Prof. J. Esparza
S. Sickert, J. Krämer

Technische Universität München
ABGABE Montag, 3. Juli 2017, 10 Uhr, **25 Punkte**

Einführung in die theoretische Informatik

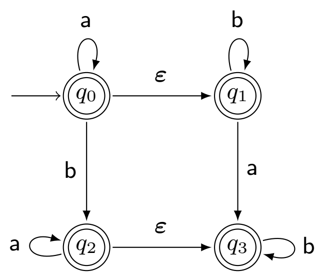
Sommersemester 2017 – Abgabeblatt 2

Abgabeblatt

Wir unterscheiden zwischen Übungs- und Abgabeblättern. Mit diesem *Abgabeblatt* fragen wir eine *Auswahl* an Kernaspekten ab, die auf den Trainingsblättern 4 bis 9 zu finden sind. Auch die Kernaspekte, die sich nicht auf Abgabeblättern finden, sind für die Klausur relevant. Wenn Sie die Abgabeblätter regelmäßig abgeben, können Sie zusätzlich einen *Bonus* von 0,3 (nach Bestehen) für beide Klausuren erreichen. Die genauen Regularien, wann Sie einen Notenbonus erhalten und wann nicht, finden Sie auf der Vorlesungshomepage.

AUFGABE 2.1. (*Myhill-Nerode-Äquivalenzklassen*) 4 Punkte

Gegeben sei folgender ε-NFA $\mathsf{N} = (\mathsf{Q}, \Sigma, \delta, \mathsf{q}_0, \mathsf{F})$ über dem Alphabet $\Sigma = \{\mathsf{a}, \mathsf{b}\}$:



Bestimmen Sie die Anzahl der Myhill-Nerode-Äquivalenzklassen der Sprache $\mathsf{L}(\mathsf{N})$, indem Sie den kanonischen Minimalautomaten (Definition 3.58) für $\mathsf{L}(\mathsf{N})$ konstruieren. Geben Sie zu jeder Klasse einen regulären Ausdruck an, der alle Wörter der Klasse beschreibt.

Hinweise:

- Sie erhalten 0 Punkte auf diese Aufgabe, falls Ihr konstruierter DFA nicht isomorph zu dem kanonischen Minimalautomaten für L ist. Prüfen Sie deshalb Ihre Lösung für verschiedene Wörter $w \in \Sigma^*$.
- Sie erhalten 0 Punkte auf diese Aufgabe, wenn Sie nicht die Tabelle, die Sie zur Minimierung Ihres DFAs verwendet haben, angegeben.
- Beschriften Sie die Zustände des Automaten mit dem lexikographisch kleinsten Element der Äquivalenzklasse.

AUFGABE 2.2. (*Kontextfreie Grammatiken*)

Die Definition aus den Folien für Typ-2 Grammatiken verbietet ε-Produktionen (mit der Ausnahme $\mathsf{S} \to \varepsilon$). In der Literatur werden aber auch Grammatiken $\mathsf{G} = (\mathsf{V}, \Sigma, \mathsf{P}, \mathsf{S})$ mit $\mathsf{P} \subseteq \mathsf{V} \times (\mathsf{V} \cup \Sigma)^*$ als kontextfrei bezeichnet. Zeigen Sie, dass diese Definitionen die gleiche Klasse an Sprachen beschreibt, d.h. $\mathsf{L}(\mathsf{G})$ ist kontextfrei für jedes G mit $\mathsf{P} \subseteq \mathsf{V} \times (\mathsf{V} \cup \Sigma)^*$. 1 Punkte

AUFGABE 2.3. (*Chomsky-Normalform, Cocke-Younger-Kasami-Algorithmus*) 8 Punkte

Wir betrachten folgende Grammatik $\mathsf{G} = (V, \Sigma, P, S)$ für (stark vereinfachte) arithmetische Ausdrücke mit $\Sigma = \{a, b, 0, 1, 2, +, -, \cdot, (,)\}$:

$$
\begin{aligned}
S &\to M + M \mid T + T \mid M \mid T \\
D &\to M - M \mid T - T \mid M \mid T \\
M &\to T \cdot M \mid M \\
T &\to F\,V \mid K \mid (S) \\
F &\to \varepsilon \mid K \\
K &\to 0 \mid 1 \mid 2 \\
V &\to a \mid b
\end{aligned}
$$

(a) Geben Sie eine Grammatik G' mit $\mathsf{L}(\mathsf{G}) = \mathsf{L}(\mathsf{G}')$, die nur noch nützliche Symbole enthält, an.

(b) Geben Sie eine Grammatik G'' mit $\mathsf{L}(\mathsf{G}') = \mathsf{L}(\mathsf{G}'')$, die in Chomsky-Normalform ist. Geben Sie hierbei alle vier Schritte an.

(c) Bestimmen Sie unter Verwendung des Cocke-Younger-Kasami-Algorithmus, ob $a + b + 2$ und $(2b)$ in $\mathsf{L}(\mathsf{G})$ enthalten sind. Geben Sie hierbei auch Ihre verwendeten Tabellen an.

