Introduction aux grammaires catégorielles

Université de Nice - Sophia Antipolis

15 décembre 2006

Notation polonaise

Leśniewski puis Ajdukewicz

- catégories primitives : s, n
- catégories dérivées : si α, β sont deux catégories, α/β est une catégorie
- modus ponens :

$$\frac{(\alpha/\beta) \quad \beta}{\alpha} /_e$$

Notation polonaise

Leśniewski puis Ajdukewicz

Exemple

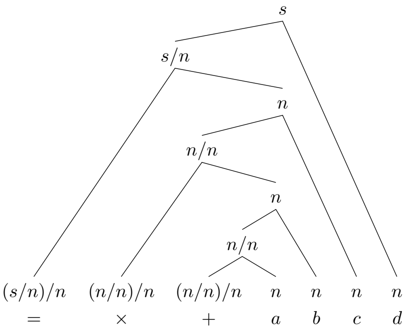
symbole	types
a, b, c, d	n
$+$	$(n/n)/n$
$\times$	$(n/n)/n$
$=$	$(s/n)/n$

Notation polonaise

Leśniewski puis Ajdukewicz

Exemple

symbole	types
a, b, c, d	n
$+$	$(n/n)/n$
$\times$	$(n/n)/n$
$=$	$(s/n)/n$



Grammaires AB catégorielles (BCG)

Bar-Hillel

- nouvel opérateur : $\alpha\backslash\beta$
- nouveau modus ponens :

$$\frac{\alpha \quad (\alpha\backslash\beta)}{\beta}\backslash_e$$

Exercice

mot	types
John	n
She	$s/(n\backslash s)$
likes	$(n\backslash s)/n$
fresh	n/n
him	$(s/n)\backslash s$
milk	n

Équivalence avec les CFGs

Bar-Hillel, Gaifman et Shamir

Théorème

Toute grammaire AB catégorielle est fortement équivalente à une grammaire algébrique sous forme normale de Chomsky.

Théorème

Toute grammaire algébrique sans production du vide sous forme normale de Greibach est fortement équivalente à une grammaire AB catégorielle.

Limites des grammaires AB catégorielles

mot	types
She	$s/(n\backslash s)$
likes	$(n\backslash s)/n$
him	$(s/n)\backslash s$

Et la phrase “She likes him.” ?

Un détour par la déduction naturelle

Gentzen

$$\frac{}{A \vdash A}\,\mathrm{id}$$

$$\frac{\Gamma, B \vdash A}{\Gamma \vdash B \longrightarrow A}\,{\longrightarrow_i} \qquad \frac{\Gamma \vdash B \longrightarrow A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A}\,{\longrightarrow_e}$$

$$\frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B}\,{\wedge_i} \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}\,{\wedge_{e1}} \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}\,{\wedge_{e2}}$$

Un détour par la déduction naturelle

Gentzen

$$\frac{\Gamma, B \vdash A}{\Gamma \vdash B \rightarrow A}\rightarrow_i \qquad \frac{\Gamma \vdash B \rightarrow A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A}\rightarrow_e$$

Version non commutative ?

Le calcul de Lambek (LCG)

Lambek

$$\frac{}{A \vdash A}\text{id}$$

$$\frac{A, \Gamma \vdash B}{\Gamma \vdash A \backslash B}\backslash_i, \Gamma \neq \varepsilon \qquad \frac{\Gamma \vdash A \quad \Delta \vdash A \backslash B}{\Gamma, \Delta \vdash B}\backslash_e$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash B / A}/_i, \Gamma \neq \varepsilon \qquad \frac{\Delta \vdash B / A \quad \Gamma \vdash A}{\Delta, \Gamma \vdash B}/_e$$

$$\frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma, \Delta \vdash A \cdot B}\cdot_i \qquad \frac{\Gamma \vdash A \cdot B \quad \Delta, A, B, \Delta' \vdash C}{\Delta, \Gamma, \Delta' \vdash C}\cdot_e$$

Exemple

"She likes him."

$$\dfrac{\dfrac{\dfrac{\dfrac{\text{She}}{s/(n\backslash s)\vdash s/(n\backslash s)}\,\text{id} \quad \dfrac{\dfrac{\text{likes}}{(n\backslash s)/n\vdash (n\backslash s)/n}\,\text{id} \quad \dfrac{}{n\vdash n}\,\text{id}}{(n\backslash s)/n, n\vdash n\backslash s}\,/_e}{s/(n\backslash s),(n\backslash s)/n,n\vdash s}\,/_e}{s/(n\backslash s),(n\backslash s)/n\vdash s/n}\,/_i \quad \dfrac{\text{him}}{(s/n)\backslash s\vdash (s/n)\backslash s}\,\text{id}}{s/(n\backslash s),(n\backslash s)/n,(s/n)\backslash s\vdash s}\,\backslash_e$$

Équivalence avec les CFGs

Pentus

- Chomsky (1963) : conjecture : les grammaires de Lambek génèrent des langages algébriques
- Cohen (1967) : toute grammaire AB catégorielle est une grammaire de Lambek
- Pentus (1993) : toute grammaire de Lambek est équivalente à une grammaire algébrique

Théorème (Pentus)

Décider si une phrase est acceptée par une grammaire de Lambek est un problème NP-complet.

