Theorie der Informatik

G. Röger
Frühjahrssemester 2019

Universität Basel
Fachbereich Informatik

Übungsblatt 6 — Lösungen

Aufgabe 6.1 (Chomsky-Normalform; 2 Punkte)

Geben Sie eine Grammatik in Chomsky-Normalform an, die die gleiche Sprache generiert wie die Grammatik $G = \langle \Sigma, V, P, S \rangle$ mit $\Sigma = \{a, b, c\}$, $V = \{S, X, Y\}$ und den folgenden Regeln P:

$$\begin{aligned}
S &\to XY \\
X &\to c \\
X &\to cS \\
Y &\to abb \\
Y &\to aYb \\
Y &\to \epsilon
\end{aligned}$$

Lösung:

$G' = \langle \Sigma, V', P', S \rangle$ mit $V = \{S, X, Y, Z, A, B, C\}$ und P' besteht aus folgenden Regeln:

$$\begin{aligned}
S &\to XY \mid c \mid CS & \quad A &\to a \\
X &\to c \mid CS & B &\to b \\
Y &\to AZ \mid AB & C &\to c \\
Z &\to BB \mid YB & &
\end{aligned}$$

Aufgabe 6.2 (Länge von Ableitungen bei Chomsky-Normalform; 2 Punkte)

Sei G eine Grammatik in Chomsky-Normalform und $w \in \mathcal{L}(G)$ ein nicht-leeres Wort ($w \neq \varepsilon$), das von G erzeugt wird. Zeigen Sie, dass jede Ableitung von w aus der Startvariable von G genau $2|w| - 1$ Ableitungsschritte hat.

Lösung:

Eine Grammatik ist in Chomsky-Normalform, wenn jede Regel die Form (a) $A \to BC$ oder (b) $A \to a$ oder (c) $S \to \varepsilon$ hat, wobei A, B, C Nichtterminale (Variablen) sind, S die Startvariable ist, B und C nicht die Startvariable sind und a ein Terminalzeichen ist.

Sei w ein Wort der Länge $n > 0$ aus der Sprache, die von der Grammatik generiert wird. Da Regeln der Sorte (c) nur verwendet werden können, um das leere Wort abzuleiten, sind sie hier nicht relevant. Um die n Terminale von w einzuführen, muss jede Ableitung genau n mal Regeln der Sorte (b) anwenden. Jede dieser Regelanwendungen entfernt ein Nichtterminal, was bedeutet, dass für jede dieser Regelanwendungen vorher ein Nichtterminal existieren muss. Diese Nichtterminale müssen durch die Anwendung von Regeln vom Typ (a) eingeführt worden sein, wobei jede Anwendung die Zahl der Nichtterminale um 1 erhöht. Da die Startvariable bereits ein Nichtterminal ist, muss also *mindestens* $n - 1$ mal eine solche Regel angewendet werden. Da Nichtterminale nur von Regeln vom Typ (b) eliminiert werden können, können auch *höchstens* $n - 1$ Regelanwendungen vom Typ (a) in der Ableitung vorkommen. Daher besteht jede Ableitung von w aus exakt $n + n - 1 = 2n - 1$ Ableitungsschritten.

Aufgabe 6.3 (Kellerautomaten; 2 Punkte)

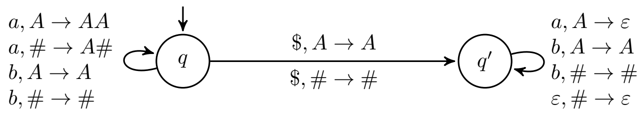
Geben Sie einen Kellerautomaten (PDA) an, der die Sprache

$$L = \{w_1\$w_2 \mid w_1, w_2 \in \{a, b\}^* \text{ und } w_1 \text{ und } w_2 \text{ enthalten gleich viele } a\}$$

über $\Sigma = \{a, b, \$\}$ akzeptiert.

