

Einführung in die Grammatikentwicklung: Übung 1 (27. Mai 2020)

Ziele:

1. Vertraut werden mit dem TRALE-Grammatikentwicklungssystem.
2. Lernen, eine Grammatik durch das Hinzufügen von Lexikoneinträgen zu erweitern.
3. Erweitern der Grammatik, so dass sie mit ditransitiven Verben umgehen kann.
4. Erweitern der Grammatik, so dass sie mit Präpositionalphrasenargumenten umgehen kann.

Materialien:

1. Kurs-Seite: <https://hpsg.hu-berlin.de/~stefan/Lehre/GE-Trale/>
2. Virtuelle Maschine mit Grammix: <https://hpsg.hu-berlin.de/Software/Grammix/>
3. Trale-Standaloneversion für Linux: <https://hpsg.hu-berlin.de/Software/Trale/>

Übungen:

1. TRALE starten und Grammatik 1 laden:
2. Nach der Installation von VirtualBox und der virtuellen Maschine für Grammix (ein kleiner Linux-Computer in Ihrem Computer), starten Sie die virtuelle Maschine und klicken auf das grüne T, unter dem *Task 1* steht. Es sollten zwei Fenster aufgehen: ein Gralej-Fenster und ein Emacs-Fenster. Im Emacs-Fenster wird die Grammatik 1 geladen und der Eingabe-Prompt >>> erscheint. Sie sind jetzt im Analyse-Modus, in dem man Sätze eingeben kann. Im Analysemodus versucht das Prolog-Programm, das unsere Grammatik verarbeitet, die Eingaben als Wörter zu interpretieren und diese dann zu Wortgruppen zusammenzubauen.
3. Testen Sie das System, in dem Sie den Satz *Jagt der Hund die Katze?* analysieren:
 - (a) Geben Sie nach dem Prompt (>>>) den Satz *Jagt der Hund die Katze?* ein.
 - (b) Drücken Sie Enter.

Das System wird den Satz verarbeiten und ein kleines Fenster mit einem Parsebaum öffnen. In diesem Baum kann man die einzelnen Knoten anklicken und die entsprechenden Merkmalstrukturen anzeigen lassen. Bei erneutem Klicken verschwinden die Strukturen wieder.

Außerdem öffnet sich das chart display. Das chart display hilft einem zu sehen, welche Konstituenten das Grammtikverarbeitungssystem zu größeren Einheiten zusammenbaut.

4. Probieren Sie den einfachen *batch*-Parse-Mechanismus:
 - (a) Wählen Sie oben im Emacs-Fenster das TRALE-Menü aus und dort den Unterpunkt `Test items`.¹
 - (b) Unten im Emacs-Fenster steht jetzt „Enter first item to test [test all]:“. In solchen Eingabeaufforderungen steht in eckigen Klammern immer der Default-Wert, also das, was passiert, wenn man einfach Enter drückt (die Eingabetaste). Drücken Sie jetzt einfach Enter.
 - (c) Das System verarbeitet alle Testsätze, die in der Datei `test_items.pl` stehen. Die Ausgaben erscheinen im `*prolog*`-Emacs-Puffer. Hinter den meisten Sätzen steht ein grünes OK. Nur hinter einem Satz steht „1 solutions (!!! 0 expected !!!)“. Wir werden in späteren Übungen sehen, wie wir verhindern können, dass für diesen Satz Analysen gefunden werden.
5. Fügen Sie einen Lexikoneintrag für ein weiteres Nomen, das ein Tier bezeichnet, hinzu. Wählen Sie ein Nomen mit dem Genus Neutrum.

¹Hinter allen Emacs-Menüpunkten stehen immer Sequenzen, mit denen man den Menüpunkt schneller aufrufen kann. Ctrl-c Ctrl-t (oder Strg-c Strg-t) ruft ebenfalls das Test-Kommando auf. In den Emacs-Menüs stehen entsprechende Abkürzungen: c-c c-t.

