Categorial Grammar

Raffaella Bernardi

Contents

1. Recognition Device

- **Aim:** To build a language recognition device.
- **Who:** Lesniewski (1929), Ajdukiewicz (1935), Bar-Hillel (1953).
- **How:** Linguistic strings are seen as the result of function applications starting from the categories assigned to lexicon items.

2. Classical Categorial Grammar

- **Language:** Given a set of basic categories ATOM, the set of categories CAT is the smallest set such that:
  - ▷ if $X \in \mathsf{ATOM}$, then $X \in \mathsf{CAT}$;
  - ▷ if $X, Y \in \mathsf{ATOM}$, then $X/Y, Y\backslash X \in \mathsf{CAT}$
- **Rules:** The above categories can be composed by means of functional application rules

$$X/Y, Y \Rightarrow X \qquad \mathrm{MP_r}$$
$$Y, Y\backslash X \Rightarrow X \qquad \mathrm{MP_l}$$

$$\frac{X/Y \quad Y}{X}\,[\mathrm{MP_r}] \qquad \frac{Y \quad Y\backslash X}{X}\,[\mathrm{MP_l}]$$

3. Classical Categorial Grammar. Examples

Given $\mathsf{ATOM} = \{np, s, n\}$, we can build the following lexicon:

Lexicon

John, Mary	$\in$	np	the	$\in$	np/n
student	$\in$	n	some	$\in$	$(s/(np\backslash s))/n$
walks	$\in$	$np\backslash s$			
sees	$\in$	$(np\backslash s)/np$			

Analysis

John walks $\in s$? $\quad\rightsquigarrow\quad np, np\backslash s \Rightarrow s$? $\quad$ Yes

$$\frac{np \quad np\backslash s}{s}\,[\mathrm{MP_l}]$$

John sees Mary $\in s$? $\quad\rightsquigarrow\quad np, (np\backslash s)/np, np \Rightarrow s$? $\quad$ Yes

$$\frac{np \quad \dfrac{(np\backslash s)/np \quad np}{np\backslash s}\,[\mathrm{MP_r}]}{s}\,[\mathrm{MP_l}]$$

who knows Lori $\in n\backslash n$? $\quad\rightsquigarrow\quad (n\backslash n)/(np\backslash s), (np\backslash s)/np, np \Rightarrow n\backslash n$?

$$\dfrac{\dfrac{\text{who}}{(n\backslash n)/(np\backslash s)} \quad \dfrac{\dfrac{\text{knows}}{(np\backslash s)/np} \quad \dfrac{\text{Lori}}{np}}{np\backslash s}[\text{MP}_\text{r}]}{n\backslash n}[\text{MP}_\text{r}]$$

which Sara wrote $[\ldots] \in n\backslash n$?

Modus ponens corresponds to functional application.

$$\frac{X/Y : t \quad Y : r}{X : t(r)}[\text{MP}_\text{r}] \qquad\qquad \frac{Y : r \quad Y\backslash X : t}{X : t(r)}[\text{MP}_\text{l}]$$

Example

$$\frac{np : \texttt{john} \quad np\backslash s : \texttt{walk}}{s : \texttt{walk}(\texttt{john})}[\text{MP}_\text{l}]$$

$$np\backslash s : \lambda x.\texttt{walk}(x) \quad (\lambda x.\texttt{walk}(x))(\texttt{john}) \rightsquigarrow_{\lambda-\text{conv.}} \texttt{walk}(\texttt{john})$$

$$\dfrac{np : \texttt{john} \qquad \dfrac{(np\backslash s)/np : \texttt{know} \quad np : \texttt{mary}}{np\backslash s : \texttt{know}(\texttt{mary})}\,[\mathrm{MP_r}]}{s : \texttt{know}(\texttt{mary})(\texttt{john})}\,[\mathrm{MP_l}]$$

4. Logic Grammar

- **Aim:** To define the logic behind CG.
- **How:** Considering categories as formulae; $\backslash, /$ as logic connectives.
- **Who:** Jim Lambek [1958]

Lambek Calculus (Rules): Natural Deduction proof format [Elimination and Introduction rules]

Besides functional applications rules – which correspond to the elimination of $\backslash, /$ – we have their introduction rules. $\Gamma \vdash A$ means that A derives from Γ; Γ, Δ stand for structures, A, B, C for logic formulae.

