

Übungsblatt 12

Aufgabe 1. Geben Sie Kellerautomaten und kontextfreie Grammatiken an, die die folgenden Sprachen akzeptieren.

- (a) $\{a^n b^n \mid n \geq 0\}$
- (b) $\{w \in \{a, b\}^* \mid \#_a(w) = \#_b(w)\}$
- (c) $\{a^m b^n \mid m, n \geq 0, m \neq n\}$

Lösung zu Aufgabe 1. 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:

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

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**

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

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

- $G_3 = (V, \Sigma, P, S)$
- $V = \{S, A, B\}$
- $\Sigma = \{a, b\}$
- $P = \{S \rightarrow aSb \mid A \mid B, A \rightarrow aA \mid a, B \rightarrow bB \mid b\}$

Über die Regel $S \rightarrow aSb$ erzeugen wir Wörter mit gleich vielen a 's und b 's.

Die Übergänge $S \rightarrow A$ und $S \rightarrow B$ fügen in der Mitte dann noch ein Wort a^n oder b^n mit $n \geq 1$ ein, so dass am Ende entweder mehr a 's oder mehr b 's erzeugt wurden.

Kellerautomat

- $M_3 = (Z, \Sigma, \Gamma, \delta, z_0, \#)$
- $Z = \{z_0, z_1, z_2, z_3, s_0, s_1, q_0\}$
- $\Sigma = \{a, b\}$
- $\Gamma = \{\#, A\}$

Dabei ist δ wie folgt definiert:

$$\delta : (z_0, a, \#) \rightarrow (z_1, \#) \quad (1)$$

$$(z_1, a, \#) \rightarrow (z_1, \#) \quad (1)$$

$$(z_1, \varepsilon, \#) \rightarrow (z_1, \varepsilon) \quad (1)$$

$$(z_0, b, \#) \rightarrow (z_2, \#) \quad (2)$$

$$(z_2, b, \#) \rightarrow (z_2, \#) \quad (2/4)$$

$$(z_2, \varepsilon, \#) \rightarrow (z_2, \varepsilon) \quad (2/4)$$

$$(z_0, a, \#) \rightarrow (z_3, A\#) \quad (3/4)$$

$$(z_3, a, A) \rightarrow (z_3, AA) \quad (3/4)$$

$$(z_3, b, A) \rightarrow (s_0, \varepsilon) \quad (3)$$

$$(s_0, b, A) \rightarrow (s_0, \varepsilon) \quad (3)$$

$$(s_0, \varepsilon, A) \rightarrow (s_1, \varepsilon) \quad (3)$$

$$(s_1, \varepsilon, A) \rightarrow (s_1, \varepsilon) \quad (3)$$

$$(s_1, \varepsilon, \#) \rightarrow (s_1, \varepsilon) \quad (3)$$

$$(z_3, b, A) \rightarrow (q_0, \varepsilon) \quad (4)$$

$$(q_0, b, A) \rightarrow (q_0, \varepsilon) \quad (4)$$

$$(q_0, b, \#) \rightarrow (z_2, \#) \quad (4)$$

Lassen Sie sich nicht abschrecken von den vielen Transitionen, denn das Prinzip ist einfacher als es aussieht. Wir unterscheiden nichtdeterministisch 4 Typen von Wörtern, die zur Sprache gehören:

- (1) a^n für $n \geq 1$
- (2) b^n für $n \geq 1$
- (3) $a^m b^n$ mit $m, n \geq 1$ und $m > n$

(4) $a^m b^n$ mit $m, n \geq 1$ und $m < n$

Die ersten beiden Fälle (1) und (2) sind einfach, da entweder nur a 's oder b 's gelesen werden und anschließend kann man das Kellerbodensymbol entfernen. In den Fällen (3) und (4) wird zu Beginn für jedes gelesene a ein A auf den Keller gelegt bis das erste b gelesen wird. Anschließend wird im Fall (3) für jedes b ein A entfernt, und dann wechselt man mit einem ε -Übergang in den Zustand s_1 falls weiterhin ein A auf dem Keller liegt (und somit mehr a 's als b 's gelesen wurden). Der Zustand s_1 leert anschließend den Keller ohne weitere Buchstaben zu lesen. Beachten Sie, dass falls man den ε -Übergang anwendet bevor das Wort zu Ende gelesen wurde, so wird am Ende nicht akzeptiert, da im Zustand s_1 keine weiteren Buchstaben gelesen werden.

