

Lösungen zum Aufgabenblatt 9

Syntax natürlicher Sprachen

Universität München, CIS, WS 2016/17

Hans Leiß

Abgabetermin: Mi, 25.1.2017

Aufgabe 9.1 Auf der Kursseite finden Sie eine Datei `saetze-mit-adverbialen.pl`, in der DCG-Regeln für einfache Sätze mit einem oder zwei Adverbialen angegeben sind. Dazu gibt es eine Datei von Beispielsätzen,

```
beispielsaetze-mit-adverbialen.pl
```

die festlegt, welche Syntaxregeln und welche lexikalischen Regeln geladen werden sollen:

```
% Grammatiken: Nominalphrasen, Adverbiale, Saetze
```

```
:- ['np.folien', % wegen multifile-Deklarationen zuerst!  
    adverbiale,  
    saetze-mit-adverbialen  
    ].
```

```
% Laden der Lexika für die beispieladverbiale.txt
```

```
:- [lexikon_np,  
    lexikon-beispielnominalphrasen,  
    lexikon-beispielsaetze,  
    lexikon-beispieladverbiale  
    ].
```

```
% Startsymbole überschreiben:
```

```
startsymbol(advp(_, [])).  
startsymbol(s(_, _)).
```

Außerdem gibt es eine Datei `beispielsaetze-mit-adverbialen.txt` mit ein paar Beispielen:

dann arbeiten sie.

wenn die Glocken läuten, arbeiten sie.

arbeiten sie, läuten die Glocken.

sie arbeiten nicht.
heute arbeiten sie nicht.
sie arbeiten heute nicht.
weil die Glocken läuten, arbeiten die Menschen nicht.

die Kinder lesen heute das Buch.
die Kinder lesen das Buch heute.
die Kinder lesen das Buch nicht.
die Kinder lesen heute nicht das Buch.
heute lesen die Kinder das Buch.
heute lesen die Kinder das Buch nicht.
nicht lesen die Kinder das Buch. % ungrammatisch!

Sie können dann in Prolog die Grammatik laden

```
?- [beispielsaetze-mit-adverbialen].
```

und mit

```
?- parse('beispielsaetze-mit-adverbialen.txt').
```

die Beispiele analysieren und die Syntaxbäume ausgeben lassen (analog mit `parsed/1` für die Ausgabe mit Graphik).

- (a) In `saetze-mit-adverbialen.pl` gibt es nur eine Satzregel für Sätze aus Vollverben mit einem Präpositionalobjekt:

```
% TODO: 2-stellige Verben mit Präpositionalobjekt:
```

```
s([def],[Temp,Mod,vz]) --> % def = Aussagesatz
  np([Pers,_Gen,Num],[nom]),
  v([[nom,Prep,Kas]],[Pers,Num,Mod,Temp]), % Verb mit Objekt-PP
  ( [] ; advp([Art1],[ ]) ),
  { praeposition(Prep,Kas) },
  [Prep],
  np([_Pers2,_Gen2,_Num2],[Kas]),
  ( [] ; advp([Art2],[ ]), { Art2 \= Art1 } ).
```

Hier fehlen die (optionalen) Adverbiale und die entsprechenden Regeln für die Verberst- und die Verbletzt-Stellung. Ergänzen Sie die fehlenden Regeln und fügen Sie (analog zu den Sätzen aus transitiven Verben) optionale `advp`-Konstituenten ein. Ergänzen Sie passende Beispiele in `beispielsaetze-mit-adverbien.txt` und erzeugen Sie die Syntaxbäume dazu.

- (b) Ebenso fehlen optionale Adverbiale in den Satzregeln für Relativsätze. Fügen Sie hier ebenfalls optionale Adverbiale ein (siehe TODO) und testen Sie Ihre Regeln mit passenden neuen Beispielen und `?- parse('beispielsaetze-mit-adverbialen.txt')`.

- (c) Auf den Folien zur Koordination von Ausdrücken wurden ein paar Regeln angegeben, bei denen durch Einschränkung von Merkmalen vermieden werden soll, daß der DCG-Parser wegen einer Linksrekursion nicht terminiert. Die Regeln finden Sie in `koordination.pl` auf der Kursseite, wobei die kritischen (linksrekursiven) Regeln auskommentiert sind.

Schreiben Sie eine Datei von Beispielen und eine Ladedatei

```
beispielkoordinationen.txt
beispielkoordinationen.pl
```

die u.a. `koordination.pl` lädt, und testen Sie, ob ihre Beispiele analysiert werden.

