Surface Reasoning

Lecture 2: Logic and Grammar

Thomas Icard

June 18-22, 2012

(Ajdukiewicz/Bar-Hillel) Categorial Grammar

Define a set CAT of categories as follows:

- Some set of basic categories is in CAT.
- If $A, B \in \mathrm{CAT}$, then both $A/B \in \mathrm{CAT}$ and $A\backslash B \in \mathrm{CAT}$.

Two basic rules:

(FA) $A/B, B \Rightarrow A$.
(BA) $B, A\backslash B \Rightarrow A$.

If we add to these two more rules we obtain a basic proof system:

(id) $A \Rightarrow A$.
(cut) If $\Gamma, A, \Gamma' \Rightarrow B$ and $\Delta \Rightarrow A$, then $\Gamma, \Delta, \Gamma' \Rightarrow B$.
Here Γ and Δ are finite sequences of categories.

Definition

CG is the smallest relation containing (id), (FA), and (BA), and closed under (cut).

- Given a set Σ of basic lexical items, e.g. natural language expressions, a *lexicon* is an assignment of a finite number of categories to each lexical item:

$$\text{LEX} \subseteq \Sigma \times \text{CAT}.$$

- A string $w_1, ..., w_n \in \Sigma^+$ is an expression of type B just in case there is a sequence of categories $A_1, ..., A_n$ such that $\langle w_i, A_i \rangle \in$ LEX, for each $i \leq n$, and $A_1, ..., A_n \Rightarrow B$.

A toy lexicon:

- Theodore, np
- candidate, n
- every, some, $(s/(s\backslash np))/n$
- broccoli, np
- likes, $(s\backslash np)/np$
- who, $(n\backslash n)/(s\backslash np)$

Or, abbreviating $iv = s\backslash np$ and $tv = iv/np$, this simplifies to:

- Theodore, np
- candidate, n
- every, some, $(s/iv)/n$
- broccoli, np
- likes, tv
- who, $(n\backslash n)/iv$

Example:

$$\dfrac{\dfrac{\text{Theodore}}{np} \qquad \dfrac{\dfrac{\text{likes}}{(s\backslash np)/np} \qquad \dfrac{\text{broccoli}}{np}}{s\backslash np}}{s}$$

A toy lexicon:

- Theodore, np
- candidate, n
- every, some, $(s/(s\backslash np))/n$
- broccoli, np
- likes, $(s\backslash np)/np$
- who, $(n\backslash n)/(s\backslash np)$

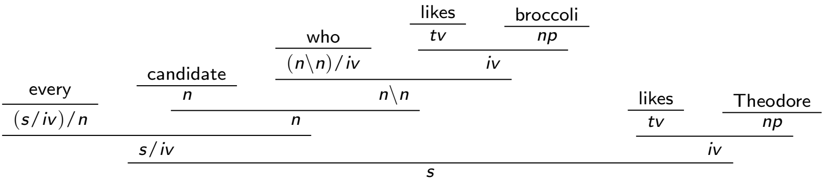
Or, abbreviating $iv = s\backslash np$ and $tv = iv/np$, this simplifies to:

- Theodore, np
- candidate, n
- every, some, $(s/iv)/n$
- broccoli, np
- likes, tv
- who, $(n\backslash n)/iv$

Example:

$$\dfrac{\dfrac{\text{Theodore}}{np} \qquad \dfrac{\dfrac{\text{likes}}{tv} \qquad \dfrac{\text{broccoli}}{np}}{iv}}{s}$$

Longer example:



CG versus CFG

Theorem (Gaifman)

The class of languages generated by context free grammars coincides with the class of languages accepted by categorial grammars.

- Recall our lexicon LEX:
  - Theodore, *np*
  - candidate, *n*
  - every, some, $(s/iv)/n$
  - broccoli, *np*
  - likes, *tv*
  - who, $(n\backslash n)/iv$
- A context free grammar generating the same set of strings would be:

$$\begin{array}{rcl} S & \longrightarrow & NP\ VP \\ NP & \longrightarrow & \text{every } N \mid \text{some } N \mid PN \mid NP \text{ who } VP \\ N & \longrightarrow & \text{candidate} \\ PN & \longrightarrow & \text{Theodore} \mid \text{broccoli} \\ VP & \longrightarrow & \text{likes } PN \end{array}$$

- Problem: the following are not strings in the language:
  - 'who Theodore likes'
  - 'likes some candidate'
- In particular, we cannot parse:
  - 'Every candidate who Theodore likes likes some candidate'
- For 'who Theodore likes' we would need 'who' to have category $((n\backslash n)/tv)/np$ in addition to $(n\backslash n)/iv$:

$$\dfrac{\dfrac{\dfrac{\text{who}}{((n\backslash n)/tv)/np} \qquad \dfrac{\text{Theodore}}{np}}{(n\backslash n)/tv} \qquad \dfrac{\text{likes}}{tv}}{n/n}$$

- Similarly, 'all' and 'some' would have to have a second category $(iv\backslash tv)/n$ for object position, in addition to $(s/iv)/n$.
- This is inelegant and seems to miss some cross-categorial generalizations.