Hinweise:

- Beachten Sie die aktualisierten Folien! Das Verfahren eliminiert nun erst am Ende Kettenproduktionen.

Definition (Hamming-Distanz)

Seien $u = a_1 a_2 \cdots a_n$ und $v = b_1 b_2 \cdots b_n$ zwei Wörter derselben Länge über dem Alphabet Σ. Die Hamming-Distanz $h(u, v)$ ist die Anzahl der Indizes $1 \leq i \leq n$ mit $a_i \neq b_i$. Die Hamming-Nachbarschaft ist dann definiert als $H(\mathsf{L}, i) = \{w \in \Sigma^* \mid \exists w' \in \mathsf{L}.\ |w| = |w'| \wedge h(w, w') \leq i\}$.
Beispiel: Sei $\Sigma = \{\mathsf{a}, \mathsf{b}\}$. Dann $h(\mathsf{aaa}, \mathsf{aaa}) = 0$, $h(\mathsf{aaa}, \mathsf{aba}) = 1$ und $h(\mathsf{aab}, \mathsf{baa}) = 2$.

AUFGABE 2.4.

5 Punkte

Zeigen Sie, dass die Hamming-Nachbarschaft mit $i = 1$ einer kontextfreien Sprachen wieder kontextfrei ist.

(a) Geben Sie hierzu ein allgemeines Verfahren an, dass eine Grammatik G in Chomsky-Normalform in eine Grammatik G' mit $\mathsf{L}(\mathsf{G}') = H(\mathsf{L}(\mathsf{G}), 1)$ übersetzt.

(b) Wenden Sie Ihr Verfahren auf die Grammatik $\mathsf{G} = (\{\mathsf{S}\}, \{(,)\}, \mathsf{P}, \mathsf{S})$ mit $\mathsf{P} : \mathsf{S} \to \mathsf{SS} \mid (\mathsf{S}) \mid ()$ an. Hierbei müssen Sie vorher G in Chomsky-Normalform umwandeln.

(c) Begründen Sie die Korrektheit Ihres allgemeinenen Verfahrens.

AUFGABE 2.5.

4 Punkte

Wir betrachten folgende (Pfeil-)Sprache über dem Alphabet $\Sigma = \{\uparrow, \downarrow, \leftarrow, \rightarrow\}$:

$$\mathsf{L} = \{w \in \Sigma^* \mid |w|_{\leftarrow} \neq |w|_{\rightarrow} \vee |w|_{\uparrow} \neq |w|_{\downarrow}\}$$

L beschreibt alle Pfade, die *nicht* im Ursprung enden. Zeigen Sie, dass L in der Klasse CFL, aber nicht in der Klasse DCFL liegt.

(a) Geben Sie hierzu einen PDA M mit $\mathsf{L} = \mathsf{L}_F(M)$ an.

(b) Zeigen Sie unter Verwendung der Abschlusseigenschaften von DCFLs, dass L nicht in der Klasse DCFL liegt.

Hinweise:

- Stellen Sie sicher, dass ihr PDA folgende Wörter korrekt akzeptiert bzw. ablehnt: $\varepsilon, \uparrow, \downarrow\uparrow\downarrow\uparrow, \downarrow\leftarrow\uparrow\rightarrow\uparrow$. Falls dies nicht der Fall ist, wird die Aufgabe mit 0 Punkten bewertet.
- Verwenden Sie maximal 8 Zustände und maximal 5 Stacksymbole für Ihren PDA. Falls Sie mehr verwenden, wird die Aufgabe mit 0 Punkten gewertet.
- Sie dürfen in dieser Aufgabe Resultate, die in den Übungen bewiesen worden sind, verwenden.

AUFGABE 2.6.

3 Punkte

Wir betrachten folgende Sprache über dem Alphabet $\Sigma = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$:

$$\mathsf{L} = \{\mathsf{a}^n \mathsf{b}^m \mathsf{c}^m \mid n, m \in \mathbb{N} \wedge m \text{ gerade}\} \cup \{\mathsf{a}^n \mathsf{b}^m \mathsf{c}^n \mid n, m \in \mathbb{N} \wedge m \text{ ungerade}\}$$

Zeigen Sie, dass L in der Klasse DCFL liegt, indem Sie einen DPDA für L angeben.

Hinweise:

- Falls Ihr PDA nicht der Definition eines DPDAs entspricht, wird die Aufgabe mit 0 Punkten gewertet.
- Stellen Sie sicher, dass ihr DPDA folgende Wörter korrekt akzeptiert bzw. ablehnt: $\varepsilon, \mathsf{abc}, \mathsf{cba}, \mathsf{abbccc}, \mathsf{aabcc}$. Falls dies nicht der Fall ist, wird die Aufgabe mit 0 Punkten bewertet.
- Verwenden Sie maximal 8 Zustände und maximal 4 Stacksymbole für Ihren DPDA. Falls Sie mehr verwenden, wird die Aufgabe mit 0 Punkten gewertet.