Dazu müssen Sie eine der kritischen Regeln aktivieren (die Auskommentierung aufheben). Falls die Analyse nicht terminiert, müssen Sie die kritische Regel weiter einschränken. (Versuchen Sie das mit einer einzelnen Regel, sodaß immer klar ist, an welcher Regel das Nichtterminieren liegen muß.)

Lösung von Aufgabe 9.1

- (a) Eine Regel für Aussagesätze mit einem Verb mit Präpositionalobjekt und Adverbialen ist:

```
s([def],[Temp,Mod,vz]) --> % def = Aussagesatz
  advp([Art1],[ ]), % Advp im Vorfeld
  { Art1 \= neg },
  v([[nom,Prep,Kas]], [Pers,Num,Mod,Temp]), % Verb mit Objekt-PP
  { praeposition(Prep,Kas) },
  np([Pers,_Gen,Num],[nom]),
  ([ ] ; advp([Art2],[ ]), { Art2 \= Art1 } ), % Stellung!
  [Prep],
  np([_Pers2,_Gen2,_Num2],[Kas]).
```

Mit ihr kann man absichtlich nur die ersten beiden der folgenden Beispiele erkennen:

deshalb basiert die Finanzverfassung auf den gesetzlichen Vorgaben.
deshalb basiert die Finanzverfassung nicht auf den gesetzlichen Vorgaben.
deshalb basiert die Finanzverfassung auf den gesetzlichen Vorgaben nicht.
nicht basiert die Finanzverfassung deshalb auf den gesetzlichen Vorgaben.

Für untergeordnete Sätze braucht man die Verbletzstellung; in diesem Fall stehen die Adverbien m.M. vor dem finiten Verb:

```
s([def],[Temp,Mod,v1]) --> % def = Aussagesatz
  np([Pers,_Gen,Num],[nom]),
  ([ ] ; advp([Art1],[ ])),
  ([ ] ; advp([Art2],[ ]), { Art2 \= Art1 } ),
  [Prep], { praeposition(Prep,Kas) },
  np([_Pers2,_Gen2,_Num2],[Kas]),
  v([[nom,Prep,Kas]], [Pers,Num,Mod,Temp]). % Verb mit Objekt-PP
```

Mit dieser Regel könnte man die für Adverbialsätze nötige Stellung erkennen:

(falls) die Finanzverfassung deshalb auf den gesetzlichen Vorgaben basiert.
 (falls) die Finanzverfassung deshalb nicht auf den gesetzlichen Vorgaben basiert.

- (b) Relativsätze mit Adverbialen. Wir relativieren nur mit Personalpronomen, also nur das Subjekt, wie in *(die Finanzverfassung,) die deshalb auf den Vorgaben basiert*, und von den Objekten nur die Nominalobjekte, wie *(das Buch,) das die Kinder lesen*, aber nicht die Präpositionalobjekte, wie *(die Vorgaben,) auf denen die Finanzverfassung basiert*.

Entsprechende Relativsatzregeln für Relativsätze mit intransitivem, transitivem und 2-stelligem Verb mit Präpositionalobjekt sind dann:

% Relativsätze mit Adverbialen:

```
s([rel(Gen,Num)], [Temp,Mod,v1]) --> % Relativsatz mit RelPron.nom
  pron([rel], [Pers,Gen,Num,nom]),
  { Pers = 3 },
  ( [] ; advp([Art1],[]) ),
  ( [] ; advp([Art2],[]), { Art2 \= Art1 } ),
  v([[nom]], [Pers,Num,Mod,Temp]).

s([rel(Gen,Num)], [Temp,Mod,v1]) --> % Relativsatz mit RelPron.nom
  pron([rel], [Pers,Gen,Num,nom]),
  { Pers = 3 },
  np([_Pers2,_Gen2,_Num2], [Kas]),
  ( [] ; advp([Art1],[]) ),
  ( [] ; advp([Art2],[]), { Art2 \= Art1 } ),
  v([[nom,Kas]], [Pers,Num,Mod,Temp]).

s([rel(Gen2,Num2)], [Temp,Mod,v1]) --> % Relativsatz mit RelPron /= nom
  pron([rel], [3,Gen2,Num2,Kas]),
  np([Pers,_Gen,Num], [nom]),
  ( [] ; advp([Art1],[]) ),
  ( [] ; advp([Art2],[]), { Art2 \= Art1 } ),
  v([[nom,Kas]], [Pers,Num,Mod,Temp]).

s([rel(Gen,Num)], [Temp,Mod,v1]) --> % mit RelPron.nom + Präp.objekt
  pron([rel], [Pers,Gen,Num,nom]),
  ( [] ; advp([Art1],[]) ),
  ( [] ; advp([Art2],[]), { Art2 \= Art1 } ),
  [Prep], { praeposition(Prep,Kas) },
  np([_Pers2,_Gen2,_Num2], [Kas]),
  v([[nom,Prep,Kas]], [Pers,Num,Mod,Temp]). % Verb mit Objekt-PP
```

