Übungsblatt 11

Aufgabe 1. Gegeben ist die kontextfreie Grammatik $G = (V, \Sigma, P, S)$ in Chomsky-Normalform über $\Sigma = \{a, b\}$ mit $V = \{S, X, Y, A, B\}$ und den folgenden Produktionen:

$$\begin{aligned} P: \ S &\to a \mid b \mid AA \mid BB \mid XA \mid YB \\ X &\to AS \\ Y &\to BS \\ A &\to a \\ B &\to b \end{aligned}$$

(a) Überprüfen Sie mit dem CYK-Algorithmus, ob $abbbba \in L(G)$ gilt.

(b) Welche Sprache erzeugt G?

Lösung zu Aufgabe 1. (a)

	a	b	b	b	b	a
j = 1	S,A	S,B	S,B	S,B	S,B	S, A
j = 2	X	Y,S	Y,S	Y,S	Y	
j = 3	X	Y,S	Y,S	∅		
j = 4	X	Y,S	∅			
j = 5	X	∅				
j = 6	S					

Da S im letzten Feld steht, gilt $abbbba \in L(G)$.

Wir verwenden die Anleitung aus dem FSA-Skript (ab Folie 235) um die obige Tabelle auszufülllen.

(b) Wir erhalten einen besseren Überblick über die von G erzeugte Sprache, indem wir einige Nicht-Terminale eliminieren.

Zunächst streichen wir die Produktionen $A \to a$ und $B \to b$ und setzen den jeweiligen Buchstaben dort ein wo zuvor A oder B stand:

$$
\begin{aligned}
P' : S &\rightarrow a|b|aa|bb|Xa|Yb \\
X &\rightarrow aS \\
Y &\rightarrow bS
\end{aligned}
$$

Dann können wir X auf rechten Seiten durch aS und Y durch bS ersetzen und die entsprechenden Regeln für X und Y löschen:

$$P'' : S \rightarrow a|b|aa|bb|aSa|bSb$$

Die Grammatik $G' = (V, \Sigma, P'', S)$ ist eine zu G äquivalente Grammatik, also gilt $L(G) = L(G')$. In dieser vereinfachten Form sehen wir, dass G' alle nicht-leeren Palindrome über dem Alphabet $\Sigma = \{a, b\}$ erzeugt.

In Mengenschreibweise:

$L(G') = L(G) = \{w \in \Sigma^+ \mid w = w^r\}$.

Aufgabe 2. Geben Sie Kellerautomaten und kontextfreie Grammatiken an, die die folgenden Sprachen akzeptieren.

(a) $\{a^n b^n \mid n \geq 0\}$

(b) $\{w \in \{a, b\}^* \mid \text{Die Anzahl der } a\text{'s und } b\text{'s ist gleich.}\}$

(c) $\{w \in \{a, b\}^* \mid w = w^r\}$

Lösung zu Aufgabe 2. Die Grammatiken, die im Folgenden angegeben werden, erfüllen zum Teil nicht die ε-Sonderregel und sind daher streng genommen nicht kontextfrei im Sinne der Definition von Folie 32. Die Grammatiken können jedoch in kontextfreie Grammatiken, die die ε-Sonderregel erfüllen, umgewandelt werden (siehe dazu Folie 193).
Zusätzliche Aufgabe (Bonus): Modifizieren Sie die angegebenen Grammatiken so, dass die ε-Sonderregel erfüllt ist und die Grammatiken somit auch im strengeren Sinne kontextfrei sind.

(a) **Grammatik**

- $G_1 = (V, \Sigma, P, S)$
- $V = \{S\}$
- $\Sigma = \{a, b\}$
- $P = \{S \rightarrow \varepsilon \mid aSb\}$

Kellerautomat

- $M_1 = (Z, \Sigma, \Gamma, \delta, z_0, \#)$
- $Z = \{z_0, z_1\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{\#, A\}$

Dabei ist δ wie folgt definiert:

$$\begin{aligned}\delta : (z_0, \varepsilon, \#) &\to (z_1, \varepsilon)\\ (z_0, a, \#) &\to (z_0, A\#)\\ (z_0, a, A) &\to (z_0, AA)\\ (z_0, b, A) &\to (z_1, \varepsilon)\\ (z_1, b, A) &\to (z_1, \varepsilon)\\ (z_1, \varepsilon, \#) &\to (z_1, \varepsilon)\end{aligned}$$

Die Idee hinter diesem Kellerautomaten ist, dass für jedes a des Teilworts a^n ein A auf den Keller gelegt wird, und für jedes b des Teilworts b^n ein A vom Keller entfernt wird.

(b) **Grammatik**

- $G_2 = (V, \Sigma, P, S)$
- $V = \{S\}$
- $\Sigma = \{a, b\}$
- $P = \{S \to \varepsilon \mid aSbS \mid bSaS\}$

Idee: Auf jedes a muss irgendwann ein zugehöriges b folgen und umgekehrt. Davor und dahinter können wir mit einem S das Wort weiter verlängern oder die Konstruktion durch die ε-Transition beenden (siehe Blatt 3, Aufgabe 4).

Kellerautomat

- $M_2 = (Z, \Sigma, \Gamma, \delta, z_0, \#)$
- $Z = \{z_0\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{\#, P, N\}$

Dabei ist δ wie folgt definiert:

$$
\begin{aligned}
\delta : &(z_0, \varepsilon, \#) \rightarrow (z_0, \varepsilon)\\
&(z_0, a, \#) \rightarrow (z_0, P\#)\\
&(z_0, b, \#) \rightarrow (z_0, N\#)\\
&(z_0, a, P) \rightarrow (z_0, PP)\\
&(z_0, b, P) \rightarrow (z_0, \varepsilon)\\
&(z_0, a, N) \rightarrow (z_0, \varepsilon)\\
&(z_0, b, N) \rightarrow (z_0, NN)
\end{aligned}
$$

Wir verwenden den Keller um die Differenz $\#_a(w) - \#_b(w)$ zu speichern, wobei $\#_a(w)$ die Anzahl a's im Wort w ist, $\#_b(w)$ die Anzahl an b's.

