

Einführung in die Computerlinguistik

Hausaufgabe 2, Abgabe 30.04.2018

Laura Kallmeyer

SoSe 2018, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 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.

1. $b|\emptyset|\varepsilon$ 2. $a(c|\emptyset)bb$ 3. $(a|bc)(\varepsilon|b\varepsilon)$ 4. $(a\emptyset b)^+|(c\varepsilon c)^+|(dd)^*$ 5. $(bc|c\emptyset)^+d|d$ 6. $(a(bb)^+|b|c)^*$

Lösung:

1. $\{b, \varepsilon\}$
2. $\{acbb\}$
3. $\{a, bc, ab, bcb\}$
4. $\{(cc)^n \mid n > 0\} \cup \{(dd)^n \mid n \geq 0\}$
5. $\{(bc)^n d \mid n \geq 0\}$
6. $\{w \in \{a, b, c\}^* \mid \text{jedes } a \text{ in } w \text{ ist notwendig von (mindestens) zwei } bs \text{ gefolgt}\}$

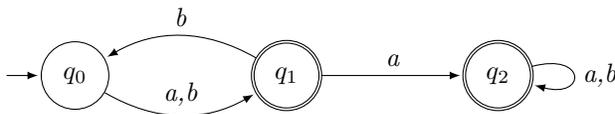
Aufgabe 2 Geben Sie für die folgenden Mengen jeweils einen regulären Ausdruck an, der die jeweilige Menge denotiert.

1. $\{w \mid w \in \{a, b, c\}^* \text{ und } w \text{ enthält mindestens einmal die Folge } ab\}$
2. $\{w \mid w \in \{a, b, c\}^* \text{ und } w \text{ hat eine gerade Anzahl von } as \text{ (0 ist auch erlaubt)}\}$
3. $\{w \mid w \in \{a, b, c\}^+, \text{ alle } as \text{ in } w \text{ sind unmittelbar von } bc \text{ oder } cb \text{ gefolgt}\}$

Lösung:

1. $(a|b|c)^*ab(a|b|c)^*$ 2. $(b|c)^*(a(b|c)^*a(b|c)^*)^*$ 3. $(b|c|abc|acb)^+$

Aufgabe 3 Gegeben sei der folgende DFA:



Berechnen Sie einen regulären Ausdruck, der die Sprache beschreibt, die von dem Automaten akzeptiert wird, indem Sie den Algorithmus aus der Vorlesung anwenden.

Hinweis: Wenn es keinen Pfad gibt, lautet der entsprechende reguläre Ausdruck \emptyset .

Lösung:

Wir haben zwei Endzustände, der Ausdruck, den wir benötigen, ist also $r_{0,1}^2|r_{0,2}^2$

- $r_{0,1}^2 = r_{0,1}^1|r_{0,2}^1(r_{22}^1)^*r_{21}^1$
- $r_{0,1}^1 = (a|b)(b(a|b))^*$
- $r_{0,2}^1 = r_{0,2}^0|r_{0,1}^0(r_{11}^0)^*r_{12}^0 = \emptyset|(a|b)(b(a|b))^*a = (a|b)(b(a|b))^*a$
- $r_{2,2}^1 = (a|b)$
- $r_{2,1}^1 = \emptyset$
- $r_{0,1}^2 = (a|b)(b(a|b))^*|(a|b)(b(a|b))^*a(a|b)^*\emptyset = (a|b)(b(a|b))^*$

- $r_{0,2}^2 = r_{0,2}^1 | r_{0,2}^1 (r_{22}^1)^* r_{22}^1$
 $r_{0,2}^1 = (a|b)(b(a|b))^* a$ (s.o.)
 $r_{22}^1 = (a|b)$ (s.o.)
 $r_{0,2}^2 = (a|b)(b(a|b))^* a | (a|b)(b(a|b))^* a (a|b)^* (a|b) = (a|b)(b(a|b))^* a | (a|b)(b(a|b))^* a (a|b)^+$
 $= (a|b)(b(a|b))^* a (a|b)^*$

Regulärer Ausdruck:

$$r_{0,1}^2 | r_{0,2}^2 = (a|b)(b(a|b))^* | (a|b)(b(a|b))^* a (a|b)^* = (a|b)(b(a|b))^* (\varepsilon | a (a|b)^*)$$

Aufgabe 4 Geben Sie für die folgenden Mengen jeweils eine reguläre Grammatik an, die diese Menge generiert.

1. $\{w \mid w \in \{a, b\}^* \text{ und } w \text{ enthält eine gerade Anzahl Zeichen (0 ist auch erlaubt)}\}$
2. $L(a^*(bc)^+)$

Lösung:

1. $G = \langle N, T, P, S \rangle$ mit $N = \{A, S\}$, $T = \{a, b\}$ und
 $P = \{S \rightarrow aA, S \rightarrow bA, S \rightarrow \varepsilon, A \rightarrow aS, A \rightarrow bS\}$
2. $G = \langle N, T, P, S \rangle$ mit $N = \{A, B, S\}$, $T = \{a, b\}$ und
 $P = \{S \rightarrow aS, S \rightarrow bcA, A \rightarrow bcA, A \rightarrow \varepsilon\}$