- Combinatory Categorial Grammar (CCG) is an extension of CG with several further rules. (For more go to Mark Steedman's course!)

(>B) $A/B,\ B/C \Rightarrow A/C$

(<B) $B\backslash C,\ A\backslash B \Rightarrow A\backslash C$

(>T) $A \Rightarrow B/(B\backslash A)$

(<T) $A \Rightarrow B\backslash(B/A)$

- Using >**B** and >**T** we can now parse 'who Theodore likes':

$$\dfrac{\dfrac{\text{who}}{(n\backslash n)/(s/np)} \qquad \dfrac{\dfrac{\dfrac{\text{Theodore}}{np}}{s/(s\backslash np)}\,(>\mathbf{T}) \qquad \dfrac{\text{likes}}{(s\backslash np)/np}}{s/np}\,(>\mathbf{B})}{n\backslash n}$$

- CCG can also capture quantifiers in object position by assigning 'some' and 'all' an only slightly adjusted category $(s\backslash(s/np))/n$, in addition to $(s/(s\backslash np))/n$ for subject position:

Theodore: np; ($>$**T**) $s/(s\backslash np)$
likes: $(s\backslash np)/np$
some: $(s\backslash(s/np))/n$
candidate: n

($>$**B**) $s/(s\backslash np)$, $(s\backslash np)/np$ ⟹ s/np

$(s\backslash(s/np))/n$, n ⟹ $s\backslash(s/np)$

s/np, $s\backslash(s/np)$ ⟹ s

- CCG has another rule:

($<$**S**$_x$) $B/C,\ (A\backslash B)/C \Rightarrow A/C$

- In general, CCG is stronger than context free, equivalent to so called *linear index grammars* (like TAG and other grammatical formalisms).

- Lambek Calculus is an alternative to CCG. The main idea is that categories correspond to logical formulas, and category forming operators correspond to logical constants. *"Parsing as deduction"*
- The setting is Gentzen-style Natural Deduction, where $\Gamma \blacktriangleright A$ means the sequence Γ is of category A.
- The basic Lambek Calculus L is given by the following rules:

$$(\mathrm{Ax})\ \frac{}{A \blacktriangleright A}$$

$$(/E)\ \frac{\Delta \blacktriangleright A/B \qquad \Gamma \blacktriangleright B}{\Delta, \Gamma \blacktriangleright A} \qquad \frac{\Gamma \blacktriangleright B \qquad \Delta \blacktriangleright A\backslash B}{\Gamma, \Delta \blacktriangleright A}\ (\backslash E)$$

$$(/I)\ \frac{\Delta, B \blacktriangleright A}{\Delta \blacktriangleright A/B} \qquad \frac{B, \Delta \blacktriangleright A}{\Delta \blacktriangleright A\backslash B}\ (\backslash I)$$

$$(\text{Ax})\ \frac{}{A \blacktriangleright A}$$

$$(/E)\ \frac{\Delta \blacktriangleright A/B \qquad \Gamma \blacktriangleright B}{\Delta, \Gamma \blacktriangleright A} \qquad \frac{\Gamma \blacktriangleright B \qquad \Delta \blacktriangleright A\backslash B}{\Gamma, \Delta \blacktriangleright A}\ (\backslash E)$$

$$(/I)\ \frac{\Delta, B \blacktriangleright A}{\Delta \blacktriangleright A/B} \qquad \frac{B, \Delta \blacktriangleright A}{\Delta \blacktriangleright A\backslash B}\ (\backslash I)$$

- From these follow all of the CCG rules, with the exception of $<\mathbf{S}_x$.

$$\cfrac{\cfrac{\Delta \blacktriangleright A/B \qquad \cfrac{\Gamma \blacktriangleright B/C \qquad [C \blacktriangleright C]^1}{\Gamma, C \blacktriangleright B}\ (/E)}{\Delta, \Gamma, C \blacktriangleright A}\ (/E)}{\Delta, \Gamma \blacktriangleright A/C}\ (/I)^1$$

- That is, if Δ is of category A/B and Γ is of category B/C, then Δ, Γ is of category A/C. This is just rule $>\mathbf{B}$.

$$(\text{Ax})\ \frac{}{A \blacktriangleright A}$$

$$(/E)\ \frac{\Delta \blacktriangleright A/B \qquad \Gamma \blacktriangleright B}{\Delta, \Gamma \blacktriangleright A} \qquad \frac{\Gamma \blacktriangleright B \qquad \Delta \blacktriangleright A\backslash B}{\Gamma, \Delta \blacktriangleright A}\ (\backslash E)$$

$$(/I)\ \frac{\Delta, B \blacktriangleright A}{\Delta \blacktriangleright A/B} \qquad \frac{B, \Delta \blacktriangleright A}{\Delta \blacktriangleright A\backslash B}\ (\backslash I)$$

- From these follow all of the CCG rules, with the exception of $<\mathbf{S}_x$.

$$\cfrac{\cfrac{\Delta \blacktriangleright A \qquad [B\backslash A \blacktriangleright B\backslash A]^1}{\Delta, B\backslash A \blacktriangleright B}\ (\backslash E)}{\Delta \blacktriangleright B/(B\backslash A)}\ (/I)^1$$

- That is, if Δ is of category A, then it is also of category $B/(B\backslash A)$. This is rule $>\mathbf{T}$.

