

Cvičení 2

Příklad 1: Napište regulární výrazy pro následující jazyky:

- Jazyk $\{ab, ba, abb, bab, abbb, babb\}$
- Jazyk nad abecedou $\{a, b, c\}$ obsahující právě ta slova, která obsahují podslovo abb .
- Jazyk nad abecedou $\{a, b, c\}$ obsahující právě ta slova, která začínají prefixem bca nebo končí sufixem $ccab$.
- Jazyk $\{w \in \{0, 1\}^* \mid |w|_0 \bmod 2 = 0\}$.
- Jazyk $\{w \in \{0, 1\}^* \mid |w|_0 \bmod 3 = 1\}$.
- Jazyk $\{w \in \{0, 1\}^* \mid w \text{ obsahuje podslova } 010 \text{ a } 111\}$
- Jazyk $\{w \in \{a, b\}^* \mid w \text{ obsahuje podslovo } bab \text{ nebo } |w|_b \leq 3\}$
- Jazyk $\{w \in \{a, b\}^* \mid w \text{ obsahuje podslovo } bab \text{ a } |w|_b \leq 3\}$
- Jazyk všech slov nad abecedou $\{a, b, c\}$, ve kterých se nikde nevyskytují dva znaky a hned za sebou.

Příklad 2: Mějme dva jazyky L_1 a L_2 popsané regulárními výrazy

$$L_1 = \mathcal{L}(0^*1^*0^*1^*0^*), \quad L_2 = \mathcal{L}((01 + 10)^*).$$

- Jaké je nejkratší a nejdelší slovo v průniku $L_1 \cap L_2$?
- Proč žádný z těchto jazyků L_1 a L_2 není podmnožinou toho druhého?
- Jaké je nejkratší slovo, které nepatří do sjednocení $L_1 \cup L_2$? Je to jednoznačné?

Příklad 3: Řekněme, že bychom chtěli navrhnout syntaxi pro zápis jednoduchých aritmetických výrazů pomocí slov nad abecedou

$$\Sigma = \{A, B, \dots, Z, a, b, \dots, z, 0, 1, \dots, 9, ., +, -, *, /, (,)\}.$$

- Navrhněte, jak budou vypadat identifikátory, a popište to pomocí regulárního výrazu.
- Navrhněte, jak budou vypadat číselné konstanty, a popište to pomocí regulárního výrazu.

Poznámka: Při popisu číselných konstant umožněte jak celočíselné konstanty, např. 129 nebo 0, tak neceločíselné konstanty, např. 3.14, $-1e10$ nebo $4.2E-23$. Zvažte i možnost zápisu číselných konstant v dalších číselných soustavách kromě desítkové (např. hexadecimální, oktálové, binární).

Příklad 4: Pro každý z následujících jazyků sestrojte deterministický konečný automat (DKA), který ho rozpoznává. Vytvořené automaty znázorněte grafem a zapište tabulkou.

- $L_1 = \{w \in \{a, b\}^* \mid w = a\}$

- b) $L_2 = \{b, ab\}$
- c) $L_3 = \{w \in \{a, b\}^* \mid \exists n \in \mathbb{N} : w = a^n\}$
- d) $L_4 = \{w \in \{a, b, c\}^* \mid |w|_a \geq 1\}$
- e) $L_5 = \{w \in \{0, 1\}^* \mid w \text{ obsahuje podslovo } 011\}$
- f) $L_6 = \{w \in \{a, b, c\}^* \mid |w| > 0 \wedge |w|_a = 0\}$
- g) $L_7 = \{w \in \{a, b\}^* \mid |w| \geq 2 \text{ a poslední dva symboly slova } w \text{ nejsou stejné}\}$
- h) $L_8 = \{w \in \{a, b\}^* \mid |w|_a \bmod 3 = 1\}$

Příklad 5: Sestrojte DKA přijímající slova začínající $abaab$, končící $abaab$ a obsahující $abaab$, tj. sestrojte DKA rozpoznávající následující tři jazyky:

- a) $L_1 = \{abaabw \mid w \in \{a, b\}^*\}$
- b) $L_2 = \{wabaab \mid w \in \{a, b\}^*\}$
- c) $L_3 = \{w_1abaabw_2 \mid w_1, w_2 \in \{a, b\}^*\}$

Příklad 6: Navrhněte obecný postup, jak pro daný DKA $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ zjistit, zda:

- a) $\mathcal{L}(\mathcal{A}) = \emptyset$
- b) $\mathcal{L}(\mathcal{A}) = \Sigma^*$

Příklad 7: Navrhněte DKA $\mathcal{A}_1, \mathcal{A}_2$ takové, že:

$$\mathcal{L}(\mathcal{A}_1) = \{w \in \{a, b\}^* \mid |w|_a \bmod 2 = 0\}$$

$$\mathcal{L}(\mathcal{A}_2) = \{w \in \{a, b\}^* \mid \text{ve } w \text{ je každý výskyt symbolu } b \text{ následován symbolem } a\}$$

S využitím automatů $\mathcal{A}_1, \mathcal{A}_2$ sestrojte DKA rozpoznávající následující jazyky:

- a) $L_1 = \{w \in \{a, b\}^* \mid |w|_a \bmod 2 = 0 \text{ a ve } w \text{ je každý výskyt symbolu } b \text{ následován symbolem } a\}$
- b) $L_2 = \{w \in \{a, b\}^* \mid |w|_a \bmod 2 = 0 \text{ nebo je ve } w \text{ každý výskyt symbolu } b \text{ následován symbolem } a\}$
- c) $L_3 = \{w \in \{a, b\}^* \mid \text{ve } w \text{ není nějaký výskyt symbolu } b \text{ následován symbolem } a\}$
- d) $L_4 = \{w \in \{a, b\}^* \mid |w|_a \bmod 2 = 0 \text{ a ve } w \text{ není nějaký výskyt symbolu } b \text{ následován symbolem } a\}$
- e) $L_5 = \{w \in \{a, b\}^* \mid \text{jestliže } |w|_a \bmod 2 = 0, \text{ pak je ve } w \text{ každý výskyt symbolu } b \text{ následován symbolem } a\}$

- f) $L_6 = \{w \in \{a, b\}^* \mid |w|_a \bmod 2 = 0 \text{ právě, když je ve } w \text{ každý výskyt symbolu } b \text{ následován symbolem } a\}$

Příklad 8: Pro každý z následujících jazyků sestrojte DKA, který ho rozpoznává. Vytvořené automaty znázorněte grafem a zapište tabulkou.

- a) $L_1 = \{w \in \{a, b\}^* \mid |w| \geq 4 \text{ a druhý, třetí a čtvrtý symbol slova } w \text{ jsou stejné}\}$
b) $L_2 = \{w \in \{a, b\}^* \mid |w| \geq 4 \text{ a třetí a poslední symbol slova } w \text{ jsou stejné}\}$
c) $L_3 = \{w \in \{a, b, c, d\}^* \mid w \text{ nezačíná } a, \text{ druhý znak není } b, \text{ třetí znak není } c \text{ a čtvrtý znak není } d\}$

Poznámka: Tento jazyk zahrnuje i ta slova w , kde $|w| < 4$.

- d) $L_4 = \{w \in \{a, b, c, d\}^* \mid w \text{ nezačíná } a \text{ nebo druhý znak není } b \text{ nebo třetí znak není } c \text{ nebo čtvrtý znak není } d\}$

Příklad 9: Navrhněte obecný postup, jak pro dané DKA $\mathcal{A}_1 = (Q_1, \Sigma, \delta_1, q_1, F_1)$ a $\mathcal{A}_2 = (Q_2, \Sigma, \delta_2, q_2, F_2)$ zjistit, zda $\mathcal{L}(\mathcal{A}_1) = \mathcal{L}(\mathcal{A}_2)$.