

Einführung in die Grammatikentwicklung: Übung 2 (27. Mai 2020)

Ziele:

1. Etwas über getypte Merkmalstrukturen und Unifikation lernen.
2. Die Grammatik um Beschränkungen (*constraints*) erweitern, die Kongruenz sicherstellen.
3. Einführung von Kasuszuweisung.

Übungen:

1. TRALE starten und Grammatik 2 laden: Klicken Sie auf das grüne T, unter dem Task 2 steht.
2. Testen Sie das System, indem Sie den Satz *Erinnert der Hund die Katze an den Hund?* analysieren.
 - (a) Geben Sie im Emacs-Fenster nach dem Prompt (`>>>`) den Satz *Erinnert der Hund die Katze an den Hund?* ein.
 - (b) Drücken Sie Enter.
3. Machen Sie sich mit der Datenstruktur, mit der wir arbeiten, vertraut.
 - (a) In der Datei `signature` befindet sich die Definition der Datenstruktur, mit der wir arbeiten. Zur Zeit sieht die Datei wie folgt aus:

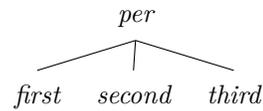
```
type_hierarchy
bot
  feat_struct
    syn_struct phon:list head:pos subcat:list
    phrase dtrs:list
    word
  pos
    noun
    verb
    det
    prep
  list
    e_list
    ne_list hd:bot tl:list
.
```

bot ist der oberste Typ, die Untertypen sind jeweils zwei Zeichen weit eingerückt. Wichtig ist der Punkt am Ende. Er darf nicht entfernt werden. Die direkten Untertypen von *bot* sind also *feat_struct* und *list*. *feat_struct* hat selbst Untertypen und zwar: *syn_struct* und *pos*. Die Typhierarchie kann man sich grafisch anzeigen lassen. Rufen Sie dazu im Trale-Menü Trale|Draw|Hierarchy auf. Unten im Emacs-Buffer erscheint dann eine Eingabeaufforderung. Hier können Sie einen Typ eingeben oder einfach Enter drücken. Wenn Sie Enter drücken, wird der Default-Wert genommen (der steht in eckigen Klammern). Für die Typhierarchieausgabe ist der Defaultwert „bot“. *bot* ist der allgemeinste Typ, weshalb dann die gesamte Hierarchie ausgegeben wird. Die Tastatur-Abkürzung für die Ausgabe der Typhierarchie ist Ctrl-c Ctrl-h.

Merkmalstrukturen vom Typ *syn_struct* besitzen die Merkmale PHON, HEAD und SUBCAT. Welche Merkmale zu einem Typ gehören, wird in der Signatur-Datei nach diesem Typ angegeben. Des Weiteren wird angegeben, von welchem Typ die Werte der Merkmale sein müssen. So ist z. B. der Wert des PHON-Merkmals vom Typ *list*. Die gesamte Signatur, d. h. Information darüber, welche Typen es gibt und welche Merkmale sie einführen, kann man sich mit (Trale|Draw|Signature) bzw. Ctrl-c Ctrl-o ausgeben lassen:

4. Fügen Sie der Grammatik Beschränkungen für Subjekt-Verb-Kongruenz hinzu, so dass zwar *Bellt der Hund?* nicht aber *Bellen der Hund?* analysiert wird.

- (a) Fügen Sie in der Datei `signature` die Merkmale `per` und `num` zum Typ `pos` hinzu. Die Werte werden jeweils als vom Typ `per` bzw. `num` definiert. Zu einem Typ gehörige Merkmale müssen hinter diesem in derselben Zeile mit dem allgemeinsten zulässigen Wert angegeben werden (`typ merkm1:wert1 merkm2:wert2`).
- (b) Fügen Sie die beiden folgenden Typhierarchien ein:



Dabei sind `num` und `per` als direkte Untertypen von `bot` in `signature` aufzunehmen, d. h. sie müssen zwei Leerzeichen eingerückt stehen.