Lösung:

$M = (\{q, q'\}, \{a, b, \$\}, \{A, \#\}, \delta, q, \#)$ mit folgendem δ:

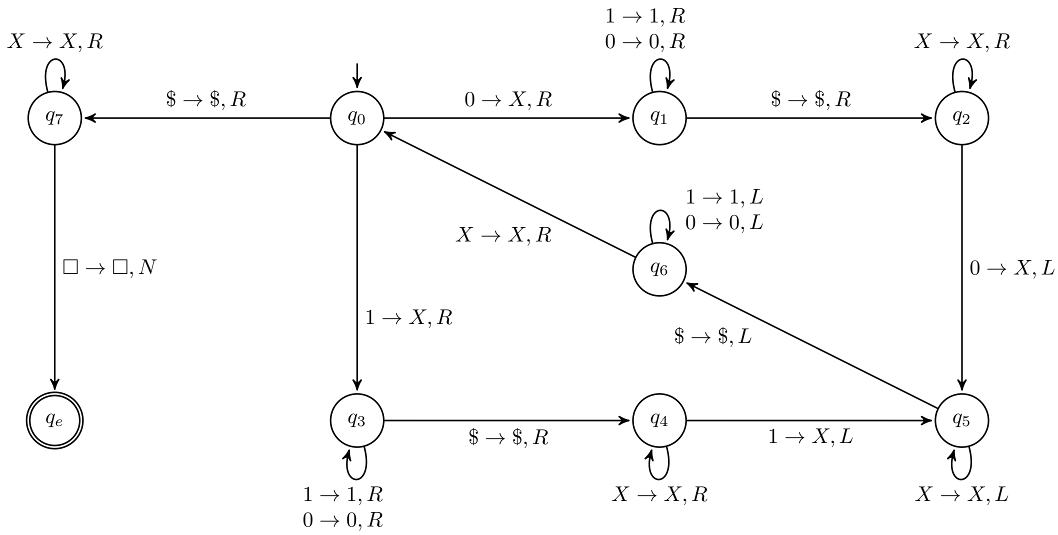


Aufgabe 6.4 (Nichtdeterministische Turing-Maschinen; 4 Punkte)

Betrachten Sie die Sprache $L = \{w\$w \mid w \in \{0,1\}^*\}$ über Alphabet $\{0, 1, \$\}$. Geben Sie das Zustandsdiagramm einer NTM M mit $\mathcal{L}(M) = L$ an. Erläutern Sie das Verhalten Ihrer TM auch in Worten.

Lösung:

$M = \langle \{q_0, q_1, \ldots, q_7, q_e\}, \{0, 1, \$\}, \{0, 1, \$, X, \square\}, \delta, q_0, \square, \{q_e\} \rangle$ mit folgendem δ:



Die TM streicht jeweils das nächste Symbol des linken Eingabeteils mit einem X ab, merkt es sich und versucht das gleiche Symbol im rechten Teilwort abzustreichen. Kann es auf diese Weise alle Symbole streichen, akzeptiert sie die Eingabe.

Immer wenn die TM in Zustand q_0 ist, steht der Schreib-Lesekopf auf dem ersten Eingabezeichen, das noch nicht abgestrichen wurde. Steht dort eine 0, wird sie durch ein X ersetzt und der Schreiblesekopf wird nach rechts über alle übrigen Zeichen des linken Eingabeteils (q_1), das Trennzeichen und alle bereits abgestrichenen Positionen des rechten Teilworts (q_2) bewegt. Wird dann eine 0 gelesen, wird sie durch ein X ersetzt und der Schreiblesekopf wird wieder nach links bewegt (über die abgestrichenen Teile des rechten Teilworts in q_5, das Trennzeichen und die noch nicht verarbeiteten Zeichen des linken Teilworts in q_6). Der Zykel $q_0, q_3, q_4, q_5, q_6, q_0$ arbeitet analog, streicht aber 1en ab.

Sind alle Zeichen des linken Teilworts auf diese Weise abgestrichen, liest die TM in q_0 ein \$. Sie muss nun noch verifizieren, dass auch rechts alle Zeichen abgestrichen sind, was auf dem Pfad über q_7 zu q_e geschieht.