- Again, we cannot derive $<\mathbf{S}_x$, which means L is strictly weaker than CCG.

Theorem (Pentus)

L *is context free.*

- Still, it allows for elegant derivations without excess categories:

$$\frac{\text{likes} \blacktriangleright tv \qquad [np \blacktriangleright np]^1}{\text{likes}, np \blacktriangleright s\backslash np}\;(/E)$$

$$\frac{\text{Theodore} \blacktriangleright np \qquad \text{likes}, np \blacktriangleright s\backslash np}{\text{Theodore likes}, np \blacktriangleright s}\;(\backslash E)$$

$$\frac{\text{Theodore likes}, np \blacktriangleright s}{\text{Theodore likes} \blacktriangleright s/np}\;(/I)^1$$

$$\frac{\text{who} \blacktriangleright (n\backslash n)/(s/np) \qquad \text{Theodore likes} \blacktriangleright s/np}{\text{who Theodore likes} \blacktriangleright n\backslash n}\;(\backslash E)$$

- As is well known, in natural language syntax tree structure matters. We sometimes cannot assume our sequences satisfy associativity.
- The weakest of the *Categorial Type Logics* is NL:

$$(\mathrm{Ax})\ \frac{}{A \blacktriangleright A}$$

$$(/E)\ \frac{\Delta \blacktriangleright A/B \qquad \Gamma \blacktriangleright B}{(\Delta \circ \Gamma) \blacktriangleright A} \qquad \frac{\Gamma \blacktriangleright B \qquad \Delta \blacktriangleright A\backslash B}{(\Gamma \circ \Delta) \blacktriangleright A}\ (\backslash E)$$

$$(/I)\ \frac{(\Delta \circ B) \blacktriangleright A}{\Delta \blacktriangleright A/B} \qquad \frac{(B \circ \Delta) \blacktriangleright A}{\Delta \blacktriangleright A\backslash B}\ (\backslash I)$$

- Adding associativity gives us back L:

$$\frac{\Gamma[\Delta_1 \circ (\Delta_2 \circ \Delta_3)] \blacktriangleright C}{\Gamma[(\Delta_1 \circ \Delta_2) \circ \Delta_3] \blacktriangleright C}$$

- Adding commutativity gives a system called LP:

$$\frac{\Gamma[(\Delta_1 \circ \Delta_2)] \blacktriangleright C}{\Gamma[(\Delta_2 \circ \Delta_1)] \blacktriangleright C}$$

- Clearly, in LP forward and backward slash collapse into a single binary operator.

Review of Types

- Recall the simple type system $\mathcal{T}$:
  - Basic types, throughout these notes e and t, are in $\mathcal{T}$;
  - If $\tau, \sigma \in \mathcal{T}$, then $(\tau \rightarrow \sigma) \in \mathcal{T}$.
- We can define a function $\textsc{type}$: CAT $\rightarrow \mathcal{T}$ such that:
  - $\textsc{type}(np) = e$;
  - $\textsc{type}(s) = t$;
  - $\textsc{type}(n) = (e \rightarrow t)$;
  - $\textsc{type}(A/B) = \textsc{type}(A\backslash B) = (\textsc{type}(B) \rightarrow \textsc{type}(A))$.

Review of Lambda Calculus

- We now define the class of λ-terms of type τ, denoted Λ_τ:
  - Variables of type τ are in Λ_τ.
  - Constants of type τ are in Λ_τ.
  - If $\alpha \in \Lambda_{\tau\to\sigma}$ and $\beta \in \Lambda_\tau$, then $\alpha(\beta) \in \Lambda_\sigma$.
  - If x is a variable of type τ and $\alpha \in \Lambda_\sigma$, then $\lambda x.\alpha \in \Lambda_{\tau\to\sigma}$.

- β and η reduction rules:

  (β) $(\lambda x_\tau.\alpha_\sigma)(\beta_\tau) \Longrightarrow \alpha_\sigma[\beta_\tau/x_\tau]$, provided x_τ is free for β_τ in α_σ.

  (η) $\lambda x_\tau.\alpha_{\tau\to\sigma}(x_\tau) \Longrightarrow \alpha_{\tau\to\sigma}$, provided x_τ is not free in $\alpha_{\tau\to\sigma}$.

- The domain $\mathcal{D} = \bigcup_{\tau\in\mathcal{T}} D_\tau$ is given by:
  - D_e is assumed to be fixed set E of entities.
  - $D_t = \{0, 1\}$.
  - $D_{\tau\to\sigma} = D_\sigma^{D_\tau}$.

- A model is a pair $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$, with $\mathcal{D}$ a domain and $\mathcal{I}$: LEX$\to \mathcal{D}$, so that if $\textsc{type}(\text{A}) = \tau$, then $\mathcal{I}(\langle w, A\rangle) \in D_\tau$.

- Now grammars must be given by the set of lexical items, their categories, *and corresponding λ-terms*.