- (c) Fügen Sie in der Datei `lexicon.pl` in allen Lexikoneinträgen für Verben Beschränkungen für den PER- und NUM-Wert des Subjekts in der SUBCAT-Liste des Verbs ein. Dabei kann eine Merkmalstruktur vom Typ `noun`, die die Merkmale NUM und PER hat wie folgt beschrieben werden:

```
(noun,
  per:third,
  num:sg)
```

Dinge, die zur selben Einbettungstiefe gehören, werden geklammert. Der Typ einer Struktur und Merkmal-Wert-Paare werden durch Komma getrennt.

- (d) Fügen Sie in der Datei `lexicon.pl` in allen Lexikoneinträgen für Nomina die PER- und NUM-Werte ein.
- (e) Speichern Sie die Änderungen ab, laden Sie die Grammatik neu, lassen Sie einen Test-Parse der Beispiele 12–19 aus der Datei `test_items.pl` laufen, indem Sie Ctrl-c Ctrl-t eingeben und dann 12 Enter und 19 Enter. Machen Sie die erforderlichen Korrekturen. (**alle Hund* wird in Aufgabe 6 zur Determinator-Nomen-Kongruenz geklärt.)
5. Fügen Sie Lexikoneinträge für *gab* ein.

- (a) Fügen Sie Lexikoneinträge für *gab* in die Datei `lexicon.pl` ein, so dass *Gab ich dem Hund die Katze?* und *Gab der Hund der Katze das Schaf?* analysierbar sind.

6. Erweitern Sie die Analyse um Determinator-Nomen-Kongruenz (vorerst nur Person und Numerus), so dass *Bellen alle Hunde?* nicht aber *Bellt alle Hund?* zugelassen wird.

- (a) Fügen Sie in der Datei `lexicon.pl` allen Lexikoneinträge für Nomina Beschränkungen für den NUM-Wert des Determinators in der SUBCAT-Liste des Nomens ein.
- (b) Fügen Sie in der Datei `lexicon.pl` die NUM-Werte für die Lexikoneinträge der Determinatoren ein.
- (c) Lassen Sie einen Batch-Parse mit den entsprechenden Beispielen laufen und machen Sie die erforderlichen Korrekturen.
- (d) Fügen Sie weitere Beispiele für Kongruenzverletzungen zwischen Determinatoren, Nomina und Verben in die Datei `'test_items.pl'` ein und führen Sie einen erneuten Testlauf durch.

7. Stellen Sie Kasuskongruenz innerhalb der Nominalphrase sicher:

- (a) Fügen Sie in `'signature'` genauso wie Sie das Person- und Numerus-Merkmal eingefügt haben auch ein Kasus-Merkmal ein (`CASE`).
- (b) Definieren Sie in `'signature'` den Typ `case` mit den Untertypen `nom`, `gen`, `dat` und `acc` für Nominativ, Genitiv, Dativ und Akkusativ.
- (c) Erzeugen Sie neue Lexikoneinträge für Determinatoren wie *der*, so dass zwischen Nominativ Singular und Genitiv bzw. Dativ Plural unterschieden wird.
- (d) Führen Sie einen Testlauf mit den Beispielen 20–44 durch (`testt([20,44]).`) und verbessern Sie die Grammatik entsprechend. Manche Testbeispiele haben zu viele Lösungen, was daran liegt, dass wir bisher kein Genusmerkmal haben. Ignorieren Sie das vorerst.

8. Stellen Sie die Kasuszuweisung durch Verben sicher:

- (a) Parsen Sie die Sätze: *Gibt der Hund dem Schaf die Katze?* und *Gibt der Hund die Katze dem Schaf?*.
- (b) Spezifizieren Sie die Kasus-Werte aller Argumente von Verben in der Datei 'lexicon.pl'.
- (c) Überlegen Sie warum einerseits der Satz *Gibt der Hund dem Schaf die Katze?* von der Grammatik verarbeitet werden kann und andererseits der Satz *Gibt der Hund die Katze dem Schaf?* nicht analysiert wird.
- (d) Passen Sie die Grammatik so an, dass wieder beide Sätze analysiert werden können.