$$\frac{\Delta \vdash B/A \quad \Gamma \vdash A}{\Delta, \Gamma \vdash B}\,[/\mathrm{E}] \qquad \frac{\Gamma \vdash A \quad \Delta \vdash A\backslash B}{\Gamma, \Delta \vdash B}\,[\backslash\mathrm{E}]$$

$$\frac{\Delta, B \vdash C}{\Delta \vdash C/B}\,[/\mathrm{I}] \qquad \frac{B, \Delta \vdash C}{\Delta \vdash B\backslash C}\,[\backslash\mathrm{I}]$$

5. Lambek calculus. Examples

which Sara wrote $\in n\backslash n$?

$$\dfrac{\text{which} \vdash (n\backslash n)/(s/np) \qquad \dfrac{\dfrac{\text{Sara} \vdash np \qquad \dfrac{\text{wrote} \vdash (np\backslash s)/np \quad [np \vdash np]^1}{\text{wrote } np \vdash np\backslash s}\,[/E]}{\text{Sara wrote } np \vdash s}\,[\backslash E]}{\text{Sara wrote} \vdash s/np}\,[/I]^1}{\text{which Sara wrote} \vdash n\backslash n}\,[/E]$$

The logical formulas built from $(\backslash, \bullet /)$ are interpreted using Kripke Models as below:

$$\begin{array}{lcl} V(A \bullet B) & = & \{z \mid \exists x \exists y[R^3zxy \;\&\; x \in V(A) \;\&\; y \in V(B)]\} \\ V(C/B) & = & \{x \mid \forall y \forall z[(R^3zxy \;\&\; y \in V(B)) \Rightarrow z \in V(C)]\} \\ V(A\backslash C) & = & \{y \mid \forall x \forall z[(R^3zxy \;\&\; x \in V(A)) \Rightarrow z \in V(C)]\} \end{array}$$

NL is sound and complete with respect to Kripke models.

Extractions are accounted for by means of introduction rules.

$$\frac{\text{john} \in np}{np \vdash np}\,\text{Lex} \quad \rightsquigarrow \quad \text{john} \vdash np$$

6. Lambek calculus. Semantics

$$\dfrac{\dfrac{\text{john} \vdash np : \mathtt{john} \quad [P \vdash np\backslash s : P]^1}{\text{john } P \vdash s : P(\mathtt{john})}[\backslash E]}{\text{john} \vdash s/(np\backslash s) : \lambda P.P(\mathtt{john})}[/I]^1$$

$$\dfrac{\dfrac{np \vdash np : \mathtt{john} \qquad \dfrac{\text{knows} \vdash (np\backslash s)/np : \mathtt{know} \quad [z \vdash np : z]^1}{\text{john knows } z \vdash np\backslash s : \mathtt{know}(z)(\mathtt{john})}[/E]}{\text{john knows } z \vdash s : \mathtt{know}(z)(\mathtt{john})}[\backslash E]}{\text{john knows} \vdash s/np : \lambda z.\mathtt{know}(z)(\mathtt{john})}[/I]^1$$

$$\Downarrow$$

The introduction rules correspond to λ-abstraction.

7. Lambek calculus. Advantages

- **Hypothetical reasoning:** Having added $[\backslash I], [/I]$ gives the system the right expressiveness to reason about hypothesis and abstract over them.
- **Curry Howard Correspondence:** Curry-Howard correspondence holds between proofs and terms. This means that parsed structures are assigned an interpretation into a model via the connection 'categories-terms'.
- **Logic:** We have moved from a grammar to a logic. Hence its behavior can be studied. The system is sound, complete and decidable.

8. Derivations

$$\dfrac{\dfrac{A \vdash B}{\langle A \rangle \vdash \Diamond B}\,[\Diamond R]}{\Diamond A \vdash \Diamond B}\,[\Diamond L] \qquad \dfrac{\dfrac{A \vdash A}{\langle \Box^{\downarrow} A \rangle \vdash A}\,[\Box^{\downarrow} L]}{\Diamond \Box^{\downarrow} A \vdash A}\,[\Diamond L]$$

$$\dfrac{\dfrac{A \vdash B}{\langle \Box^{\downarrow} A \rangle \vdash B}\,[\Box^{\downarrow} L]}{\Box^{\downarrow} A \vdash \Box^{\downarrow} B}\,[\Box^{\downarrow} R] \qquad \dfrac{\dfrac{A \vdash A}{\langle \Box^{\downarrow} A \rangle \vdash A}\,[\Diamond R]}{A \vdash A}\,[\Box^{\downarrow} R]$$

$$\dfrac{\dfrac{A \vdash A}{(A)^0 \vdash \sharp A}\,[(\cdot)^0 L]}{A \vdash {}^0((A)^0)}\,[{}^0(\cdot) R] \qquad \dfrac{\dfrac{A \vdash A}{{}^0(A) \vdash \flat A}\,[{}^0(\cdot) L]}{A \vdash ({}^0(A))^0}\,[(\cdot)^0 R]$$