Theodore	np	`theo`
broccoli	np	`broc`
candidate	n	`cand`
likes	$(s\backslash np)/np$	`like`
who	$(n\backslash n)/(s\backslash np)$	$\lambda x.\lambda y.\lambda z.x(z) \wedge y(z)$
every	$(s/(s\backslash np))/n$	$\lambda x.\lambda y.\forall z(x(z) \rightarrow y(z))$
some	$(s/(s\backslash np))/n$	$\lambda x.\lambda y.\exists z(x(z) \wedge y(z))$
no	$(s/(s\backslash np))/n$	$\lambda x.\lambda y.\neg\exists z(x(z) \wedge y(z))$

- To use quantifiers in object position we could add:

every	$(s\backslash(s/np))/n$	$\lambda x.\lambda y.\forall z(x(z) \rightarrow y(z))$
some	$(s\backslash(s/np))/n$	$\lambda x.\lambda y.\exists z(x(z) \wedge y(z))$
no	$(s\backslash(s/np))/n$	$\lambda x.\lambda y.\neg\exists z(x(z) \wedge y(z))$

- In NL, L, LP and other Categorial Type Logics, because the syntactic rules are logical rules, semantics comes 'for free' from the Curry-Howard Correspondence between natural deduction proofs in intuitionistic implicational logic and typed λ-terms.
- Because all these systems are weaker than IIL, we must take a sublanguage of full λ-calculus. Johan van Benthem proved that the correspondence holds for this fragment.
- Our rules for NL now become:

$$(\mathrm{Ax})\ \frac{}{x : A \blacktriangleright x : A}$$

$$(/E)\ \frac{\Delta \blacktriangleright t : A/B \qquad \Gamma \blacktriangleright u : B}{(\Delta \circ \Gamma) \blacktriangleright t(u) : A} \qquad \frac{(\Delta \circ x : B) \blacktriangleright t : A}{\Delta \blacktriangleright \lambda x.t : A/B}\ (/I)$$

$$(\backslash E)\ \frac{\Gamma \blacktriangleright u : B \qquad \Delta \blacktriangleright t : A\backslash B}{(\Gamma \circ \Delta) \blacktriangleright t(u) : A} \qquad \frac{(x : B \circ \Delta) \blacktriangleright t : A}{\Delta \blacktriangleright \lambda x.t : A\backslash B}\ (\backslash I)$$

- We write $\mathrm{NL} \vdash \Gamma \blacktriangleright t : A$, and likewise for L and LP.

Type Raising

Semantically, type-raising corresponds to a certain λ-abstraction.

$$\dfrac{\dfrac{\Delta \blacktriangleright t : A \qquad [x : B\backslash A \blacktriangleright x : B\backslash A]^1}{(\Delta \circ x : B\backslash A) \blacktriangleright x(t) : B}\,(\backslash E)}{\Delta \blacktriangleright \lambda x.x(t) : B/(B\backslash A)}\,(/I)^1$$

Example

$$\dfrac{\text{likes} \blacktriangleright \texttt{like} : tv \qquad [w : np \blacktriangleright w : np]^1}{\text{likes},\ w : np \blacktriangleright \texttt{like}(w) : s\backslash np}\ (/E)$$

$$\dfrac{\text{Theodore} \blacktriangleright \texttt{theo} : np \qquad \text{likes},\ w : np \blacktriangleright \texttt{like}(w) : s\backslash np}{\text{Theodore likes},\ w : np \blacktriangleright \texttt{like}(\texttt{theo}, w) : s}\ (\backslash E)$$

$$\dfrac{\text{Theodore likes},\ w : np \blacktriangleright \texttt{like}(\texttt{theo}, w) : s}{\text{Theodore likes} \blacktriangleright \lambda w.\texttt{like}(\texttt{theo}, w) : s/np}\ (/I)^1$$

$$\dfrac{\text{who} \blacktriangleright \lambda x.\lambda y.\lambda z.x(z) \wedge y(z) : (n\backslash n)/(s/np) \qquad \text{Theodore likes} \blacktriangleright \lambda w.\texttt{like}(\texttt{theo}, w) : s/np}{\text{who Theodore likes} \blacktriangleright \lambda y.\lambda z.\texttt{like}(\texttt{theo}, z) \wedge y(z) : n\backslash n}\ (\backslash E)$$

- We can combine this with ‘candidate’ to form a complex predicate:

  $$\text{NL} \vdash \text{candidate who Theodore likes} \blacktriangleright \lambda z.\texttt{like}(\texttt{theo}, z) \wedge \texttt{cand}(z) : n$$

  which is exactly the right result.

- Slogan: *“Meaning is a by-product of syntactic derivation.”*

- The fundamental idea of van Benthem and Sánchez-Valencia's Natural Logic is to forget about the λ-terms, shifting a small amount of the semantics into the syntax, in particular into the categories.

- The crucial features are monotonicity properties of functions.

- Consider the meaning of 'every': $\lambda x.\lambda y.\forall z(x(z) \rightarrow y(z))$.
  As we saw on the first day, this function is antitone in its first argument, monotone in its second, if we order the domains as usual.

- To capture this, let us write the category of 'every' as

$$(s/(s\backslash np)^{+})/n^{-}$$

  or abbreviating,

$$(s/iv^{+})/n^{-}.$$

- We can say more generally that A/B^{+} and $A\backslash B^{+}$ are categories of monotone functional items, and A/B^{-} and $A\backslash B^{-}$ of antitone.