In diese Regeln ist keine Einschränkung an die Stellung des Negationsadverbs eingebaut. Da das im Komplementrahmen des Verbs vorkommende **Kas** stets im Formmerkmal einer Nominalphrase auftritt, braucht man keine Nebenbedingung { **Kas** = **nom** ; ... ; **Kas** = **akk** }. Analog für **Temp** usw.

Mit diesen Regeln kann man z.B. analysieren:

der ihm deshalb einen Bären aufbindet
dem er deshalb einen Bären aufbindet
den er ihm deshalb aufbindet die deshalb auf den gesetzlichen Vorgaben basiert
die deshalb nicht auf den gesetzlichen Vorgaben basiert

Aber (absichtlich) nicht erkannt werden gewisse andere Stellungen der Adverbiale wie

die auf den gesetzlichen Vorgaben deshalb basiert

- (c) Koordinationsregeln: Auf den Vorlesungsfolien waren Koordinationsregeln für Nominalphrasen, Sätze und Adjektivphrasen angegeben. Eine Datei `beispielkoordinationen.txt` von Testbeispielen sollte also typische Koordinationen solcher Ausdrücke (mit dem Vokabular der Beispielllexika) enthalten; da Personalpronomen und gewisse durch Adjektive oder Relativsätze modifizierte Nomen als Nominalphrasen gelten, sollte man auch deren Koordination testen, etwa:

`% Nomenkoordinationen`

`Frauen und Männer.`

`Frau und Kind.`

`Frau oder Kind.`

`% Koordination adj.erweiterter Nomen`

`großen Männern und kleinen Kindern.`

`großen Männern und Kindern.`

`Männern und großen Kindern.`

`Frauen, Männern und kleinen Kindern.`

`% Nominalphrasenkoordination`

`er und sie. % Num = pl`

`er und sie las ihn. % darf nicht erkannt werden`

`er oder sie. % Num = sg`

`er oder sie lasen ihn. % darf nicht erkannt werden`

`ihn oder die Kinder.`

`mich oder ihn.`

`ich und wir.`

`den Mann und die Frau.`

`viele Kinder und Fragen.`

`die Frau, das Kind und die Männer.`

`die Frau, das Kind oder die Männer.`

`die Frau, das Kind oder der Mann.`

`% Satzkoordination`

`er las ihn und sie las ihn.`

`sie arbeitete und das Kind las ein Buch.`

arbeitete sie und las das Kind ein Buch.
ob das Kind ein Buch las und sie arbeitete.

er und sie lasen ihn. % koordiniertes Subjekt
er oder sie las ihn.

% Relativsatzkoordinationen

der das Buch las.
den er las und der arbeitet.
der es las, und der auf der Vorgabe basierte.
dem Buch, das er liest und das auf der Vorgabe basierte.
der Autor, der arbeitete, den er liest, und der ihn las.
der Mensch, der ihn liest.

% NP- oder RelS-Koordinationen ?

der Autor und der Mann, der ihn las.
der Autor und der Mann, die das Buch lasen.
der Frau, die das Buch las.
der Frauen, die das Buch lasen.

Bei den Koordinationsregeln der Vorlesungsfolien war das Problem der Linksrekursion dadurch behoben, daß die Merkmalwerte des linken Koordinationsglieds von denen des koordinierten Ausdrucks verschieden waren, hier die *und*-Regel:

% Koordination von Nominalphrasen mit Verhinderung der Linksrekursion

```
np([Per,-,pl],[Kas]) -->
  { member(Gen1,[mask,fem,neut]) % linke NP hat echtes Genus,
    ; Gen1=(-), Per1=1, Per=(-) }, % oder: falls linke NP 1.Person,
  np([Per1,Gen1,_],[Kas]), % soll die gesamte NP Person - haben
  [und], ([auch];[]),
  np([Per2,_,_],[Kas]),
  { agreeOrNone(Per1,Per2,Per) }.
```

Diese Regel erlaubt es z.B. nicht, Reihungen wie *Frauen, Männern und kleinen Kindern* zu analysieren, und sie erkennt (*viele Kinder*) und *Fragen* nur in der angegebenen Klammerung. Unschön ist natürlich, daß die Bedingungen an die Merkmale zwar Linksrekursion ausschließen, aber etwas willkürlich: z.B. wird, weil im Plural der Genuswert - vergeben wird, keine Koordination wie *euch und sie* (mit linker Konstituente in 2.Person) erkannt.

