Einführung in die Computerlinguistik
Zwischenklausur

21.11.2011

Laura Kallmeyer

WS 2011/2012, Heinrich-Heine-Universität Düsseldorf

Erlaubte Hilfsmittel: Eine Din-A4 Seite mit Notizen.

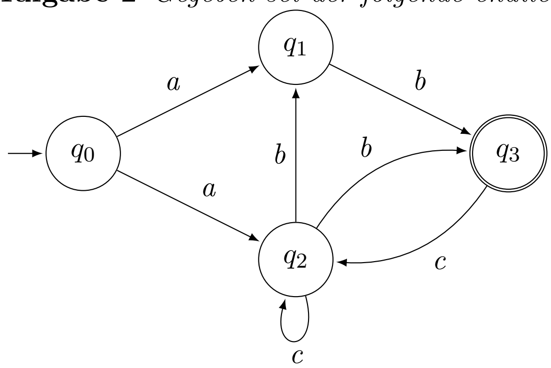
Aufgabe 1 *Sei* $\Sigma_1 = \{a, b, c, d\}$*,* $\Sigma_2 = \{e, f\}$*,* $L_1 = \{a^n b^m c^n \mid n, m \geq 0\}$ *und* $L_2 = \{\varepsilon, e\}$*. Welche der folgenden Aussagen ist wahr?*

(a) $\varepsilon \in L_1$ *(b)* $L_1 \in \Sigma_1^*$ *(c)* $L_1 \cap L_2 = \emptyset$ *(d)* $L_2 \subset \Sigma_2^*$

(e) $|L_1 \cap L_2| = 1$ *(f)* $\Sigma_1^* \cap \Sigma_2^* = \emptyset$ *(g)* $L_2^* \subset \Sigma_2^*$ *(h)* $\{d\}^* \subset (\Sigma_1^* \setminus L_1)$

Lösung: Wahr sind (a), (d), (e), (g) 8

Aufgabe 2 *Gegeben sei der folgende endliche Automat:*



(a) Geben Sie den Automaten in Tupelschreibweise an. D.h., als Tupel $\langle Q, \Sigma, \delta, q_0, F\rangle$*.*

(b) Geben Sie einen äquivalenten deterministischen endlichen Automaten an.

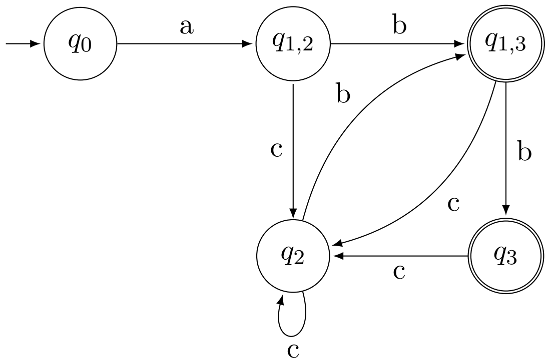
Lösung:

(a) $\langle Q, \Sigma, \delta, q_0, F\rangle$ mit $Q = \{q_0, q_1, q_2, q_3\}$, $\Sigma = \{a, b, c\}$, $F = \{q_3\}$ und δ wie folgt definiert:

$$\delta(q_0, a) = \{q_1, q_2\} \quad \delta(q_1, b) = \{q_3\} \quad \delta(q_2, b) = \{q_1, q_3\} \quad \delta(q_2, c) = \{q_2\} \quad \delta(q_3, c) = \{q_2\}$$

Alle anderen Werte von δ sind $\emptyset$. 4

(b) DFA:



6

Aufgabe 3 *Geben Sie die Denotate der folgenden regulären Ausdrücke an, indem Sie die jeweilige Menge, falls sie endlich ist, explizit auflisten, und sonst in Worten beschreiben.*

(a) $a|b(c|b|\varepsilon)$ *(b)* $a(b|\varepsilon)b\emptyset$ *(c)* $\emptyset|b(\varepsilon|b|aaa)$ *(d)* $(\varepsilon\emptyset|b)^+(\varepsilon|b)$ *(e)* $(a|b)^*(\emptyset|b)$

Lösung:

(a) $L(a|b(c|b|\varepsilon)) = \{a, bc, bb, b\}$ 2

(b) $L(a(b|\varepsilon)b\emptyset) = \emptyset$ 1

(c) $L(\emptyset|b(\varepsilon|b|aaa)) = \{b, bb, baaa\}$ 2

(d) $L((\varepsilon\emptyset|b)^+(\varepsilon|b)) = \{b^n \,|\, n \geq 1\}$, also die Menge aller Folgen von mindestens einem und beliebig vielen bs. 2

(e) $L((a|b)^*(\emptyset|b))$ ist die Menge aller Folgen von as und bs, die eine Länge von mindestens 1 haben und mit einem b enden. 2

Aufgabe 4 *Gegeben seien die folgenden Grammatiken:*

(a) $G_1 = \langle\{S, A\}, \{a, b, c\}, S, P\rangle$ *mit den folgenden Produktionen in* P*:*

$$S \to baS \quad S \to A \quad A \to c$$

(b) $G_2 = \langle\{S, A\}, \{a, b, c\}, S, P\rangle$ *mit den folgenden Produktionen in* P*:*

$$S \to bSA \quad S \to aa \quad A \to c$$

(c) $G_3 = \langle\{S, A\}, \{a, b\}, S, P\rangle$ *mit den folgenden Produktionen in* P*:*

$$S \to Ab \quad S \to ab \quad A \to aS$$

(d) $G_4 = \langle\{S, A\}, \{a, b, c\}, S, P\rangle$ *mit den folgenden Produktionen in* P*:*

$$S \to Abc \quad S \to abc \quad A \to Sa$$

1. *Geben Sie für jede der Grammatiken an, ob es sich um eine Typ 3 Grammatik handelt oder um eine Typ 2 Grammatik, die keine Typ 3 Grammatik ist. Begründen Sie Ihre Antwort.*
2. *Geben Sie für alle Grammatiken in der obigen Liste an, welche Kettensprache (*string language*) die jeweilige Grammatik generiert. Beschreiben Sie die Sprache in impliziter Mengendarstellung, also in der Form* $\{\ldots \mid \ldots\}$.

Lösung:

(a) G_1 ist eine Typ 3 Grammatik, da alle linken Seiten von Produktionen einzelne Nichtterminale sind und da es in jeder rechten Seite nur höchstens ein Nichtterminales gibt, das immer ganz rechts steht. 1

Sprache $\{(ba)^n c \mid n \geq 0\}$ 2

(b) G_2 ist keine Typ 3 Grammatik, da die erste Produktion 2 Nichtterminale in ihrer rechten Seite enthält. Sie ist aber eine Typ 2 Grammatik, da alle linken Seiten einzelne Nichtterminale sind. 1

Sprache $\{b^n aac^n \mid n \geq 0\}$ 2

(c) G_3 ist keine Typ 3 Grammatik, da die Nichttermnalen in den rechten Seiten von Produktionen mal ein echtes Präfix, mal ein echtes Suffix sind. Sie ist aber eine Typ 2 Grammatik, da alle linken Seiten einzelne Nichtterminale sind. 1

Sprache $\{a^n b^n \mid n > 0\}$ 2

(d) G_4 ist eine Typ 3 Grammatik, da alle linken Seiten von Produktionen einzelne Nichtterminale sind und da es in jeder rechten Seite nur höchstens ein Nichtterminales gibt, das immer ganz links steht. 1

Sprache $\{(abc)^n \mid n > 0\}$ 2

Aufgabe 5 *Geben Sie kontextfreie Grammatiken für folgende Sprachen an:*

(a) $\{a^n b^m c^n \mid n \geq 0, m \geq 1\}$

(b) $\{a^n b^m c^m d^n \mid n, m \geq 1\}$

(c) $\{ww^R \mid w \in \{a, bb\}^*\}$

Lösung:

(a) $\langle \{S, A\}, \{a, b, c\}, P, S \rangle$ mit folgenden Produktionen in P:

$$S \to aSc \quad S \to A \quad A \to bA \quad A \to b$$

3

(b) $\langle \{S\}, \{a, b, c, d\}, P, S \rangle$ mit folgenden Produktionen in P:

$$S \to aSd \quad S \to aAd \quad A \to bAc \quad A \to bc$$

4

(c) $\langle \{S, A\}, \{a, b\}, P, S \rangle$ mit folgenden Produktionen in P:

$$S \to aSa \quad S \to bbSbb \quad S \to \varepsilon$$

4