- The steps of Sánchez-Valencia's polarity marking algorithm are:
  1. Assign lexical items their appropriate marked types.
  2. Propagate these markings down the proof tree.
  3. The polarity of each node is computed.
- The result is a proof tree with just enough information to support some basic inferential patterns (stay tuned).
- For Step 1 we might label our lexicon as follows:
  - Theodore, np
  - candidate, n
  - every, $(s/iv^+)/n^-$
  - some, $(s/iv^+)/n^+$
  - broccoli, np
  - likes, iv/np^+
  - who, $(n\backslash n^+)/iv^+$
  - no, $(s/iv^-)/n^-$
- When interpreting such terms in models we require terms of category A/B^+ and $A\backslash B^+$ are mapped to monotone functions, and those of A/B^- and $A\backslash B^-$ are mapped to antitone functions.

Step 2

Where $* \in \{+, -\}$:

- (/E):

$$\frac{\Delta \blacktriangleright A/B \qquad \Gamma \blacktriangleright B}{(\Delta \circ \Gamma) \blacktriangleright A} \quad \Longrightarrow \quad \frac{\overset{\Delta \blacktriangleright A/B}{+} \qquad \Gamma \blacktriangleright B}{(\Delta \circ \Gamma) \blacktriangleright A}$$

$$\frac{\Delta \blacktriangleright A/B^{*} \qquad \Gamma \blacktriangleright B}{(\Delta \circ \Gamma) \blacktriangleright A} \quad \Longrightarrow \quad \frac{\overset{\Delta \blacktriangleright A/B^{*}}{+} \qquad \overset{\Gamma \blacktriangleright B}{*}}{(\Delta \circ \Gamma) \blacktriangleright A}$$

- (/I):

$$\frac{\begin{matrix}[B \blacktriangleright B]^{i} \\ \vdots \\ \Delta \circ B \blacktriangleright A\end{matrix}}{\Delta \blacktriangleright A/B} \quad \Longrightarrow \quad \frac{\begin{matrix}[B \blacktriangleright B]^{i} \\ \vdots \\ \Delta \circ B \blacktriangleright A \\ +\end{matrix}}{\Delta \blacktriangleright A/B^{m}}$$

- m is $-$ (resp. $+$) if all the nodes on the path from $\Delta \circ B \blacktriangleright A$ to $[B \blacktriangleright B]^{i}$ are marked, and an odd (resp. even) number are $-$.

Step 2

Where $* \in \{+, -\}$:

- ($\backslash$E):

$$\frac{\Delta \blacktriangleright A\backslash B \qquad \Gamma \blacktriangleright B}{(\Gamma \circ \Delta) \blacktriangleright A} \Longrightarrow \frac{\overset{\Delta \blacktriangleright A\backslash B}{+} \qquad \Gamma \blacktriangleright B}{(\Gamma \circ \Delta) \blacktriangleright A}$$

$$\frac{\Delta \blacktriangleright A\backslash B^{*} \qquad \Gamma \blacktriangleright B}{(\Gamma \circ \Delta) \blacktriangleright A} \Longrightarrow \frac{\overset{\Delta \blacktriangleright A\backslash B^{*}}{+} \qquad \overset{\Gamma \blacktriangleright B}{*}}{(\Gamma \circ \Delta) \blacktriangleright A}$$

- ($\backslash$I):

$$\frac{\begin{matrix}[B \blacktriangleright B]^{i} \\ \vdots \\ B \circ \Delta \blacktriangleright A\end{matrix}}{\Delta \blacktriangleright A\backslash B} \Longrightarrow \frac{\begin{matrix}[B \blacktriangleright B]^{i} \\ \vdots \\ B \circ \Delta \blacktriangleright A \\ +\end{matrix}}{\Delta \blacktriangleright A\backslash B^{m}}$$

- m is $-$ (resp. $+$) if all the nodes on the path from $\Delta \circ B \blacktriangleright A$ to $[B \blacktriangleright B]^{i}$ are marked, and an odd (resp. even) number are $-$.

Step 3

- The final step is quite simple:
  1. Mark the root node with $+$.
  2. Starting at the leaf nodes, check whether every node along the path to the root is marked.
  3. If it is, and there are an odd number of nodes marked $-$, label the node with $-$. If there are an even number, label it with $+$.

- The result is a parsed expression with monotonicity information explicitly represented.

- Using this we can build a simple *Monotonicity Calculus*:

$$\frac{[_S \dots X^+ \dots] \qquad [\![X]\!] \subseteq [\![Y]\!]}{[_S \dots Y^+ \dots]} \qquad \frac{[_S \dots X^- \dots] \qquad [\![Y]\!] \subseteq [\![X]\!]}{[_S \dots Y^- \dots]}$$

- Sánchez-Valencia proved a Soundness Theorem [4]. We may also have time to prove one in this course.

