

Lösungen zum Aufgabenblatt 7

Syntax natürlicher Sprachen

Universität München, CIS, WS 2016/17

Hans Leiß

Abgabetermin: Mi, 12.1.2017

Aufgabe 7.1 Sei $G = (\Sigma, T, S)$ die Kategorialgrammatik nach Ajdukiewicz/Bar-Hillel mit

- $\Sigma := \{a, b, c\}$
- $T(a) := \{A/B, C \setminus (A/B)\}$, $T(b) := \{B, (C/B) \setminus A, A \setminus B\}$, $T(c) := \{B \setminus C, A/A, C/B\}$
- $S := A/B$

Die Typisierung T kann man als eine Menge von Axiomen oder Annahmen verstehen, nämlich

$$a : A/B, \quad a : C \setminus (A/B), \quad b : B, \quad b : (C/B) \setminus A, \quad b : A \setminus B, \quad c : B \setminus C, \quad c : A/A, \quad c : C/B.$$

- (a) Zeigen Sie durch Angeben eines Beweises, daß $cbbca \in L(G)$. Der Beweis muß $cbbca : S$ aus den Axiomen $x : X$ für $x \in \Sigma$ und $X \in T(x)$ und den Schlußregeln

$$\frac{x : X \quad y : X \setminus Y}{xy : Y} (\backslash) \quad \text{und} \quad \frac{y : Y/X \quad x : X}{yx : Y} (/)$$

zeigen.

- (b) Im Vorlesungsteil wurde skizziert, wie man zu einer Kategorialgrammatik $G \in KG$ eine kontextfreie Grammatik $G' \in CFG$ findet, die dieselbe Sprache definiert. Führen Sie das konkret für die Kategorialgrammatik G durch.

Beschreiben Sie die Systematik Ihres Vorgehens: Welche Nonterminale und welche Grammatikregeln werden warum eingeführt?

Geben Sie einen Beweis für $bca \in L(G)$ und einen Syntaxbaum für $bca \in L(G')$ an.

Lösung von Aufgabe 7.1 Sei G die obige Kategorialgrammatik.

- (a) Ein Beweis für $cbbca \in L(KG)$ ist der folgende:

$$\frac{\frac{\frac{c : A/A \quad \frac{c : C/B \quad b : (C/B) \setminus A}{cb : A} (\backslash)}{ccb : A} (/)}{cbb : B} (\backslash) \quad b : A \setminus B}{cbbc : C} (\backslash) \quad c : B \setminus C}{cbbca : A/B} (\backslash) \quad a : C \setminus (A/B) (\backslash)$$

(b) Die entsprechende kontextfreie Grammatik $G' = (\Sigma', N', P', S')$ wird wie folgt gebildet:

- (i) Die Menge Σ' der Terminale von G' ist $\Sigma' := \Sigma = \{a, b, c\}$.
- (ii) Die Menge N der Nonterminale bilden die in den Kategorisierungen $T(a), T(b), T(c)$ vorkommenden Kategorien (einschließlich der Unterausdrücke), also

$$\{A/B, A, B, C \setminus (A/B), C, (C/B) \setminus A, C/B, A \setminus B, B \setminus C, A/A\}.$$

Außenklammern sind weggelassen; d.h. (A/B) und A/B sind dasselbe Nonterminal.

- (iii) Das Startsymbol S' ist $S' := S = A/B$.
- (iv) Die Regelmengemenge P' umfaßt einerseits lexikalische Regeln, die den Kategorisierungen der Terminale entsprechen, also $X \rightarrow x$ für $X \in T(x)$:

$$\begin{array}{lll} A/B \rightarrow a & B \rightarrow b & B \setminus C \rightarrow c \\ C \setminus (A/B) \rightarrow a & (C/B) \setminus A \rightarrow b & A/A \rightarrow c \\ & A \setminus B \rightarrow b & C/B \rightarrow c \end{array}$$

Andererseits umfaßt P' die den Kürzungsregeln entsprechenden kontextfreien Regeln, für die Kategorien, die jetzt Nonterminale sind. Deutlichkeitshalber schreibe ich \cdot als Trenner auf den rechten Regelseiten:

$$\begin{array}{lll} A \rightarrow A/B \cdot B & A \rightarrow C/B \cdot (C/B) \setminus A & C \rightarrow B \cdot B \setminus C \\ A/B \rightarrow C \cdot C \setminus (A/B) & C \rightarrow C/B \cdot B & A \rightarrow A/A \cdot A \\ & B \rightarrow A \cdot A \setminus B & \end{array}$$