- (a) Öffnen Sie die Datei 'lexicon.pl' im Emacs-Fenster zum Editieren (siehe Emacs-Handout).
 - (b) Kopieren Sie in dem lexicon.pl-Puffer, den Sie erhalten, drei Zeilen, die den Lexikoneintrag für *Hund* definieren, und modifizieren Sie die Kopie so, dass die Zeichenfolge vor dem ---> für ein anderes Tier passend ist.
 - (c) Speichern Sie die abgeänderte Version der Datei.
6. Laden Sie die Grammatik neu und testen Sie das Ergebnis Ihrer Änderung:
- (a) Geben Sie Ctrl-c Ctrl-a ein. Wenn Sie nicht im *prolog*-Puffer sind, sollte Emacs dorthin umschalten und die Grammatik laden.
 - (b) Parsen Sie den Satz *Jagt die Katze das <Ihr-Tier>*.
 - (c) Nehmen Sie diesen Satz in die Datei test_items.pl auf.
Der Eintrag muss die folgende Form haben:

```
t(8,"Jagt die Katze das <Ihr-Tier>?",@root,1,transitive).
```
- Dabei ist '8' eine fortlaufende Nummer. An zweiter Stelle steht der jeweilige Testsatz, danach kommt ein Symbol, das vollständige Äußerungen beschreibt (später mehr), danach die Anzahl der Lesarten des Satzes (in diesem Fall '1') und zum Schluß eine Einordnung des Satzes in eine Phänomenklasse.
- (d) Laden Sie die Grammatik neu und führen Sie erneut den Batch-Parse aus.
7. Schauen Sie sich die Grammatik an, um eine intuitive Vorstellung davon zu bekommen, wie sie funktioniert (formale Details werden später diskutiert werden). Schauen Sie sich die folgenden Wortgruppen an und versuchen Sie herauszufinden, warum Sie analysiert werden können oder warum das nicht möglich ist:
- bellt die Katze
 - jagt die Katze
 - bellt Katze
 - bellen die Katze
 - bellen
- Man beachte, dass die Parse-Tabelle (*chart*) auch dann zur Verfügung steht, wenn der Satz nicht erfolgreich verarbeitet wurde. In der Chart kann man auf Knoten klicken, um sich Merkmalstrukturen anzusehen. Man beachte, dass die Grammatik einige Sätze unkorrekterweise parst (Übergenerierung) und manche grammatischen Sätze nicht analysieren kann (Untergenerierung).
8. Die Regel, die für Sätze mit ditransitiven Verben gebraucht wird, ist bereits Bestandteil der Grammatik, aber es gibt keine Lexikoneinträge, die sie benutzen. Fügen Sie einen Lexikoneintrag für *gibt* mit Selektion eines Subjekts, direkten Objekts und indirekten Objekts hinzu (d. h. den Eintrag, den man braucht um *Gibt der Hund der Katze das <Ihr-Tier>* zu parsen).
- (a) Kopieren Sie den Eintrag für *jagt* in lexicon.pl.
 - (b) Ersetzen Sie den Wert vor dem Pfeil durch *gibt*.
 - (c) Fügen Sie ein zusätzliches Element in die SUBCAT-Liste ein. Dieses Element ist einfach eine Kopie eines der bereits vorhandenen Elemente. Man beachte: Listen werden durch '[' und ']' begrenzt. Die Elemente in Listen werden durch Kommata getrennt.
 - (d) Testen Sie, indem Sie *Gibt der Hund der Katze das <Ihr-Tier>* parsen. Testen Sie auch auf Übergenerierung, indem Sie testen, ob *Gibt der Hund der Katze* geparst werden kann.
 - (e) Fügen Sie der Datei 'test_items.pl' entsprechende Sätze hinzu.
9. Schreiben Sie einen neuen Typ und zwei neue Lexikoneinträge, um *Erinnert der Hund die Katze an das <Ihr-Tier>* verarbeiten zu können.

- (a) Fügen Sie in Analogie zu **noun** den Typ **prep** als neuen Untertyp des Typs **pos** (part of speech) in der Datei 'signature' ein.
- (b) Schreiben Sie einen Lexikoneintrag für die Präposition *an*. Dieser sollte dem Eintrag für *belien* ähneln, d. h. *an* verlangt eine Nominalphrase als abhängiges Element in der SUBCAT-Liste. Der HEAD-Wert von *an* muss **prep** sein.
- (c) Schreiben Sie einen Eintrag für *erinnert*. Sie können den Eintrag für *geben* als Vorlage verwenden. Das Wort vor dem Pfeil muss wieder angepaßt werden. Das dritte Element in der Subcat-Liste muss ebenfalls geändert werden, damit *erinnern* eine PP verlangt (Zur Zeit stellen wir nicht sicher, dass es sich auch wirklich um eine *an*-PP handeln muss).
- (d) Fügen Sie sowohl grammatische als auch ungrammatische Testwortgruppen in 'test_items.pl' ein, die eine Überprüfung der Korrektheit der Grammatikerweiterung erlauben.
- (e) Lassen Sie erneut einen Batch-Parse mit 'test_items.pl' laufen und sehen Sie sich die Ergebnisse an.
- (f) Feiern Sie!

Anhang (für Menschen, die TRALE unter ihrem eigenen Linux nutzen)

1. Öffnen Sie ein xterm-Fenster (auf Computer-Icon für Konsole links unten klicken)
2. Setzen Sie sich ins Verzeichnis GE (Grammatikentwicklung), indem Sie


```
cd GE
```

 eingeben. (`cd` steht für „change directory“ (Wechsle das Verzeichnis)).
3. Setzen Sie sich ins Verzeichnis **Grammatik1**, indem Sie in einem xterm-Fenster


```
cd Grammatik1
```

 eingeben.
4. Starten Sie TRALE, indem Sie im xterm-Fenster `trale` eingeben. Dieser Befehl öffnet ein Emacs-Fenster und lädt die in `theory.pl` spezifizierte Grammatik. Nachrichten in bezug auf den Ladevorgang werden im Emacs-Fenster erscheinen.
5. Testen Sie das System, in dem Sie den Satz *Jagt der Hund die Katze?* analysieren:
 - (a) Geben Sie im Emacs


```
go.
```

 ein. Sie sind jetzt im Analyse-Modus, in dem man Sätze eingeben kann. Im Analysemodus versucht das Prolog-Programm, das unsere Grammatik verarbeitet, die Eingaben als Wörter zu interpretieren und diese dann zu Wortgruppen zusammenzubauen.