Example 1

every ►$(s/iv^{+})/n^{-}$ candidate ►n likes ►iv/np^{+} broccoli ►np

every candidate ►s/iv^{+} likes broccoli ►iv

every candidate likes broccoli ►s

$\Downarrow$

every ►$(s/iv^{+})/n^{-}$ candidate ►n likes ►iv/np^{+} broccoli ►np

$+$ $-$ $+$ $+$

every candidate ►s/iv^{+} likes broccoli ►iv

$+$ $+$

every candidate likes broccoli ►s

$\Downarrow$

every ►$(s/iv^{+})/n^{-}$ candidate ►n likes ►iv/np^{+} broccoli ►np

$+$ $-$ $+$ $+$

every candidate ►s/iv^{+} likes broccoli ►iv

$+$ $+$

every candidate likes broccoli ►s

$+$

Example 1

- We can thus write this expression as

$$((\text{every}^{+}\text{candidate}^{-})^{+}(\text{likes}^{+}\text{broccoli}^{+})^{+})^{+}$$

- This means, if we replace 'candidate' with something smaller, the resulting expression is entailed by this one.
- For any of the subexpressions labeled with $+$ (which includes all others in this example), replacing them with something of the same type with larger extension preserves validity.
- For instance:

$$\frac{\text{every candidate}^{-}\ \text{likes broccoli} \qquad [\![\text{hopeful candidate}]\!] \subseteq [\![\text{candidate}]\!]}{\text{every (hopeful candidate)}^{-}\ \text{likes broccoli}}$$

- While:

$$\frac{\text{every candidate likes}^{+}\ \text{broccoli} \qquad [\![\text{likes}]\!] \subseteq [\![\text{tolerates}]\!]}{\text{every candidate tolerates}^{+}\ \text{broccoli}}$$

Example 2

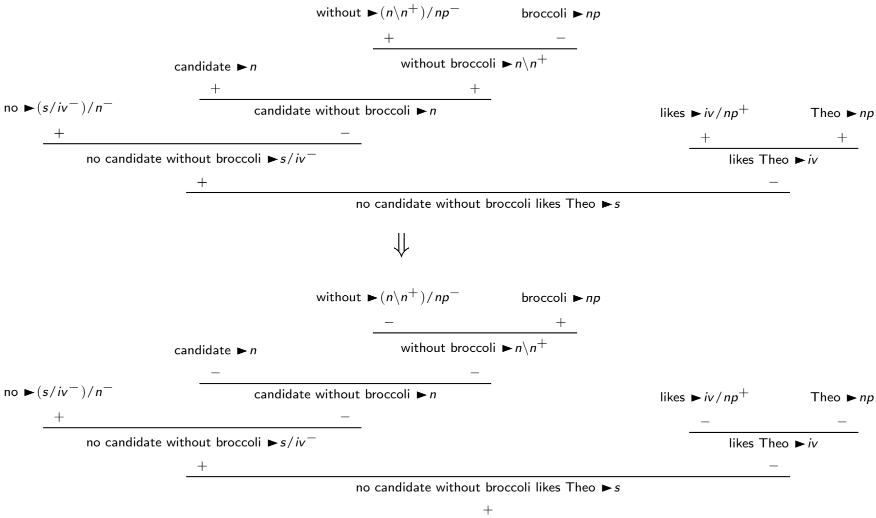
For a slightly more interesting example, let us add one word to the lexicon, without : $(n\backslash n^{+})/np^{-}$.

$$\dfrac{\dfrac{\text{no} \blacktriangleright (s/iv^{-})/n^{-} \qquad \dfrac{\text{candidate} \blacktriangleright n \qquad \dfrac{\text{without} \blacktriangleright (n\backslash n^{+})/np^{-} \qquad \text{broccoli} \blacktriangleright np}{\text{without broccoli} \blacktriangleright n\backslash n^{+}}}{\text{candidate without broccoli} \blacktriangleright n}}{\text{no candidate without broccoli} \blacktriangleright s/iv^{-}} \qquad \dfrac{\text{likes} \blacktriangleright iv/np^{+} \qquad \text{Theo} \blacktriangleright np}{\text{likes Theo} \blacktriangleright iv}}{\text{no candidate without broccoli likes Theo} \blacktriangleright s}$$

$$\Downarrow$$

$$\dfrac{\underset{+}{\dfrac{\underset{+}{\text{no} \blacktriangleright (s/iv^{-})/n^{-}} \qquad \underset{-}{\dfrac{\underset{+}{\text{candidate} \blacktriangleright n} \qquad \underset{+}{\dfrac{\underset{+}{\text{without} \blacktriangleright (n\backslash n^{+})/np^{-}} \qquad \underset{-}{\text{broccoli} \blacktriangleright np}}{\text{without broccoli} \blacktriangleright n\backslash n^{+}}}}{\text{candidate without broccoli} \blacktriangleright n}}}{\text{no candidate without broccoli} \blacktriangleright s/iv^{-}}} \qquad \underset{-}{\dfrac{\underset{+}{\text{likes} \blacktriangleright iv/np^{+}} \qquad \underset{+}{\text{Theo} \blacktriangleright np}}{\text{likes Theo} \blacktriangleright iv}}}{\text{no candidate without broccoli likes Theo} \blacktriangleright s}$$

