

Übungsblatt 4, Abgabe Montag 14. November 18 Uhr

Aufgabe 1 (nach Prof. Pinkal CL Kurs in Saarbrücken <http://www.coli.uni-saarland.de>)

Natürlich-sprachliche Systeme verwenden entweder Vollformenlexika oder Kombinationen aus einem Stammlexikon und einem Lemmatisierer. Die Einsparung an zu speichernden Wörtern ist je nach Sprache unterschiedlich. Für Englisch < Faktor 2 für Finnisch eher 50 (siehe Klabunde et al.)

Versuchen Sie die das Verhältnis Wortformen: Stämme für das Deutsche abzuschätzen. Wichtig ist, dass sie ihr Vorgehen motivieren können. Hinweis, wie produktiv sind die verschiedenen Wortklassen...wichtig ist nicht die Endzahl sondern dass das Schätzverfahren plausibel ist.

Aufgabe 2

Stellen Sie die vollständigen Paradigmen für mager und mein (= das deutsche Possessivpronomen) zusammen.

Aufgabe 3

Analysieren Sie folgende Wörter (in der Art von Beispiel 4.2 auf Seite 55 des Skripts von Jörg Schuster):

Mitgiftjäger, Weihnachtsbescherung, Liebesbrief, chemotechnisch
oblematische Fälle bitte kommentieren.

Aufgabe 4

„Lesen Sie den Artikel von Prof. Günthner. ‚Computational Linguistics: A personal View‘. Geben Sie an wie sich die maschinelle Sprachverarbeitung im Artikel darstellt und wie die einzelnen Gebiete miteinander verzahnt sind.

Haben Sie schon eine Idee was sich an diesem Programm inzwischen verändert hat.“

Falls noch jemand den Artikel braucht bitte bei mir im Büro abholen. Zimmer C112

Aufgabe 5

(a) Versuchen Sie das folgende Prologprogramm zum Laufen zu bringen.

```
%Startzustand
start(q0).
%Finalzustand
final(q2).

%Uebergaenge
%transition(Herkunftszustand, Symbol, Folgezustand)

transition(q0,a,q1).
transition(q1,b,q1).
transition(q1,c,q2).

%accept wählt den Startzustand und startet den Automaten
%%
%%consume_and_move arbeitet den Eingabestring rekursiv ab
%(rekursiv: ruft sich selbst wieder auf)
%bis entweder ein Finalzustand erreicht ist und alle Symbole des Eingabestrings konsumiert
%und (er ist leer) oder aber keine gültige transition mehr erreichbar ist; dann scheitert der
%Automat für den Eingabestring

accept(Symbols) :- start(StartState), consume_and_move(Symbols, StartState).

consume_and_move([],State) :- final(State).
consume_and_move([Symbol|RestSymbols], State):-
    transition(State, Symbol, NextState), consume_and_move(RestSymbols, NextState).

%%benutzen Sie entweder die grafische Oberfläche xpcce indem sie sie
%% mit emacs. Starten. Schreiben Sie ihr programm und rufen dann unter compile consult
selection auf oder direkt im Prolog Prompt

?- consult('programmname.pl').

%%Aufruf des Programms: Prolog Prompt
%%
%%
%% ?- accept([a,b,b,c]).      %%Eingabestring als Liste in Eckigen Klammern,
kommagetrennt

%%Ausgabe yes.

%% ?- accept([a,b,b,c,b]).    %%Eingabestring als Liste in Eckigen Klammern,
kommagetrennt

%%Ausgabe no.
```

(b) welche Sprache erkennt dieser Automat.

(c) was muss im Programm geändert werden, dass der Automat auch Strings mit beliebig vielen a's vor den b's akzeptiert, also z.B. das Wort aabbc