

Programmierkurs Python II – SS 2013

Übung 7

1 Chomsky-Normalform (3 + 2 Punkte)

1. Implementiere eine Funktion `is_cnf(rules)`, die testet, ob eine Grammatik in CNF vorliegt (keine Tilgungs-Regeln, genau zwei NTe auf der rechten Seite) (1 Punkt)
2. Implementiere eine Funktion `cnf(rules)`, die eine Grammatik als Eingabe nimmt und eine äquivalente Grammatik zurückgibt, die nur Regeln enthält, deren rechte Seite aus ein oder zwei NT-Symbolen besteht (fast CNF). (2 Punkte)
3. *Bonus*: Erweitere die Implementierung von `cnf` so, dass die Ausgabegrammatik keine Kettenregeln mehr enthält (also tatsächlich in CNF vorliegt). (2 Punkte)

Für alle Aufgaben soll die Eingabegrammatik als Liste von Regeln repräsentiert werden (siehe Template). Der Einfachheit halber kannst Du davon ausgehen, dass a) es keine zyklischen Regeln gibt ($A \rightarrow B, B \rightarrow A$) und b) das eigentliche Lexikon (Regeln wie $DET \rightarrow \text{der}$) vom Rest der Grammatik getrennt wurde, in den Regeln selbst also keine Terminal-Symbole vorkommen.

2 CYK (3 + 2 Punkte)

Implementiere eine Methode `recognize_cyk(start, rules, lexicon, sentence)`, die den CYK-Algorithmus als Erkenner umsetzt. Eingabe sind ein Startsymbol, eine Grammatik, ein Lexikon und einen Satz nimmt und `True` zurückgibt, falls der Eingabesatz in der Sprache der Grammatik enthalten ist (sonst `False`).

Die Methode soll eine `NotWellFormedException` werfen, falls die Grammatik nicht in CNF vorliegt. (Siehe Kriterien aus Aufgabe 1.1.)

Bonus: Implementiere eine Methode `parse_cyk(start, rules, lexicon, sentence)`, die den Erkenner zum Parser erweitert. Der Parser soll als Generator implementiert werden und alle Ableitungsbäume für den Eingabesatz zurückgeben. Benutze die vorgegebene Klasse für Parsebäume. (Explizite Back-Pointer mit nummerierten Spans sind nicht notwendig; die korrekten Bäume reichen aus.)

3 Earley-Parser (Bonus, 4 Punkte)

Implementiere eine Methode `parse_earley(sentence, rules, lexicon)`, die den Earley-Parser umsetzt und als Generator alle Parsebäume zurückgibt. Es bietet sich hierbei an, als Hilfs-Struktur eine Chart und spezielle Chart-Items zu implementieren.

Abgabe bis Donnerstag, 20.06.2013, 11:00 Uhr