Im Fall (4) wird ebenfalls für jedes b ein A vom Keller entfernt, aber anschließend wechselt man in den Zustand z_2 , falls man das Kellerbodensymbol $\#$ erreicht und ein b gelesen wird (und somit mehr b 's als a 's gelesen wurden). Im Zustand z_2 werden dann noch beliebig viele b 's gelesen bevor man das Kellerbodensymbol entfernt.

Sie dürfen für besseres Verständnis von δ auch die Diagramm-Repräsentation nutzen (Erklärung siehe Seite 6):

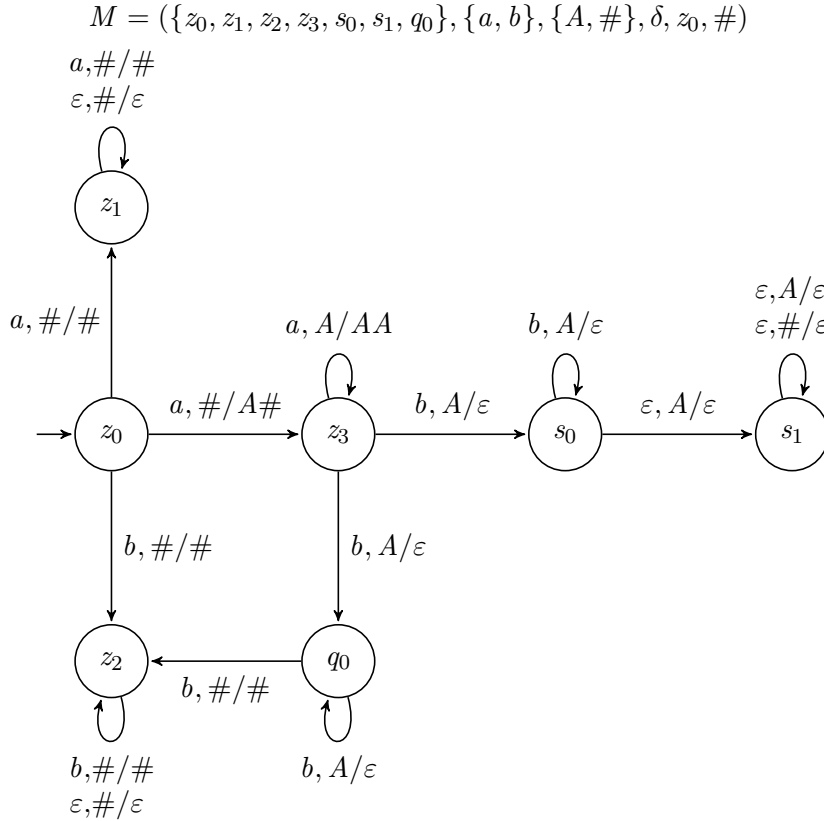
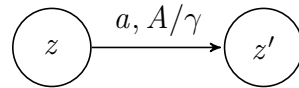


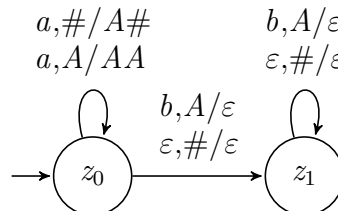
Diagramm-Repräsentation von Kellerautomaten