Um solche willkürlichen Einschränkungen zu umgehen, wäre es sinnvoll, zwischen koordinierten und einfachen Nominalphrasen durch ein Artmerkmal zu unterscheiden⁵. Dann kann die NP-Koordinationsregel so formuliert werden:

```
np([Per,-,Num,co],[Kas]) --> % co = koordiniert
  np([Per1,_ ,Num1],[Kas]),
```

⁵Hier wird $NP_{Pers,Gen,Num}^{Kas}$ für einfache, $NP_{Pers,Gen,Num,co}^{Kas}$ für koordinierte verwendet

```

( [Conj],
  np([Per2,_,Num2],[Kas])
; % [,'], % beim Einlesen gelöscht
  np([Per2,_,Num2,co],[Kas])
),
{ agreeOrNone(Per1,Per2,Per) },
{ Conj = und, Num = pl
; Conj = oder, agreeOrNone(Num1,Num2,Num)
}.

```

Kritisch ist die Regel zur Satzkoordination:

```

% Koordination von Sätzen mit Verhinderung der Linksrekursion
s([Def],[-,Mod,Vst]) --> % Temp != -
  { Temp=praes ; Temp=praet }, % verhindert Linksrekursion
  s([Def1],[Temp,Mod,Vst]),
  ( [und] ; [oder] ),
  s([Def2],[Temp,Mod,Vst]),
  { agree(Def1,Def2,Def) }. % Oder?

```

Sie sollte zwar keine Linksrekursion erzeugen, aber das Parsen wird mit dieser Regel sehr langsam: ein offensichtlicher Nachteil ist, daß durch die Nebenbedingung zu *Temp* die linke Konstituente stets zweimal analysiert werden muß, obwohl höchstens einer der Tempuswerte korrekt sein kann.⁶ (Es scheint auch, daß nach jedem Nomen viele Versuche, einen Relativsatz zu finden, durchlaufen werden.)

Besser ist es wieder, zwischen einfachen und koordinierten Sätzen so zu unterscheiden, daß eine solche Alternative zu *Temp* unnötig wird, etwa mit $S_{Def}^{Temp,Mod,Vst}$ für einfache und $S_{Def}^{Vst,Mod}$ für koordinierte Sätze:

```

s([Def],[Vst,Mod]) -->
  s([Def],[Temp1,Mod,Vst]),
  ( ( [und] ; [oder] ),
    s([Def],[Temp2,Mod,Vst]) .
  ; % [,'] beim Einlesen gelöscht
    s([Def],[Vst,Mod])
  ).
s([Def],[Vst,Mod]) -->
  s([Def],[Temp1,Mod,Vst]) .

```

Man muß noch dafür sorgen, daß mit

```

s([rel(Gen,Num)],[Mod]) --> s([rel(Gen,Num)],[Vst,Mod]) .

```

die Relativsätze in der N' -Regel an Nomina angeschlossen werden können

Bei den Adjektivphrasen wäre am klarsten, einfache und koordinierte mit einem Artmerkmal zu unterscheiden und die Koordination durch

⁶Die entsprechende Nebenbedingung der Tempusalternative in Relativsätzen habe ich weglassen, damit am fehlenden Relativpronomen schneller erkannt wird, daß kein Relativsatz folgt.

```

% Koordination von einfachen Adjektivphrasen
ap([co],[pos,Dekl,Gen,Num,Kas]) -->
  ap([si],[pos,Dekl,Gen,Num,Kas]),
  ( ([und] ; [oder] ; []), % [] statt [',']
    ap([si],[pos,Dekl,Gen,Num,Kas])
  ; % [','] beim Einlesen gelöscht
    ap([co],[pos,Dekl,Gen,Num,Kas])
  ).

```

zu regeln. Dann muß man aber die bisherigen Regeln für `ap([], [Komp,Dekl,Gen,Num,Kas])` zu Regeln für `ap([si], [Komp,Dekl,Gen,Num,Kas])` ändern (oder eine Regel `ap([si], Form)` --> `ap([], Form)` ergänzen) und in der Modifikation von Nomen durch Adjektivphrasen die Kategorie `ap([Art], [Komp,Dekl,Gen,Num,Kas])` verwenden.