9. Residuated and Galois Connected Functions

Remark 2 Let $\mathcal{B}'$ be a poset s.t. $\mathcal{B}' = (B, \sqsubseteq'_B)$ where $x \sqsubseteq'_B y \stackrel{\text{def}}{=} y \sqsubseteq_B x$, and $h : B \rightarrow A$. If (f, h) is a residuated pair with respect to $\sqsubseteq_A$ and $\sqsubseteq'_B$, then it's Galois connected with respect to $\sqsubseteq_A$ and $\sqsubseteq_B$.

$$b \sqsubseteq_B f(a) \quad \text{iff} \quad f(a) \sqsubseteq'_B b \quad \text{iff} \quad a \sqsubseteq_A h(b)$$

Recall Consider two posets $\mathcal{A} = (A, \sqsubseteq_A)$ and $\mathcal{B} = (B, \sqsubseteq_B)$, and functions $f : A \rightarrow B$, $g : B \rightarrow A$. The pair (f, g) is said to be **residuated** iff $\forall a \in A, b \in B$

$$[RES_1] \qquad f(a) \sqsubseteq_B b \quad \text{iff} \quad a \sqsubseteq_A g(b)$$

The pair (f, g) is said to be **Galois connected** iff $\forall a \in A, b \in B$

$$[GC_1] \qquad b \sqsubseteq_B f(a) \quad \text{iff} \quad a \sqsubseteq_A g(b)$$

10. Interpretation of the Constants

$$
\begin{aligned}
V(\Diamond A) &= \{x \mid \exists y(R^2_\Diamond xy \ \& \ y \in V(A)\} \\
V(\Box^\downarrow A) &= \{x \mid \forall y(R^2_\Diamond yx \ \Rightarrow y \in V(A)\}
\end{aligned}
$$

$$
\begin{aligned}
V({}^{\mathbf{0}}A) &= \{x \mid \forall y(y \in V(A) \ \Rightarrow \neg R^2_0 yx\} \\
V(A^{\mathbf{0}}) &= \{x \mid \forall y(y \in V(A) \ \Rightarrow \neg R^2_0 xy\}
\end{aligned}
$$

$$
\begin{aligned}
V(A \bullet B) &= \{z \mid \exists x \exists y[R^3 zxy \ \& \ x \in V(A) \ \& \ y \in V(B)]\} \\
V(C/B) &= \{x \mid \forall y \forall z[(R^3 zxy \ \& \ y \in V(B)) \Rightarrow z \in V(C)]\} \\
V(A \backslash C) &= \{y \mid \forall x \forall z[(R^3 zxy \ \& \ x \in V(A)) \Rightarrow z \in V(C)]\}
\end{aligned}
$$

11. Nonveridical Functions

definition [(Non)veridical functions (II)]

Let $(\overrightarrow{a}_n, t)$ stand for a boolean type $(a_1, (\ldots(a_n, t)\ldots))$ where $a_1, \ldots, a_n$ are arbitrary types and $0 \leq n$. Let $f_{(\overrightarrow{a},t)}$ be a constant.

1. The expression represented by f is **veridical** in its i-argument, if a_i is a boolean type, **i.e**. $a_i = (\overrightarrow{b}, t)$, and $\forall \mathcal{M}, g$

   $$[\![f(x_{a_1}, \ldots, x_{a_{i-1}}, x_{(\overrightarrow{b},t)}, x_{a_{i+1}}, \ldots, x_{a_n})]\!]_{\mathcal{M},g} = 1 \text{ entails } [\![\exists \overrightarrow{y}_{\overrightarrow{b}} .x_{(\overrightarrow{b},t)}(\overrightarrow{y}_{\overrightarrow{b}})]\!]_{\mathcal{M},g} = 1.$$

   Otherwise f is nonveridical.