Sei $M = (Z, \Sigma, \Gamma, \delta, z_0, \#)$ ein Kellerautomat. Wir können auch Kellerautomaten in Diagrammform darstellen, wie wir das bereits bei endlichen Automaten getan haben. Zustände aus Z , sowie auch der Startzustand z_0 , werden genauso dargestellt wie bei endlichen Automaten, also als Kreise bzw. im Falle von z_0 ein Kreis mit einem eingehenden Pfeil. Für einen Übergang $(z, a, A) \rightarrow (z', \gamma)$ aus δ (d.h. $z, z' \in Z, a \in \Sigma, A \in \Gamma$ und $\gamma \in \Gamma^*$) schreiben wir:



Schleifen sind ausdrücklich erlaubt. Außerdem haben viele Kellerautomaten nur einen Zustand (wie bei der (b)), d.h. alle Übergänge gehen wieder nach z_0 . Bei solchen Beispielen hilft die graphische Veranschaulichung dem Verständnis nicht so viel, wie z.B. verglichen mit der (a):

$$M = (\{z_0, z_1\}, \{a, b\}, \{A, \#\}, \delta, z_0, \#)$$



Man erkennt hier sofort, dass es eine Phase gibt, in der nur a 's gelesen werden und eine Phase, in der b 's nur gelesen werden können, wenn zeitgleich A 's im Keller weggelöscht werden.

Bedenken Sie, dass es verschiedene Definitionen von Kellerautomaten gibt. Wir haben hier keine Endzustände, also auch keine Doppelkreise! Es ist zudem notwendig, das Tupel mit aufzuschreiben für die volle Punktzahl in der Klausur (sonst geht z.B. nicht hervor, was das Kellerbodensymbol ist).

Aufgabe 2. Eine kontextfreie Grammatik, die die Sprache

$$L = \{w \in \{a, b\}^* \mid w = w^r\}$$

erzeugt, ist $G = (V, \Sigma, P, S)$ mit $V = \{S\}$, $\Sigma = \{a, b\}$ und

$$P = \{S \rightarrow \varepsilon \mid a \mid b \mid aSa \mid bSb\}.$$

Wandeln Sie die Grammatik mit dem Verfahren der Vorlesung (Folie 279) in einen PDA um. Überprüfen Sie, ob ihr Automat richtig arbeitet, indem Sie ihn auf den Eingaben

- $w_1 = abbabba$ und
- $w_2 = abb$

laufen lassen.

Lösung zu Aufgabe 2. Nach dem Verfahren der Vorlesung erhalten wir einen Kellerautomaten $M = (\{z\}, \Sigma, V \cup \Sigma, \delta, z, S)$ mit δ gegeben durch

$$\begin{aligned}\delta(z, \varepsilon, S) &= \{(z, \varepsilon), (z, a), (z, b), (z, aSa), (z, bSb)\}, \\ \delta(z, a, a) &= \{(z, \varepsilon)\}, \\ \delta(z, b, b) &= \{(z, \varepsilon)\}.\end{aligned}$$

Wir sehen sofort, dass w_1 ein Palindrom ist. Wir haben:

$$\begin{aligned}(z, abbabba, S) &\vdash (z, abbabba, aSa) \vdash (z, bbabba, Sa) \vdash (z, bbabba, bSba) \\ &\vdash (z, babba, Sba) \vdash (z, babba, bSbba) \vdash (z, abba, Sbba) \\ &\vdash (z, abba, abba) \vdash^* (z, \varepsilon, \varepsilon)\end{aligned}$$

Also wird w_1 von M akzeptiert. Für w_2 bekommen wir zum Beispiel

$$(z, abb, S) \vdash (z, abb, aSa) \vdash (z, bb, Sa) \vdash (z, bb, ba) \vdash (z, b, a)$$

Wegen des Nichtdeterminismus müssen wir uns noch überlegen, dass es nicht doch eine Möglichkeit gibt $(z, \varepsilon, \varepsilon)$ zu erhalten: Wir müssen das Eingabewort komplett abarbeiten, daher müssen wir im ersten Schritt einen Übergang nehmen, bei dem S erhalten bleibt und dafür kommt (wegen erstes Zeichen a) nur einer in Frage ($S \rightarrow aSa$). Das Wort w_2 hat Länge 3 und wir wissen daher, dass im nächsten Schritt die Mitte gelesen werden muss. Es bleibt daher nur den zu $S \rightarrow b$ korrespondierenden Übergang zu nehmen. Obwohl wir bis hierher also stets die sinnvollsten Übergänge gewählt haben, bleibt der Automat nun bei (z, b, a) stehen.

