Institut für Informatik I
Prof. Dr. Heiko Röglin
Tobias Brunsch
Dennis den Brok

Logik und diskrete Strukturen
Wintersemester 2012/13
Abgabe: 13.11.2012, 14:00 Uhr

Übungsblatt 5

Aufgabe 5.1

2+2+2+2 Punkte

Gegeben seien zwei DFAs $M_1 = (Q_1, \Sigma, \delta_1, q_0^1, F_1)$ und $M_2 = (Q_2, \Sigma, \delta_2, q_0^2, F_2)$ über demselben Alphabet Σ.

(a) Konstruieren Sie einen DFA, der die Sprache $L(M_1) \cup L(M_2)$ entscheidet.

(b) Konstruieren Sie einen DFA, der die Sprache $L(M_1) \cap L(M_2)$ entscheidet.

(c) Konstruieren Sie einen DFA, der die Sprache $\overline{L(M_1)} := \Sigma^* \setminus L(M_1)$ entscheidet.

(d) Es sei L eine Sprache, die von einem DFA entschieden wird. Gibt es dann auch für jede Teilmenge $L' \subseteq L$ einen DFA, der L' entscheidet?

Aufgabe 5.2

3+2+3 Punkte

(a) Seien $k \in \mathbb{N}$ eine (feste) Zahl und $L \subseteq \{0,1\}^k$ eine Sprache über dem Alphabet $\{0,1\}$. Zeigen Sie, dass es einen DFA M gibt, der L entscheidet und höchstens einen akzeptierenden Zustand besitzt.

(b) Geben Sie eine feste Zahl $k \in \mathbb{N}$ und eine reguläre Sprache L an, die nur Wörter der Länge höchstens k enthält, sodass kein DFA mit höchstens einem akzeptierenden Zustand existiert, der L entscheidet.

(c) Zeigen Sie, dass Ihr Beispiel aus Aufgabenteil (b) die geforderten Bedingungen erfüllt.

Aufgabe 5.3

2+2+2+2

Wir wollen Wörter über dem Alphabet $\Sigma = \{a, b, c, d\}$ mit Hilfe von Wörtern über dem Alphabet $\hat{\Sigma} = \{0,1\}$ darstellen, indem wir jeden Buchstaben aus Σ einzeln kodieren. Die Standardvariante ist, a durch 00, b durch 01, c durch 10 und d durch 11 zu kodieren. Mit dieser Kodierung würden wir zum Beispiel das Wort $w = abba$ darstellen als $\hat{w} = 00010100$.

Zusätzlich betrachten wir die alternativen Kodierungen $C_1 \colon \Sigma \to \hat{\Sigma}^*$ und $C_2 \colon \Sigma \to \hat{\Sigma}^*$, gegeben durch

$$C_1(a) = 0\,, \quad C_1(b) = 11\,, \quad C_1(c) = 111\,, \quad C_1(d) = 10 \quad \text{und}$$
$$C_2(a) = 0\,, \quad C_2(b) = 110\,, \quad C_2(c) = 111\,, \quad C_2(d) = 10\,.$$

(a) Die Kodierung C_1 komprimiert jedes Wort mindestens genauso gut wie C_2. Dennoch ist Kodierung C_2 zu bevorzugen. Warum?

(b) Wir nehmen an, dass uns eine Statistik über die relativen Häufigkeiten der Buchstaben a, b, c und d in den darzustellenden Wörtern bekannt ist. Diese seien 0,45, 0,15, 0,10 und 0,30, das heißt der Buchstabe a kommt in diesen Wörtern dreimal so oft vor wie der Buchstabe b.

Welchen Kompressionsfaktor erhält man bei solchen Wörtern, wenn man die Kodierung C_2 verwendet, verglichen mit der Standardkodierung?

(c) Wir interessieren uns dafür, ob ein Wort aus $\hat{\Sigma}^*$ die Darstellung eines Wortes aus Σ^* bezüglich C_2 ist. Sei dazu L die Sprache dieser Darstellungen, das heißt $L = \{C_2(a_1) \ldots C_2(a_n) : n \in \mathbb{N}_0 \text{ und } a_1, \ldots, a_n \in \Sigma\}$. Geben Sie einen DFA an, der L entscheidet.

(d) Geben Sie die Sprache L aus Aufgabenteil (c) möglichst einfach an, ohne die Kodierung C_2 zu verwenden.

Aufgabe 5.4

2+4 Zusatzpunkte

Wir betrachten die Sprache $L = \{ww :\ w \in \Sigma^*\}$ über dem Alphabet Σ.

(a) Für welche Alphabete Σ wird die Sprache L von einem DFA entschieden? Geben Sie für diese Alphabete Σ einen DFA an, der die Sprache L entscheidet.

(b) Beweisen Sie für die anderen Alphabete Σ, ohne das Pumping-Lemma zu benutzen, dass die Sprache L von keinem DFA entschieden wird.