Die Sprache $L(G')$ bzw. $L(G)$ enthält bca , was mit der Ableitung und G' wie folgt gezeigt wird:

$$\begin{aligned} S = A/B &\Rightarrow C \cdot C \setminus (A/B) \\ &\Rightarrow C \cdot a \\ &\Rightarrow B \cdot B \setminus C \cdot a \\ &\Rightarrow b \cdot B \setminus C \cdot a \\ &\Rightarrow b \cdot c \cdot a \end{aligned}$$

Als Beweis mit G entspricht dem:

$$\frac{\frac{b : B \quad c : B \setminus C}{bc : C}(\setminus) \quad a : C \setminus (A/B)}{bca : A/B = S}(\setminus)$$

Aufgabe 7.2 Wir haben in der Tafelübungsstunde vom 21.12. eine DCG-Regel für Sätze mit zweistelligem Verb und einem optionalen Adverbial geschrieben:

```
s([def],[Temp,Mod,vz]) --> % def = Aussagesatz
    np([Pers,_Gen,Num],[nom]),
    v([[nom,Kas]],[Pers,Num,Mod,Temp]),
    { Temp = praes ; Temp = praet },
    np([_Pers2,_Gen2,_Num2],[Kas]),
    ([ ; advp([_Art],[ ]) ]).
```

Sei G die kontextfreie Grammatik aus der obigen Satzregel, der vereinfachten NP -Regel

$$\begin{aligned} \text{np}([\text{Pers}, \text{Gen}, \text{Num}], [\text{Kas}]) \rightarrow & \\ \text{det}([\text{Pers}, \text{Num}, \text{Dekl}], [\text{Gen}, \text{Kas}]), & \\ \text{n}([\text{Gen}], [\text{Num}, \text{Kas}]). & \end{aligned}$$

und den lexikalischen Regeln, die man zur Erkennung der Sätze *die Frau liest das Buch gründlich* und *die Frau liest das Buch oft* braucht, wobei die Merkmale auf die hierfür nötigen Fälle beschränkt seien (also $\text{Temp}=\text{praes}$, $\text{Mod}=\text{ind}$ usw.).

Unter advp zwei Arten von Adverbialen unterscheiden: solche, die das Verb (bzw. das Prädikat) modifizieren, wie *gründlich*: $\text{Adv}_{\text{modal}}$ und solche, die eine Aussage modifizieren, wie *oft*: Adv_{temp} in *die Frau liest das Buch oft* (\simeq *es ist oft der Fall, daß die Frau das Buch liest*).

Wie kann man diese Grammatik (ungefähr) mit den Mitteln der Kategorialgrammatik (KG) ausdrücken? Als Basiskategorien seien vorgesehen: N^{nom} , N^{akk} , NP^{nom} , NP^{akk} und S (implizit mit den Werten der anderen Merkmale, so, wie man sie für die verwendeten Wörter braucht).

Um die Typisierung T anzugeben, muß man sich überlegen:

- Welche KG-Kategorien haben *Frau* und *Buch*?
- Welche KG-Kategorie braucht man für die Artikel *die* und *das*?
- Welche KG-Kategorie entspricht der DCG-Kategorie $v([\text{nom}, \text{akk}], [\text{praes}, \text{ind}, \text{vz}])$, wenn kein Adverb benutzt wird? Was wäre dafür in der Typisierung von *liest* nötig? (Und was wäre $T(\textit{liest})$, wenn man auch die Verbletzstellung berücksichtigen würde?)
- Welche KG-Kategorie braucht man für das Adverb *oft*? Welche Kategorie hätten Sie für das Adverb *gründlich* erwartet?

Wie kann man mit dieser syntaktischen Typisierung T und den Regeln ($/$) und (\backslash) der Kategorialgrammatik beweisen, daß *die Frau liest das Buch*: S und *die Frau liest das Buch oft*: S ?

Satzmodifizierende Adverbien wie *oft* sollten andere Typen haben als verbmodifizierende Adverbien wie *gründlich*. Da im Beispiel aber *gründlich* nicht neben dem Verb steht, kann man mit dem naheliegenden Typ den Beispielsatz *die Frau liest das Buch gründlich* nicht beweisen. Was könnte man bei den Typannahmen ändern, damit dieser Beispielsatz von der KG als Satz erkannt wird, und wie sähe ein Beweis dafür aus?

