

Übung 2

— Parsing —

Bitte reichen Sie die Aufgaben bis Montag, den 19.11. ein.

1 Erste Schritte

Um eine kleine LFG-Grammatik erstellen zu können öffnen Sie bitte als erstes X11 auf der unteren Leiste. Tippen Sie dort ein:

`emacs & (enter)` (öffnet den üblichen Editor)

Tastenkombination `ESC-x` (versetzt sie in die untere Befehlsleiste) `lfg-mode (enter)`

Gehen Sie nun unter File/Open File und tippen Sie `MTTestgrammatik.lfg` ein. Es sollte sich nun eine Datei öffnen, in der Sie Ihre eigene kleine Grammatik entwickeln können. Die oberen Angaben in der Datei sind für das System und brauchen uns hier nicht weiter zu interessieren.

2 Syntaktische Regeln

Unter dem Abschnitt `TOY ENGLISH RULES` lassen sich nun syntaktische Regeln bestimmen. Hierfür muss ein bestimmtes Schemata benutzt werden:

```
Kategorie --> Kategorie1: Schemata1;  
                Kategorie2: Schemata2.
```

Also beispielsweise

```
S          --> NP: (^ SUBJ) = ! ;  
           VP: ^=! .
```

NP und VP sind wiederum Knoten, die weiter definiert werden müssen:

```
VP         --> V .  
NP         --> (D) .
```

(Klammern symbolisieren Optionalität; ein Semikolon wird nach ihnen nicht benötigt)

Tippen Sie diesen oberen Abschnitt nun unter Ihren RULES ein. Beachten Sie die Satzzeichen: vor jeder neuen Kategorie muss ein Semikolon stehen. Ist die Regel zu Ende, dann wird dies durch einen Punkt symbolisiert. Die LFG verknüpft in ihrer Grammatik die Konstituentenabfolge (C(onstituent)-Structure) mit den grammatischen Beziehungen wie Subjekt oder Objekt (F(unctional)-Structure); daher werden beide Aspekte in die Regeln integriert.

Wenn Sie ihre Grammatik nun laden wollen so klicken Sie im Editor oben auf 'LFG' und dann auf 'start a new XLE process'. Hinter dem Prozentzeichen tippen Sie nun 'create-parser MTTestgrammatik.lfg' um ihre kleine Grammatik hochzuladen. Wenn Sie etwas falsch gemacht haben, dann teilt das System Ihnen das nun mit; ansonsten bekommen Sie eine positive Mitteilung. Um einen Satz zu parsen tippen Sie: parse "sentence". Leider erscheint noch keine Baumstruktur, da das Lexikon leer ist: sie müssen Ihrer Grammatik erst ein paar Wörter beibringen.

3 Das Lexikon

Dazu wechseln wir zurück zu unserer Grammatik (unter 'Buffers' im Editor können Sie hin-und herwechseln).

Wörter gehören in den Abschnitt TOY ENGLISH LEXICON 1.0) und haben das folgende Schema:

Wort	Kategorie	Schemata
girls	N *	([^] PRED)='GIRLS'.
laugh	V *	([^] PRED)='LAUGH< ([^] SUBJ) >'.
the	D *	([^] DEF)=+.

Tippen Sie beides in ihr LEXICON, speichern Sie es und wechseln Sie zurück zum XLE-Buffer. Dort müssen Sie unter XLE das System restarten und Ihre Grammatik noch einmal laden um garantiert Ihre neueste Version zu bekommen. Nun versuchen Sie: parse "the girls laugh" — wenn Sie eine schöne C- und F-Struktur bekommen, dann haben Sie alles richtig gemacht; wenn nicht müssen sie nach dem Fehler suchen.

4 Weitere Aufgaben

Bis hierher haben wir nur mit dem intransitiven Verb 'laugh' gearbeitet.

Verändern Sie nun Ihre Regeln so dass diese auch ein Objekt beschreiben können. Verwenden Sie den Satz 'the girls kiss the boys'.

Hilfestellung:

1. Überlegen Sie zuerst wo sich das Objekt im Englischen Satz befindet. Welche Regel unter RULES müssen Sie verändern?
2. Welche Kategorie brauchen Sie für das Objekt?
3. Wie muss der Eintrag gestaltet sein um intransitive Verben weiterhin zu ermöglichen?
4. Vergessen Sie nicht die neuen Lexikoneinträge. Wie muss der Lexikoneintrag für 'kiss' aussehen damit er Transitivität erlaubt?

5 Zusatzaufgaben für Extracredit

- a. Außer transitiven und intransitiven Verben gibt es noch ditransitive Verben, die zwei Objekte einfordern. Das indirekte Objekt wird in der XLE-Grammatik als OBJ2 bezeichnet. Versuchen Sie, den Satz "The boy gives the dog the bone" korrekt zu parsen.
- b. Versuchen sie eine Präpositionalphrase in Ihre Grammatik einzubauen, so dass man Sätze wie "The boys laugh in the garden" ebenfalls parsen kann. Wenn man mehrere Präpositionalphrasen verwenden möchte (z.B. "the boys laugh in the garden in the city"), mit welchem Operator lässt sich das ganz einfach ausdrücken?