2. A nonveridical function represented by $f_{(\overrightarrow{a},t)}$ is **antiveridical** in its i-argument, if $a_i = (\overrightarrow{b}, t)$ and $\forall \mathcal{M}, g$

   $$[\![f(x_{a_1}, \ldots, x_{a_{i-1}}, x_{(\overrightarrow{b},t)}, x_{a_{i+1}}, \ldots, x_{a_n})]\!]_{\mathcal{M},g} = 1 \text{ entails } [\![\neg\exists. \overrightarrow{y}_{\overrightarrow{b}}\ x_{(\overrightarrow{b},t)}(\overrightarrow{y}_{\overrightarrow{b}})]\!]_{\mathcal{M},g} = 1.$$

Notice that the base case of $a_i = t$ is obtained by taking $\overrightarrow{y}$ empty.

12. Dutch

In [van Wouden] it is shown that in Dutch polarity items are sensitive to downward monotonicity. Among downward monotone functions we can distinguish the sets below:

antimorphic	antiadditive	downward monotone
$f(X \cap Y) = f(X) \cup f(Y)$ $f(X \cup Y) = f(X) \cap f(Y)$	$f(X) \cup f(Y) \subseteq f(X \cap Y)$ $f(X \cup Y) = f(X) \cap f(Y)$	$f(X) \cup f(Y) \subseteq f(X \cap Y)$ $f(X \cup Y) \subseteq f(X) \cap f(Y)$
not	nobody, never, nothing	few, seldom, hardly

13. Classification of NPIs in Dutch

This classification effects the classification of polarity items.

Negation	NPIs			PPIs		
	strong	medium	weak	strong	medium	weak
Minimal (DM)	−	−	+	−	+	+
Regular (AA)	−	+	+	−	−	+
Classical (AM)	+	+	+	−	−	−
	mals (tender)	ook maar (anything)	hoeven (need)	allerminst (not-at-all)	een beetje (a bit)	nog (still)

NPIs are **licensed**, whereas PPIs are **antilicensed** by a certain property among the ones characterizing downward monotone functions. From this it follows that

- a NPI licensed by the property of a function in DM will be grammatical also when composed with any functions belonging to a stronger set.
- if a PPI is 'allergic' to one specific property shared by the functions of a certain set, it will be ungrammatical when composed with them, but compatible with any other function in a weaker set which does not have this property.

14. Antilicensing Relation

A weak PPI is antilicensed by antimorphicity, therefore it can be constructed with any expression in a set equal to or bigger than AA, $B/^{\mathbf{0}}AA$. A medium PPI is antilicensed by antiadditivity, therefore it can be in construction with any expression in a set equal to or bigger than DM, $B/^{\mathbf{0}}DM$. From these types the following inferences derive.

Let $AM \longrightarrow AA \longrightarrow DM$.

$$\dfrac{\text{MPPI} \vdash B/^{\mathbf{0}}(DM) \qquad \dfrac{DM \vdash DM}{{}^{\mathbf{0}}(DM) \vdash {}^{\mathbf{0}}(DM)}[\downarrow \text{Mon}]}{\text{MPPI} \circ {}^{\mathbf{0}}(DM) \vdash A} \qquad \dfrac{\text{MPPI} \vdash B/^{\mathbf{0}}(DM) \qquad \dfrac{AA \vdash AA}{{}^{\mathbf{0}}(AA) \vdash {}^{\mathbf{0}}(DM)}*}{{}^{*}\text{MPPI} \circ {}^{\mathbf{0}}(AA) \vdash B}$$

$$\dfrac{\text{WPPI} \vdash B/^{\mathbf{0}}(AA) \qquad \dfrac{AA \vdash AA}{{}^{\mathbf{0}}(AA) \vdash {}^{\mathbf{0}}(AA)}[\downarrow \text{Mon}]}{\text{WPPI} \circ {}^{\mathbf{0}}(AA) \vdash A} \qquad \dfrac{\text{WPPI} \vdash B/^{\mathbf{0}}(AA) \qquad \begin{matrix}\dfrac{DM \vdash DM}{{}^{\mathbf{0}}(DM) \vdash {}^{\mathbf{0}}(DM)}[\downarrow \text{Mon}] \\ \vdots \\ {}^{\mathbf{0}}(DM) \vdash {}^{\mathbf{0}}(AA)\end{matrix}}{\text{WPPI} \circ {}^{\mathbf{0}}(DM) \vdash B}$$