

Übung zur Vorlesung "Computerlinguistik II / Sprachtechnologie"

Sommersemester 2018, Prof. Dr. Udo Hahn, Tobias Kolditz

Übungsblatt 1 vom 07.05.2018

Abgabe bis 14.05.2018 per E-Mail (PDF) an

tobias.kolditz@uni-jena.de

Aufgabe 1 : "Recap" – Beschränkungen von Sprachen

8 pt

Gegeben sei das Alphabet $\Sigma = \{a, b\}$.

a) Typ-3-Beschränkung

3 pt

Eine Typ-3-Sprache ist in ihrer Ausdrucksfähigkeit stark beschränkt. Geben Sie in Potenzschreibweise eine Sprache über Σ an, die diese Beschränktheit **möglichst einfach** darstellt. Das heißt, die Sprache soll durch eine Typ-3-Grammatik nicht erzeugt werden können.

b) *Pumping-Lemma für reguläre Sprachen

3 pt

Zeigen Sie mithilfe des Pumping-Lemmas für reguläre Sprachen (s. Definition unten), dass es sich bei dem von Ihnen gewählten Beispiel nicht um eine reguläre Sprache handelt.

Hinweis: Das Pumping-Lemma macht eine Aussage der Form "Wenn L eine reguläre Sprache ist, dann hat L Eigenschaft E ". Wenn wir nachweisen, dass E für unsere Sprache nicht gilt, können wir folgern, dass es sich nicht um eine reguläre Sprache handelt (modus tollens).

c) Typ-2-Beschränkung

2 pt

Typ-2-Sprachen sind zwar weniger beschränkt, aber auch sie scheitern für bestimmte Strukturen. Geben Sie also auch eine **möglichst einfache & eindeutige** Sprache über Σ in Potenzschreibweise an, die eine Typ-2-Grammatik nicht mehr erzeugen kann.

Aufgabe 2 : Annotation

3 pt

Betrachten Sie folgenden Text:

“Die letzten Sekunden des Bundespräsidenten Horst Köhler lösen widersprüchliche Gefühle aus. Wie er da Hand in Hand mit seiner Frau dem Ausgang des Schlosses Bellevue zustrebte, das hatte etwas Anrührendes. Man fragte sich, wie übel dem Mann mitgespielt worden sein muss, dass er so geht.”

Die Wörter des oben gegebenen Texts sollen auf ihre Wortarten hin annotiert werden. Nutzen Sie für die Annotation der Wortarten Tags (= Labels) aus folgendem Tag-Set:

Tag	Beschreibung	Beispiel
ADJA	attributives Adjektiv	[das] große [Haus]
ADJD	adverbiales oder prädikatives Adjektiv	[er fährt] schnell, [er ist] schnell
ADV	Adverb	schon, bald, doch
ART	bestimmter oder unbestimmter Artikel	der, die, das, ein, eine
K	Konjunktion	weil, dass, damit, wenn, ob
N	Nomen	Tisch, Herr, Hans
P	Pronomen	er, mein, sich, jener, kein [Mensch], irgendein [Glas]
PR	Präposition	in [der Stadt], ohne [mich]
PTKVZ	abgetrennter Verbzusatz	[er kommt] an, [er fährt] rad
V	Verb	gehen, stand, gelegen
\$,	Komma	,
\$.	Satzbeendende Interpunktion	. ? ! ; :

Bauen Sie eine Tabelle auf, in der Sie jedem Wort ein Tag zuweisen.

Pumping-Lemma für reguläre Sprachen

Wenn L eine reguläre Sprache ist, dann gibt es eine Zahl $p \in \mathbb{N}$ (Pumpzahl), so dass für jedes Wort $w \in L$ mit $|w| \geq p$ folgendes gilt: Es gibt eine Zerlegung von w in drei Teilwörter $w = xyz$ mit

1. $|xy| \leq p$
2. $|y| > 0$
3. $xy^i z \in L$ für alle $i \geq 0$.