Sémantique transparente

Montague

Retour du commutatif :

- les phrases *s* décrivent des propositions
- les noms *n* décrivent des entités
- les catégories α/β et $\beta\backslash\alpha$ décrivent des fonctions $\alpha \rightarrow \beta$

Exemple

mot	syntaxe	sémantique
John	n	John'
likes	$(n\backslash s)/n$	$\lambda x \lambda y.\text{likes}' xy$
fresh	n/n	$\lambda x.\text{fresh}' x$
milk	n	milk'

$$\dfrac{\dfrac{\text{John}}{n \vdash n : \text{John}'}\,\text{id} \qquad \dfrac{\dfrac{\text{likes}}{(n\backslash s)/n \vdash (n\backslash s)/n : \lambda x \lambda y.\text{likes}' xy}\,\text{id} \quad \dfrac{\text{milk}}{n \vdash n : \text{milk}'}\,\text{id}}{(n\backslash s)/n, n \vdash n\backslash s : \lambda y.\text{likes}' \text{milk}' y}\,/e}{n, (n\backslash s)/n, n \vdash s : \text{likes}' \text{milk}' \text{John}'}\,\backslash e$$

Grammaires combinatoires (CCGs)

Steedman

$$\frac{X/Y : f \quad Y : a}{X : fa} > \qquad \frac{Y : a \quad Y\backslash X : f}{X : fa} <$$

$$\frac{X/Y : f \quad Y/Z : g}{X/Z : \lambda x.f(gx)} >B \qquad \frac{X\backslash Y : f \quad Y\backslash Z : g}{X\backslash Z : \lambda x.g(fx)} <B$$

$$\frac{X : x}{Y/(X\backslash Y) : \lambda f.fx} >T \qquad \frac{X : x}{(Y/X)\backslash Y : \lambda f.fx} <T$$

$$\frac{(X/Y)/Z : f \quad Y/Z : g}{X/Z : \lambda x.fx(gx)} >S \qquad \frac{Z\backslash Y : f \quad Z\backslash(Y\backslash X) : g}{Z\backslash X : \lambda x.gx(fx)} <S$$

Grammaires combinatoires (CCGs)

Steedman

$$\frac{X/Y : f \quad Y : a}{X : fa} >$$

$$\frac{Y : a \quad Y\backslash X : f}{X : fa} <$$

$$\frac{X/Y : f \quad Y/Z : g}{X/Z : \lambda x.f(gx)} >B$$

$$\frac{X\backslash Y : f \quad Y\backslash Z : g}{X\backslash Z : \lambda x.g(fx)} <B$$

$$\frac{X : x}{Y/(X\backslash Y) : \lambda f.fx} >T$$

$$\frac{X : x}{(Y/X)\backslash Y : \lambda f.fx} <T$$

$$\frac{(X/Y)/Z : f \quad Y/Z : g}{X/Z : \lambda x.fx(gx)} >S$$

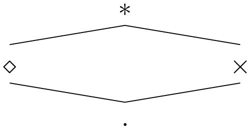
$$\frac{Z\backslash Y : f \quad Z\backslash(Y\backslash X) : g}{Z\backslash X : \lambda x.gx(fx)} <S$$

Exercice

mot	type
John	n
might	$(n\backslash s)/(n\backslash s)$
like	$(n\backslash s)/n$
fresh	n/n
milk	n

Version multi-modale (MMCCG)

Baldridge et Kruijff



- plusieurs modes pour les opérations $/$ et $\backslash$ ($/_*$, $/_\times$, $/_\diamond$, ...)
- chaque mode permet ou non l'emploi de certaines règles ($X/_*Y \quad Y \Rightarrow X$, $X/_\diamond Y \quad Y/_\diamond Z \Rightarrow X/_\diamond Z$, ...)
- contrôle fin de la grammaire

Formalisme faiblement contextuel

Joshi, Vijay-Shanker et Weir

- Équivalence faible avec les TAGs.
- Complexité d'analyse : polynomiale.