Es gäbe zum Beispiel auch die Möglichkeit das Eingabewort komplett abzuarbeiten, wodurch wir (z, ε, bba) erhalten. Hier ist allerdings der Keller nicht leer.

Aufgabe 3. Nutzen Sie Abschlusseigenschaften kontextfreier Sprachen, um zu zeigen oder zu widerlegen, dass folgende Sprachen kontextfrei sind:

- (a) $L_1 = \{a^{n_1} b^{n_1} a^{n_2} b^{n_2} \dots a^{n_k} b^{n_k} \mid n_1, n_2, \dots, n_k, k \geq 0\}$
- (b) $L_2 = \{w \in \{a, b, c\}^* \mid \#_a(w) = \#_b(w) = \#_c(w)\}$
- (c) $L_3 = \{a^n b^n c^m d^n \mid n, m \geq 0\}$
- (d) $L_4 = \{x^n y^n \mid x, y \in \Sigma, n \geq 0\}$, Σ beliebig

Hinweis: Sie dürfen das Wissen über die Sprachen von Blatt 10-11 benutzen, von dem Sie wissen, dass sie nicht kontextfrei sind.

- (e) Finden Sie den Fehler in folgendem „Beweis“: Für jedes $n \in \mathbb{N}$ ist $L_n = \{a^n b^n c^n\}$ kontextfrei (sogar regulär), da L_n eine einelementige Menge ist. Außerdem sind kontextfreie Sprachen unter Vereinigung abgeschlossen. Folglich ist induktiv $K_n = \bigcup_{k=0}^n L_k$ kontextfrei. Es folgt, dass auch der Grenzwert $K = \bigcup_{k=0}^{\infty} L_k = \{a^n b^n c^n \mid n \geq 0\}$ kontextfrei ist.

Lösung zu Aufgabe 3.

- (a) L_1 ist gleich $\{a^n b^n \mid n \geq 0\}^*$ und kontextfreie Sprachen sind unter Kleene-Stern abgeschlossen.
- (b) L_2 ist nicht kontextfrei, da $L = \{a^n b^n c^n \mid n \geq 0\}$ nicht kontextfrei ist und $L_2 \cap L(a^* b^* c^*) = L$ gilt. Denn wäre L_2 kontextfrei, so auch L , denn kontextfreie Sprachen sind unter Schnitt mit regulären Sprachen abgeschlossen.
- (c) L_3 ist nicht kontextfrei, denn es gilt $h(L_3) = L$ mit $h(a) = a$, $h(b) = b$, $h(c) = \varepsilon$, $h(d) = c$ und kontextfreie Sprachen sind unter Homomorphismenbildung abgeschlossen.
- (d) Es gilt $L_4 = \bigcup_{x,y \in \Sigma} \{x^n y^n \mid n \geq 0\}$ und alle $\{x^n y^n \mid n \geq 0\}$ sind kontextfrei. Außerdem ist Σ per Definition endlich, also handelt es sich um eine endliche Vereinigung. Also ist L_4 kontextfrei (Abschluss unter Vereinigung). Für unendliche Vereinigungen kann man nicht so argumentieren (siehe (e)).
- (e) Eigenschaften von Sprachen unter unendlichen Vereinigungen bleiben grundsätzlich nicht unbedingt erhalten. Die Konstruktionen für NFAs (reguläre Sprachen) und kontextfreie Grammatiken funktionieren gut für Vereinigung von zwei und somit für beliebig viele (aber endlich viele) Sprachen. Würde sich das Ganze für unendliche Vereinigungen verallgemeinern lassen, so könnte man zeigen, dass **jede** Sprache endlich / regulär / kontextfrei / ... ist. Dies ist ein Widerspruch.

Deswegen: Immer vorsichtig sein mit Grenzwerten!