{VP} \quad \text{VP}
\end{align*}
\]

where:

- \( \text{VP} = S/\text{NP} \)
- \( \text{TV} \) (transitive verb) = \( \text{VP}/\text{NP} \)
- \( \text{DTV} \) (ditransitive verb) = \( \text{VP}/\text{NP}/\text{NP} \)
Extraction

Try to work out derivations for the following object extractions:

(19) what John knows (NP)

(20) What does John know?

Remember:

- Categories can select for complex categories
- *wh*-pronouns (e.g., *what*) can select for gapped constructions
Non-standard surface structures are licensed throughout

- The non-traditional constituents motivated for right-node raising and similar coordinations are also possible constituents of non-coordinate sentences like *Marcel proved completeness*.

The Principle of Type Transparency guarantees that all such non-standard derivations yield identical interpretations.
Constituent Condition on Rules

Traditionally, there has been the notion of the **Constituent Condition on Rules**

*The inputs and outputs of all rules of syntax should be constituents*

Thus, CCG has to claim one of the following:

1. This claim is not true.
2. Items like *Marcel proved* can be constituents.
One potential advantage of CCG is that you can derive sentences in a completely incremental, left-to-right fashion (some type-raising omitted here) (Dowty 1988):

\[
\begin{array}{c}
\text{I} \\
\frac{\text{S/VP}}{} \\
\text{believe} \\
\frac{\text{VP/S'}}{} > \text{B} \\
\text{that} \\
\frac{\text{S'/S}}{} > \text{B} \\
\text{she} \\
\frac{\text{S/VP}}{} \\
\text{ate} \\
\frac{\text{VP/NP}}{} > \text{B} \\
\text{dinner} \\
\frac{\text{NP}}{} \\
\text{S} \\
\frac{\text{S/VP}}{} > \text{B} \\
\frac{\text{S/VP}}{} > \text{B} \\
\frac{\text{S/NP}}{} > \text{B} \\
\frac{\text{S}}{} >
\end{array}
\]
According to Steedman, the non-standard surface structures are not spurious ambiguities but relevant since they subsume the intonation structures needed to explain the possible intonation contours for sentences of English.

Intonational boundaries contribute to determining which of the possible combinatory derivations is intended.

The interpretations of the constituents that arise from these derivations are related to semantic distinctions of information structure and discourse focus.

Steedman’s claims:
- Where intonational boundaries are present, they contribute to disambiguation.
- Conversely, any such boundaries must be consistent with some syntactic derivation, or ill-formedness will result.
Examples for impossible intonation boundaries

(21)  
b.  * (Seymour prefers the nuts) (and bolts approach).
c.  * (They only asked whether John knew the woman who chaired) (the zoning board).
Steedman's claims:

- *Surface structure* and *information structure* coincide, the latter simply consisting in the interpretation associated with a constituent analysis of the sentence.

- *Intonation* coincides with *surface structure*, and hence information structure, in the sense that all intonational boundaries coincide with syntactic boundaries (but not all syntactic boundaries are intonationally marked).
Syntactic structure and intonation (cont.)

As a result, fragments such as *Marcel proved* in (22c), are not only prosodic constituents but surface syntactic constituents, complete with interpretations.

(22)  a. Marcel proved completeness.

b. \[ S \text{ Marcel} [VP \text{ proved completeness} ] \]

c. \[ S \text{ [?P Marcel proved} ] \text{ completeness} \]
To deal with the exponential potential of type-raising, CCG parsers can do different things (see Hockenmaier & Steedman 2002)

- **Logical form check:** Put logical/semantic forms into a chart & only add a new constituent when form isn’t already in there
- **Last resort:** Only use the rules of type-raising and functional composition when no other parse works

**NB:** The Penn Treebank has been put into CCG format and is available as the CCGBank.
Summary

- CG has complex lexical entries with syntactic and semantic information and a few rules to apply to them.
- CCG adds combinators, which allows for flexible constituency and all sorts of coordination phenomena.
- The “non-constituent” clustering of information is claimed to be consistent with intonational/information structure.