Lösung von Aufgabe 7.2 Bei der Typisierung T der Wörter braucht man für eine Analyse der Beispielsätze mindestens:

- $\textit{Frau} : N^{\text{nom}}$, $\textit{Buch} : N^{\text{akk}}$,
- $\textit{die} : NP^{\text{nom}}/N^{\text{nom}}$, $\textit{das} : NP^{\text{akk}}/N^{\text{akk}}$,
- $\textit{liest} : (NP^{\text{nom}} \backslash S)/NP^{\text{akk}}$,
- $\textit{oft} : S \backslash S$.

Eine Analyse von *die Frau liest das Buch* wäre damit

$$\frac{\frac{\frac{die : NP^{nom}/N^{nom} \quad Frau : N^{nom}}{die Frau : NP^{nom}} \quad \frac{liest : (NP^{nom}\backslash S)/NP^{akk} \quad \frac{\vdots}{das Buch : NP^{akk}}}{liest das Buch : NP^{nom}\backslash S}}{die Frau liest das Buch : S}}{\backslash}$$

Mit dem Satzmodifikator $oft : S\backslash S$ kann man das fortsetzen und den erweiterten Satz analysieren:

$$\frac{\frac{\frac{\vdots}{die Frau liest das Buch : S} \quad oft : S\backslash S}{die Frau liest das Buch oft : S}}{\backslash}$$

Offenbar braucht man für andere Wortstellungen andere Typannahmen, z.B. $oft : S/S$ für die Verwendung von *oft* in *oft liest die Frau das Buch*, wobei hier S für Sätze mit Verberststellung verwendet wird.

Das Adverb *gründlich* sollte die Kategorie eines nachgestellten Modifikators von transitiven Verben haben, also

$$gründlich : (NP^{nom}\backslash S)/NP^{akk} \backslash (NP^{nom}\backslash S)/NP^{akk},$$

oder die eines nachgestellten Modifikators von Prädikaten oder Verbalphrasen $VP := NP^{nom}/S$,

$$gründlich : (NP^{nom}\backslash S) \backslash (NP^{nom}\backslash S).$$

Typisiert man es als Modifikator von Verbalphrasen, so kann man den Satz *die Frau liest das Buch gründlich* wie folgt analysieren:

$$\frac{\frac{\frac{\vdots}{die Frau : NP^{nom}} \quad \frac{\frac{\frac{\vdots}{liest das Buch : VP} \quad gründlich : VP\backslash VP}{liest das Buch gründlich : VP}}{die Frau liest das Buch gründlich : S}}{\backslash}$$

Mit $TV := VP/NP^{akk}$ und der Typisierung $gründlich : TV\backslash TV$ als Verbmodifikator kann man die folgende Wortstellung analysieren:

$$\frac{\frac{\frac{liest : TV \quad gründlich : TV\backslash TV}{liest gründlich : TV} \quad \frac{\frac{\vdots}{das Buch : NP^{akk}}}{liest gründlich das Buch : VP}}{\backslash}$$

Man würde natürlich gerne einen systematischen Zusammenhang zwischen diesen beiden Verwendungen ausdrücken: wenn man eine Funktion $f : A/B$ mit einem nachgestellten Modifikator $m : (A/B)\backslash(A/B)$ modifizieren und dann auf ein Argument $b : B$ anwenden kann, also

$$\frac{\frac{f : A/B \quad m : (A/B)\backslash(A/B)}{(f \cdot m) : A/B} \quad b : B}{((f \cdot m) \cdot b) : A},$$

dann kann man auch zuerst die Funktion auf das Argument anwenden und anschließend das Ergebnis modifizieren:

$$\frac{\frac{f : A/B \quad b : B}{(f \cdot b) : A} / \quad m : A \setminus A}{((f \cdot b) \cdot m) : A} \setminus.$$

Dazu müßte man ausdrücken können, daß jedes $m : (A/B) \setminus (A/B)$ auch den Typ $m : A \setminus A$ haben soll. Das können wir in der einfachen Kategorialgrammatik von Ajdukiewicz und Bar-Hillel aber nicht formulieren.