Example 2


without ► (n\n+)/np−
broccoli ► np
+ −
without broccoli ► n\n+
candidate ► n
+ +
candidate without broccoli ► n
no ► (s/iv−)/n−
+ −
no candidate without broccoli ► s/iv−
likes ► iv/np+
Theo ► np
+ +
likes Theo ► iv
+ −
no candidate without broccoli likes Theo ► s
⇓
without ► (n\n+)/np−
broccoli ► np
− +
without broccoli ► n\n+
candidate ► n
− −
candidate without broccoli ► n
no ► (s/iv−)/n−
+ −
no candidate without broccoli ► s/iv−
likes ► iv/np+
Theo ► np
− −
likes Theo ► iv
+ −
no candidate without broccoli likes Theo ► s
+

Example 2

- The polarity profile now looks as follows:

$$((\text{no}^{+}(\text{candidate}^{-}(\text{without}^{-}\text{broccoli}^{+})^{-})^{-})^{+}(\text{likes}^{-}\text{Theo}^{-})^{-})^{+}.$$

- This is reflected in different inference patterns:

$$\frac{\text{no candidate without broccoli likes}^{-}\text{ Theo} \qquad [\![\text{adores}]\!] \subseteq [\![\text{likes}]\!]}{\text{no candidate without broccoli adores}^{-}\text{ Theo}}$$

- While:

$$\frac{\text{no candidate without broccoli}^{+}\text{ likes Theo} \qquad [\![\text{broccoli}]\!] \subseteq [\![\text{cabbage}]\!]}{\text{no candidate without cabbage}^{+}\text{ likes Theo}}$$

Example 3

$$\dfrac{\text{who} \blacktriangleright (n\backslash n^{+})/(s/np)^{+} \qquad \dfrac{\dfrac{\text{Theodore} \blacktriangleright np \qquad \dfrac{\text{likes} \blacktriangleright (s\backslash np^{+})/np^{+} \qquad [np \blacktriangleright np]^{1}}{\text{likes, } np \blacktriangleright s\backslash np^{+}}\ (/E)}{\text{Theodore likes, } np \blacktriangleright s}\ (\backslash E)}{\text{Theodore likes} \blacktriangleright s/np}\ (/I)^{1}}{\text{who Theodore likes} \blacktriangleright n\backslash n^{+}}\ (\backslash E)$$

$$\Downarrow$$

$$\dfrac{\begin{array}{c}\text{who} \blacktriangleright (n\backslash n^{+})/(s/np^{+})^{+} \qquad \dfrac{\begin{array}{c}\dfrac{\begin{array}{c}\text{Theodore} \blacktriangleright np \qquad \dfrac{\begin{array}{c}\text{likes} \blacktriangleright (s\backslash np^{+})/np^{+} \qquad [np \blacktriangleright np]^{1} \\ + \qquad\qquad\qquad\qquad + \end{array}}{\text{likes, } np \blacktriangleright s\backslash np^{+}}\ (/E) \\ + \qquad\qquad\qquad\qquad\qquad\qquad + \end{array}}{\text{Theodore likes, } np \blacktriangleright s}\ (\backslash E) \\ + \end{array}}{\text{Theodore likes} \blacktriangleright s/np^{+}}\ (/I)^{1} \\ + \qquad\qquad\qquad\qquad\qquad\qquad + \end{array}}{\text{who Theodore likes} \blacktriangleright n\backslash n^{+}}\ (\backslash E)$$

Example 3

$$\dfrac{\overset{+}{\text{who} \blacktriangleright (n\backslash n^{+})/(s/np^{+})^{+}} \qquad \overset{+}{\dfrac{\overset{+}{\dfrac{\overset{+}{\text{Theodore} \blacktriangleright np} \qquad \overset{+}{\dfrac{\overset{+}{\text{likes} \blacktriangleright (s\backslash np^{+})/np^{+}} \qquad \overset{+}{[np \blacktriangleright np]^{1}}}{\text{likes},\ np \blacktriangleright s\backslash np^{+}}\,(/E)}}{\text{Theodore likes},\ np \blacktriangleright s}\,(\backslash E)}}{\text{Theodore likes} \blacktriangleright s/np^{+}}\,(/I)^{1}}}{\text{who Theodore likes} \blacktriangleright n\backslash n^{+}}\,(\backslash E)$$

$$\Downarrow$$

$$\underset{+}{\dfrac{\overset{+}{\text{who} \blacktriangleright (n\backslash n^{+})/(s/np^{+})^{+}} \qquad \overset{+}{\dfrac{\overset{+}{\dfrac{\overset{+}{\text{Theodore} \blacktriangleright np} \qquad \overset{+}{\dfrac{\overset{+}{\text{likes} \blacktriangleright (s\backslash np^{+})/np^{+}} \qquad \overset{+}{[np \blacktriangleright np]^{1}}}{\text{likes},\ np \blacktriangleright s\backslash np^{+}}\,(/E)}}{\text{Theodore likes},\ np \blacktriangleright s}\,(\backslash E)}}{\text{Theodore likes} \blacktriangleright s/np^{+}}\,(/I)^{1}}}{\text{who Theodore likes} \blacktriangleright n\backslash n^{+}}}\,(\backslash E)$$

- Recently, Jan van Eijk devised a variation on Sánchez-Valencia's algorithm, requiring only a single, "top-down" pass.

