

4. Übungsblatt – Abgabe: 28.11.2021, 23:55

Aufgabe 4.1

Der folgende Satz hat mehrere Lesarten:

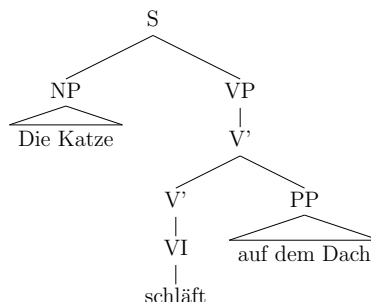
„Der Mann holte die Kiste mit der Leiter aus dem Regal“

- Charakterisieren Sie durch eindeutige Paraphrasierung die einzelnen Lesarten, dh. geben die diese in eigenen Worten wieder. (Es sind einige! Sie sollten mindestens 4 strukturell verschiedene Lesarten finden, also Lesarten mit unterschiedlichen Strukturbäumen.)
- Welche dieser Lesarten sind unwahrscheinlicher als andere? Warum? Welches Wissen braucht ein Computerprogramm, um diese Lesarten als unwahrscheinlich zu erkennen?
- Leiten Sie die Strukturbäume vom Originalsatz für zwei strukturell verschiedene Lesarten ab. Dazu können Sie die untenstehende Grammatik von den aktuellen Vorlesungsfolien benutzen; Lexikoneinträge ergänzen Sie bitte selbst. Nicht alle möglichen Lesarten haben verschiedene Strukturbäume. Bitte markieren Sie, welche Lesarten zu welchem Baum gehören.

$S \rightarrow NP VP$	$NP \rightarrow \text{Art } N' \mid EN$	$N' \rightarrow N' PP$
$N' \rightarrow N (NP)$	$VP \rightarrow V'$	$V' \rightarrow V' PP$
$V' \rightarrow VT NP$	$V' \rightarrow VI$	$PP \rightarrow P NP$

Hinweis 1: Eine Paraphrase ist eine Umformulierung eines Satzes, die eine bestimmte Lesart dieses Satzes eindeutig beschreibt, z.B durch Benutzung von Relativsätzen.

Hinweis 2: Sie dürfen in Aufgabenteil c) NPs oder PPs die keine weiteren Modifier oder Argumente haben, im Baum abkürzen:



Aufgabe 4.2

Gegeben ist die Sprache $L_1 = \{ww^R \mid w \in \{a,b\}^+\}$ über dem Alphabet a, b , wobei w^R die Spiegelung von w ist, d.h. es enthält die Zeichen von w in umgekehrter Reihenfolge. L_1 enthält also alle (nichtleeren) Wörter, die vorwärts wie rückwärts gelesen gleich sind, z.B. *abaaaaaba*.

Außerdem ist die Sprache $L_2 = \{wcw^R \mid w \in \{a,b\}^*\}$ gegeben. Wörter aus L_2 sind also z.B. *c*, *abcba*, *bbbaabacabaabbb*.

- (a) Geben Sie für L_1 und L_2 jeweils einen möglichst intuitiven PDA, der ohne Kategoriensymbole (z.B. *S*) auf dem Speicher auskommt (vgl. Folie 14).
- (b) Geben Sie außerdem für L_1 und L_2 jeweils eine kontextfreie Grammatik an.
- (c) Erzeugen Sie aus jeder dieser Grammatiken nach dem Schema auf den aktuellen Vorlesungsfolien jeweils einen PDA als Top-Down- und Bottom-Up-Parser.
- (d) Geben Sie für jeden Ihrer erzeugten PDAs (also insgesamt 6) an, ob sie deterministisch sind, oder nicht.

Aufgabe 4.3

Gegeben ist die folgende Grammatik:

$S \rightarrow NP VP$
 $VP \rightarrow VP ADV$
 $VP \rightarrow \text{geht}$
 $NP \rightarrow \text{Hans}$
 $ADV \rightarrow \text{heute} \mid \text{schnell}$

Betrachten Sie den Satz „Hans geht schnell“.

- (a) Warum kann ein Top-down-Parser (Tiefensuche mit Backtracking, wir probieren passende Regeln in der Reihenfolge aus, wie sie in der Grammatik stehen) mit dieser Grammatik den Satz nicht parsen? Wie lässt sich dieses Problem durch Anpassen der Grammatik beheben? (Tipp: Die Reihenfolge der Regeln ist entscheidend.)
- (b) Zeigen Sie, wie der Bottom-up-Parser (=Shift-Reduce-Parser) diesen Satz verarbeiten würde. Geben Sie für jeden Schritt den Stack und die verwendete Grammatikregel an. Geben Sie auch den dabei entstehenden Parsebaum an. (Nur einmal, Sie brauchen keine Zwischenergebnisse als Bäume anzugeben.)

Abgabe via Moodle. Bei Fragen posten Sie im MS Teams Fragenchannel oder besuchen Sie die Helpsession am Freitag von 16:15 bis 17:45 Uhr auf Teams.