# Parsing, WS 2006/07
# Übungsaufgaben

H.Leiß

23. Oktober 2006

**Aufg. 1** Es sei ein Vollformenlexikon als eine Liste von Prolog-Fakten der Form

⟨*Eintrag des Vollformenlexikons*⟩≡
```
word(Vollform,Kategorie).
```

gegeben, wobei `Vollform` ein Prolog-Atom und `Kategorie` ein Prolog-Term sei, beide ohne Variable. Zum Beispiel:

⟨*Lexikon*⟩≡
```
word(das,det).
word(das,rel).
word(läuft,v).
word('Programm',n).
```

(i) Schreibe ein Prolog-Programm

`eintraege(+Atomliste,+Nat)`,

das die Atome `A` der `Atomliste`, gepaart mit Nummern `N` beginnend bei der Zahl `Nat`, als Vorkommen von `A` an der Position `N` (als *Token*) versteht und aus dem Vollformenlexikon `word/2` alle Einträge zur Vollform `A` heraussucht, die gefundenen Kategorien `Cat` durch Fakten

⟨*Eintrag des Sublexikons*⟩≡
```
lexword(N,A,Cat)
```

in die Prolog-Datenbank schreibt und so ein Sublexikon `lexword/3` der Vollform-*Vorkommen* erzeugt.

Nach dem Aufruf Ihres Programms

```
?- eintraege([das,'Programm',das,läuft],0).
```

sollten Sie mit `listing(lexword)` folgende Fakten gezeigt bekommen:

⟨*Sublexikon*⟩≡
```
lexword(0,das,det).
lexword(0,das,rel).
lexword(1,'Programm',n).
lexword(2,das,det).
lexword(2,das,rel).
lexword(3,läuft,v).
```

(ii) Erweitern Sie Ihr Programm so, daß es für Wörter `W`, die mit einem Großbuchstaben anfangen und zu denen es keinen Eintrag im Lexikon `word/2` gibt, die Variante mit Kleinschreibung im Lexikon sucht und deren Kategorien in das Sublexikon als Kategorien von `W` aufnimmt.

Damit sollten Sie mit `?- eintraege(['Das','Programm',das,läuft],0).` und demselben Lexikon `word/2` folgende Einträge bekommen:

⟨*Sublexikon, b)*⟩≡
```
lexword(0,'Das',det).
lexword(0,'Das',rel).
lexword(1,'Programm',n).
lexword(2,das,det).
lexword(2,das,rel).
lexword(3,läuft,v).
```

Bem. Wenn `N` die Nummer eines Großbuchstabens ist, ist `N+32` die Nummer des entsprchenden Kleinbuchstabens.

(iii) Schreiben Sie ein Prädikat `tag_lexword`, das (a) am Satzanfang, d.h. bei Position 0, alle Lesarten als Relativpronomen `rel` entfernt, und unmittelbar vor einem Verbvorkommen die Lesarten als Artikel `det` entfernt.

Wenn Sie Ihr Prädikat bei obigem `lexword/3` aufrufen, sollte danach

⟨*Sublexikon, b), gefiltert*⟩≡
```
lexword(0,'Das',det).
lexword(1,'Programm',n).
lexword(2,das,det).
lexword(3,läuft,v).
```

übrig bleiben.

Gehen Sie davon aus, daß keine Abkürzungen oder Zahlen oder Sonderzeichen im Lexikon oder der Eingabe auftreten. Beachten Sie die skizzierten Programmteile auf den Folien.

Besprechung: Mo, 30.10. 9-10 Uhr. (Eigenen Teil beitragen!)