

4. Übungsblatt - Abgabe: 29.11.2011

Aufgabe 4.1

Auf der Webseite <http://community.languagetool.org/> können Sie eine Software für regelbasierte Grammatikprüfung testen.

Die Startseite präsentiert Ihnen drei zufällig ausgewählte Sätze, in denen die Regeln Fehler gefunden haben. Oft erkennt das System falsche Positive (korrekte Konstruktionen werden als Fehler markiert). Finden Sie drei möglichst unterschiedliche englische Beispiele für solche falsch positiven Fälle, bei denen die Regeln auf der Abfolge der Wörter basieren (nicht auf Zeichensetzungs- oder Formatierungsfehlern)¹ und geben Sie die jeweils angewendete Regel an.²³ (Es kann eine Weile dauern, bis Sie drei geeignete Beispiele finden.) Beantworten Sie für jedes Ihrer 3 Beispiele die folgenden Fragen:

- Warum funktioniert die Regel nicht wie beabsichtigt?
- Ließe sich die Regel so abändern, dass das jeweilige Beispiel korrekt behandelt wird? Wenn ja, wie? Wenn nein, warum nicht?

Aufgabe 4.2

- Schreiben Sie eine kontextfreie Grammatik, die Nominalphrasen wie die folgenden erzeugt:

Det N (*das Auto*)

Det A N (*das neue schnelle Auto*)

Det A N Prp Det N (*das grüne Auto auf dem Parkplatz*)

Det N Prp Det N Prp Det A N (*das Auto auf dem Parkplatz bei dem neuen Institutsgebäude*)

Det N Prp PN (*das Auto von Peter*) **Det N Prp Pro** (*das Auto von ihm*)

Verwenden Sie zusätzliche Kategoriensymbole (z.B. PP für Präpositionalphrase und AP für Adjektivphrase). Schreiben Sie außerdem einige lexikalische Einträge für jede lexikalische Kategorie.

- Fügen Sie die NP-Regeln aus (a) zur Grammatik G1 aus den aktuellen Vorlesungsfolien hinzu und leiten Sie drei unterschiedliche Sätze ab (bitte mit den zugehörigen Ableitungsbäumen; der komplette Ableitungsprozess braucht nicht aufgeschrieben zu werden). Mindestens zwei der drei Sätze sollen ziemlich lang sein (≥ 10 Wörter); bitte strukturell möglichst unterschiedliche Sätze ableiten.

¹Mit 'Show other random examples' können Sie neue Beispiele erzeugen.

²Klicken Sie dazu auf den Link "[Visit Rule]".

³Die Software verwendet das Penn Tagset, eine Übersicht finden Sie z.B. auf <http://www.computing.dcu.ie/~acahill/tagset.html>. Zusätzlich sind folgende Tags definiert:

NN:U - Mass noun und NN:UN - Noun used as mass

- (c) Gibt es mit der Grammatik Probleme? Ableitbare Ketten, für die es keine grammatischen Sätze gibt; grammatische Sätze, die eigentlich in den Bereich der Grammatik fallen sollten, aber nicht von ihr erzeugt werden? Bitte geben Sie jeweils ein illustrierendes Beispiel dazu!

Aufgabe 4.3

Das Stuttgart-Tübinger Tagset (STTS) wurde dazu entwickelt, deutsche Texte mit feinkörnigen Wortartinformationen zu annotieren. Bestimmen Sie für die folgenden Sätze die Wortarten nach dem STTS (vgl. das Handout aus der Übung oder <http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html>). Markieren Sie außerdem alle NPs (können geschachtelt sein). Welche werden von Ihren Regeln aus Aufgabe 4.2 erkannt (das Vorhandensein entsprechender lexikalischer Einträge vorausgesetzt), welche nicht? Abstrahieren Sie davon, dass bestimmte Kategorien im STTS anders als in unserer Grammatik benannt sind (z.B. Det vs. ART).

- (a) Die ehemaligen, vor 1815 teilweise selbständigen Territorien, die auf dem Gebiet des heutigen Saarlandes existierten, standen immer wieder zeitweise unter französischem Einfluss, insbesondere im späten 17. Jahrhundert sowie im Zeitraum zwischen der Französischen Revolution und dem Zweiten Pariser Frieden.
- (b) Peter fragte sich, warum er nicht einfach ins Schwimmbad gegangen ist, anstatt an diesem furchtbar heißen Tag für irgendeine Klausur, die er in drei Wochen schreiben würde, zu lernen.

Aufgabe 4.4

Entwerfen Sie ein Grammatikfragment für eine Sprache, die nicht Deutsch und wenn möglich auch nicht Englisch ist. Ihr Grammatikfragment sollte mindestens 8 nichtlexikalische Regeln beinhalten (d.h. Regeln, die nicht die Form wie z.B. *ART* → *der* haben), verschiedene Sätze der Sprache ableiten können und sich im Bereich der Wort-/Satzstellung vom Deutschen unterscheiden. Erklären Sie Ihre vom Deutschen abweichenden Kategoriensymbole. Geben Sie auch lexikalische Einträge an, geben Sie mindestens 3 Beispielleitungen an und erklären Sie diese. Bitte übersetzen Sie Ihre Beispiele so, dass auch jemand, der die Sprache nicht beherrscht, Ihre Beispielleitungen nachvollziehen kann, indem sie zusätzlich zu Ihrer tatsächlichen Übersetzung noch eine Wort-für-Wort-Übersetzung angeben.

Beispiel für eine Wort-für-Wort-Übersetzung:

Je le lui ai donné.

Ich es (dirObj) ihm habe gegeben.

Ich habe es ihm gegeben.

Abgabe in Gruppen von bis zu drei Studierenden bis **29.11.2011** 10 Uhr entweder als Email im pdf-Format an i2cl@coli.uni-sb.de oder auf Papier im Briefkasten an der Tür von Raum 1.04 in C7.2.