

Aufgabe 2

Entwicklung einer Computermorphologie

Programmieren Sie mit den SFST-Werkzeugen (<http://www.cis.lmu.de/~schmid/tools/SFST>) eine einfache Morphologie für eine morphologisch interessante Sprache (also nicht Englisch, Chinesisch).

Schritte

- Finden Sie heraus, welche Flexionsklassen in der Sprache existieren. Hier sind Wörterbücher und Webseiten wie Wiktionary hilfreich.
- Wählen Sie eine Teilmenge der Flexionsklassen aus.
- Implementieren Sie die Flexionsklassen und die erforderlichen morphophonologischen Regeln. (Mindestens 1 Regel sollte dabei sein.)

Empfohlene Schritte:

- Erstellen Sie eine Datei mit Stammlexikoneinträgen für jede Flexionsklasse. Sie können hier auch Stämme (glaub) auf Grundformen (glauben) abbilden durch einen Eintrag der Form: `glaube:<n:<>`
Beachten Sie dabei, dass in einer Lexikodatei nur die Symbole “:”, “<”, “>” und “\” als Operatoren interpretiert werden. Sie können also nicht schreiben `glaub{en}:{}`
- Erstellen Sie für jede Flexionsklasse einen Transducer, der Analysestrings auf Flexionsendungen abbildet, bspw. `{<pres><1><sg>}:{e} | {<pres><2><sg>}:{st}|...`
- Lesen Sie jedes Stammlexikon in einen Transducer ein und konkatenieren Sie diesen Transducer mit dem entsprechenden Flexionstransducer.
- Vereinigen Sie die Transducer für die verschiedenen Flexionsklassen mit Disjunktion.
- Schreiben Sie die notwendigen morphophonologischen Regeln, um die korrekten Wortformen zu erhalten (Bsp. happy+er → happier), und wenden Sie sie an.

Manchmal sind Triggersymbole nützlich. Beispielsweise bilden Nomen wie “Haus” oder “Buch” den Plural mit einem Umlaut. Hier kann man bspw. zu der Endung -er ein Triggersymbol hinzufügen: `{<p1>}:{er}` und dann später (nach der Konkatenation von Stämmen und Endungen) eine morphophonologische Regel anwenden, die den Stamm umlautet, wenn das Triggersymbol vorhanden ist und anschließend (oder gleichzeitig) das Triggersymbol löscht.

- Dokumentieren Sie, welche Phänomene Sie erfasst haben und an welcher Quelle (Buch, Webseite) Sie sich orientiert haben.

Schicken Sie Ihr SFST-Programm und Ihre Dokumentation an schmid@cis.lmu.de.