Positive Werte werden durch eine Folge von P's codiert, negative Werte durch eine Folge von N's.

Enthält das eingelesene Wort gleichviele a's und b's, ist der Keller schließlich leer.

Bemerkung: Diese Konstruktion ist analog zur Kodierung der ganzen Zahlen als Differenz aus zwei natürlichen Zahlen (für jede Zahl $z \in \mathbb{Z}$ gibt es natürliche Zahlen x und y mit $z = x - y$). Deswegen brauchen wir beide Kellerzeichen P und N.

(c) **Grammatik**

- $G_3 = (V, \Sigma, P, S)$
- $V = \{S\}$
- $\Sigma = \{a, b\}$
- $P = \{S \rightarrow \varepsilon \mid a \mid b \mid aSa \mid bSb\}$

Kellerautomat

- $M_3 = (Z, \Sigma, \Gamma, \delta, z_0, \#)$
- $Z = \{z_0, z_1\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{\#, A, B\}$

Dabei ist δ wie folgt definiert:

$$
\begin{aligned}
\delta : (z_0, \varepsilon, \#) &\to (z_0, \varepsilon) && (1)\\
(z_0, a, \#) &\to (z_0, A\#) && (1)\\
(z_0, b, \#) &\to (z_0, B\#) && (1)\\
(z_0, a, A) &\to (z_0, AA) && (1)\\
(z_0, b, A) &\to (z_0, BA) && (1)\\
(z_0, a, B) &\to (z_0, AB) && (1)\\
(z_0, b, B) &\to (z_0, BB) && (1)\\
\\
(z_0, \varepsilon, A) &\to (z_1, A) && (2)\\
(z_0, \varepsilon, B) &\to (z_1, B) && (2)\\
(z_0, a, \#) &\to (z_1, \varepsilon) && (2)\\
(z_0, b, \#) &\to (z_1, \varepsilon) && (2)\\
(z_0, a, A) &\to (z_1, A) && (2)\\
(z_0, b, A) &\to (z_1, A) && (2)\\
(z_0, a, B) &\to (z_1, B) && (2)\\
(z_0, b, B) &\to (z_1, B) && (2)\\
\\
(z_1, a, A) &\to (z_1, \varepsilon) && (3)\\
(z_1, b, B) &\to (z_1, \varepsilon) && (3)\\
(z_1, \varepsilon, \#) &\to (z_1, \varepsilon) && (3)
\end{aligned}
$$

Falls $w \in L$, muss $w = xx^r$ oder $w = xmx^r$ gelten, mit $x \in \{a, b\}^*$, $m \in \{a, b\}$.

In (1) wird das Wort in richtiger Reihenfolge auf den Stack gelegt und in (3) rückwärts ausgelesen und mit dem Rest des Wortes verglichen.

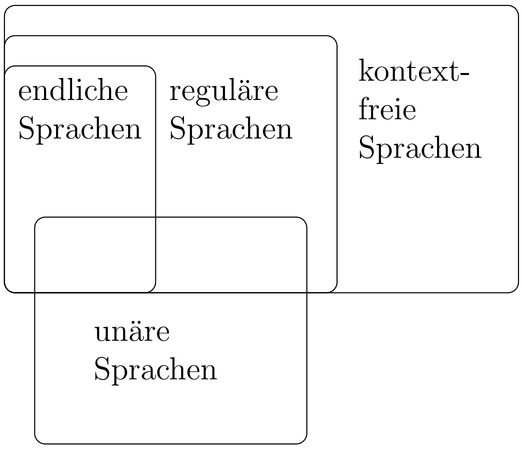
Der Übergang, d.h. die Mitte des Wortes, wird mit (2) nicht-deterministisch geraten, mit einem ε-Übergang für den Fall $w = xx^r$ und einem a/b-Übergang für den Fall $w = xmx^r$.

Es ist erlaubt, die Übergänge etwas kompakter hinzuschreiben. Zum Beispiel kann man die vierte und sechste Zeile zu $(z_0, a, C) \to (z_0, AC)$ $(C \in \{A, B\})$ zusammenfassen.

Aufgabe 3. Stellen Sie die folgenden Sprachklassen über dem Alphabet $\{a, b\}$ in einem Venn-Diagramm dar. Geben Sie zusätzlich Beispielsprachen für alle Teilbereiche in ihrem Diagramm an.

- Reguläre Sprachen
- Kontextfreie Sprachen
- Endliche Sprachen
- Unäre Sprachen (alle Sprachen L mit $L \subseteq \{a\}^*$)

Lösung zu Aufgabe 3.



- endlich und unär: $\{a\}$
- endlich und nicht unär: $\{ab\}$
- regulär, unär und nicht endich: $\{a^n \mid n \in \mathbb{N}\}$
- regulär, nicht unär und nicht endlich: $\{ab^n \mid n \in \mathbb{N}\}$
- unär, nicht kontextfrei: $\{a^p \mid p \text{ ist Primzahl}\}$
- nicht unär, nicht regulär, kontextfrei: $\{a^n b^n \mid n \in \mathbb{N}\}$
- nicht unär, nicht kontextfrei: $\{ww \mid w \in \{a, b\}^*\}$

<u>Hinweis</u>: Jede kontextfreie, unäre Sprache ist auch regulär (siehe Vorlesung).