Lambek Grammatiken

Kilian Evang, Christian Wurm

Düsseldorf, 12.2022

Kategorialgrammatiken

Kategorien in CG waren immer wie folgt definiert: *Var* ist die (endliche/unendliche) Menge der atomaren Kategorien

1. Falls $\alpha \in Var$, dann ist $\alpha \in Cat(Var)$
2. Falls $\alpha, \beta \in Cat(Var)$ dann ist α/β, $\beta\backslash\alpha \in Cat(Var)$

Applikative CG

Definition CG

Eine (applikative) Kategorialgrammatik ist ein Tupel $(Var, \mathtt{Lex}, S, \zeta)$

Das bedeutet: wir haben

1. *Var* und damit $Cat(Var)$, also die Menge der (atomaren) syntaktischen Kategorien
2. $\mathtt{Lex}$, das Lexikon. Das können Worte sein (eher linguistisch) oder Buchstaben (eher formal-sprachlich). Im letzteren Fall wäre $\mathtt{Lex} = \Sigma$.
3. S ist eine *ausgezeichnete Kategorie*, wie man sie von CFG kennt
4. ζ ist die Typzuweisung.

Applikative CG

Applikative CG haben zwei Regeln:

$$\frac{\alpha/\beta \quad \beta}{\alpha}\ /E \qquad \frac{\beta \quad \beta\backslash\alpha}{\alpha}\ \backslash E$$

Eine **Ableitung** ist induktiv definiert:

- jede Instanz von $/E$, $\backslash E$ ist eine Ableitung, und
- falls $\frac{T}{\beta}$, $\frac{T'}{\alpha/\beta}$ Ableitungen sind, dann auch $\dfrac{\dfrac{T'}{\alpha/\beta} \quad \dfrac{T}{\beta}}{\alpha}$,
- ebenso mit $\backslash$

Der AB-Kalkül

(ax) $\quad \alpha \vdash \alpha$

$(\mathbf{I} - /) \quad \dfrac{\Gamma, \alpha \vdash \beta}{\Gamma \vdash \beta/\alpha} \qquad (\mathbf{I} - \backslash) \quad \dfrac{\alpha, \Gamma \vdash \beta}{\Gamma \vdash \alpha\backslash\beta}$

$(/ - \mathbf{I}) \quad \dfrac{\Delta, \beta, \Theta \vdash \gamma \quad \Gamma \vdash \alpha}{\Delta, \beta/\alpha, \Gamma, \Theta \vdash \gamma}$

$(\backslash - \mathbf{I}) \quad \dfrac{\Delta, \beta, \Theta \vdash \gamma \quad \Gamma \vdash \alpha}{\Delta, \Gamma, \alpha\backslash\beta, \Theta \vdash \gamma}$

(schnitt) $\quad \dfrac{\Gamma, \alpha, \Gamma' \vdash \beta \quad \Delta \vdash \alpha}{\Gamma, \Delta, \Gamma' \vdash \beta}$

Natürliches Schließen

$$(\mathrm{E}/)\ \frac{\alpha/\beta \quad \beta}{\alpha} \qquad\qquad (\mathrm{E}\backslash)\ \frac{\beta \quad \beta\backslash\alpha}{\alpha}$$

$$(\mathrm{I}/)\ \frac{\begin{array}{c}\underline{\ldots \quad [\beta]}\\ \vdots \\ \alpha\end{array}}{\alpha/\beta} \qquad\qquad (\mathrm{I}\backslash)\ \frac{\begin{array}{c}\underline{[\beta] \quad \ldots}\\ \vdots \\ \beta\end{array}}{\beta\backslash\alpha}$$

Die Intro-Regeln sind etwas schwer zu interpretieren: β ist hier eine **Annahme**, die wir mit der Regel **auflösen**. Die Annahme muss dabei am linken/rechten Rand des Beweises stehen!

Lambek

Der Lambek-Kalkül erweitert unsere Sprache mit einen weiteren Konnektor $\bullet$:

- Falls $\alpha, \beta \in \mathit{Var}$, dann $\alpha \bullet \beta \in \mathit{Var}$

Im Gentzen Kalkül gibt es dazu zwei Regeln:

$$\frac{\Gamma, \alpha, \beta, \Delta \vdash \gamma}{\Gamma, \alpha \bullet \beta, \Delta \vdash \gamma} \bullet\mathbf{I} \qquad \frac{\Gamma \vdash \alpha \quad \Delta \vdash \beta}{\Gamma, \Delta \vdash \alpha \bullet \beta} \mathbf{I}\bullet$$

Damit ist natürlich klar: im Lambek-Kalkül kann *alles* eine Konstituente sein.

Lambek - natürliches Schließen

Es gibt auch natürliches Schließen zu Lambek:

$$\frac{\alpha \quad \beta}{\alpha \bullet \beta} \bullet\mathrm{I} \qquad\qquad \frac{\alpha \bullet \beta}{\alpha \qquad \beta} \bullet\mathrm{E}$$

Man beachte: das ganze ist dann kein Baum mehr, sondern ein gerichteter azyklischer Graph!

Lambek-Grammatiken

Eine Lambek-Grammatik $G = (Var, \mathtt{Lex}, S, \zeta)$ generiert ein Wort/Satz $\vec{w} = w_1 w_2 ... w_i \in \mathtt{Lex}^+$, falls gilt:

- es gibt einen Beweis im Lambek-Kalkül von $\alpha_1, ..., \alpha_i \vdash S$, und
- f.a. $j \in \{1, ..., i\}$ gilt: $\alpha_j \in \zeta(w_j)$

Lambek Grammatiken sind lustig, aber nicht besonders praktisch. Das hat zwei Gründe:

Lambek-Grammatiken

Lambek Grammatiken sind lustig, aber nicht besonders praktisch. Das hat zwei Gründe:

Theorem

(Pentus) Es gibt eine Lambek-Grammatik GL so dass $L(GL) = L$ gdw. L kontextfrei ist.

Das war 40 Jahre lang ein offenes Problem!

Theorem

$\vdash_L$ *ist NP-vollständig.*

Lambek-Grammatiken

Theorem

$\vdash_L$ *ist NP-vollständig.*

NB: CFG haben ein Erkennungs/Parsing Problem in $O(n^{2.7})$!

- Die Repräsentation ist also entscheidend für die Komplexität, nicht nur die Komplexitätsklasse!
- Das impliziert natürlich, dass die Repräsentationen nicht einfach übersetzbar sind!

Übung

Beweisen Sie folgende Sequenten:

1. $\alpha \vdash (\alpha \bullet \beta)/\beta$
2. $\alpha, (\beta/\beta) \vdash (\alpha \bullet \beta)/\beta$
3. $(\alpha \bullet \beta)\backslash\gamma \vdash \beta\backslash(\alpha\backslash\gamma)$

Zusammenfassung

Der Lambek-Kalkül ist superspannend, aber nicht als Kategorialgrammatik!