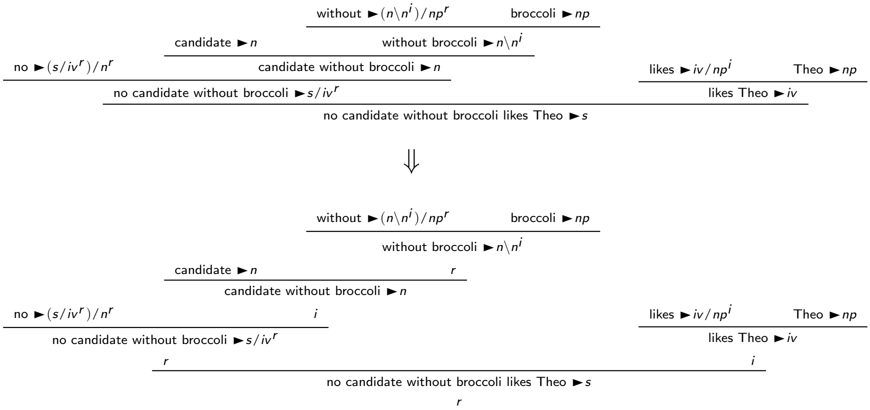
- The first step is to change the category markings. Instead of $+$ and $-$, we use three functions i, r, and b over $M = \{+, -, \#\}$, where $\#$ is the uninformative marking:
  - $i(m) = m$ for all $m \in M$, i.e. i is identity.
  - $r(+) = -$, $r(-) = +$, and $r(0) = 0$, i.e. r is reversal.
  - $b(m) = 0$ for all $m \in M$.

- Our grammar (with a few new items) then becomes:
  - Theodore, np
  - candidate, n
  - every, $(s\,/\,iv^i)\,/\,n^r$
  - some, $(s\,/\,iv^i)\,/\,n^i$
  - no, $(s\,/\,iv^r)\,/\,n^r$
  - broccoli, np
  - likes, $iv\,/\,np^i$
  - who, $(n\backslash n^i)\,/\,iv^i$
  - without, $(n\backslash n^i)\,/\,np^r$
  - most, $(s\,/\,iv^i)\,/\,n^b$

The Algorithm

- First, mark each parent node in the derivation tree with the marking for the argument category of its functional child. I.e., if A has children A/B^m and B, then A gets marking m.
- Second, from the root up, compute the polarity markings:
  - The root is assigned $+$.
  - Having marked node N with m, mark the functional child of N with m and the argument child with $f(m)$ where f is the category marking on N.

Example 2 (again)



Example 2 (again)

without ► $(n\backslash n^{i})/np^{r}$ broccoli ► np

without broccoli ► $n\backslash n^{i}$

candidate ► n r

candidate without broccoli ► n

no ► $(s/iv^{r})/n^{r}$ i

no candidate without broccoli ► s/iv^{r}

likes ► iv/np^{i} Theo ► np

likes Theo ► iv

r i

no candidate without broccoli likes Theo ► s

r

$$\Downarrow$$

without ► $(n\backslash n^{i})/np^{r}$ broccoli ► np

$-$ $+$

without broccoli ► $n\backslash n^{i}$

candidate ► n

$-$ $-$

candidate without broccoli ► n

no ► $(s/iv^{r})/n^{r}$

$+$ $-$

no candidate without broccoli ► s/iv^{r}

likes ► iv/np^{i} Theo ► np

$-$ $-$

likes Theo ► iv

$+$ $-$

no candidate without broccoli likes Theo ► s

$+$

Summary

- AB categorial grammar can be extended in a number of ways. CCG is one notable, elegant extension. Lambek Calculus is another.
- Lambek Calculus is motivated by the idea of thinking of syntactic derivation as logical proof. With this comes a very close correspondence between syntax and semantics *via* the Curry-Howard Correspondence.
- The idea behind the Monotonicity Calculus of van Benthem and Sánchez-Valencia is to forget about the λ-terms, but inject part of the semantics into the syntax. In particular monotonicity / antitonicity information is marked in the category assignments.
- The main workhorse of the Monotonicity Calculus is the polarity marking algorithm. The result is a marked expression which can be used to derive monotonicity inferences, based on background information about relations among subexpressions.
- Thus we have two proof systems working simultaneously: one to derive grammatical expressions, and one to derive inferential relations between grammatical expressions.

J. van Benthem. *Language in Action: Categories, Lambdas and Dynamic Logic*. Studies in Logic 130. Elsevier, Amsterdam, 1991.

R. Bernardi. *Reasoning with Polarity in Categorial Type Logics*, Ph.D. Thesis, UiL-OTS, Utrecht University, 2002.

J. van Eijk. 'Natural Logic for Natural Language', in B. ten Cate and H. Zeevat (eds.) *6th International Tbilisi Symposium on Logic, Language, and Computation*, Springer, 2007.

V. Sánchez-Valencia. *Studies on Natural Logic and Categorial Grammar*. Ph.D. Thesis, University of Amsterdam, 1991.