Úkol č. 1 do předmětu Teoretická informatika

Vojtěch Havlena (xhavle03)

1. Příklad

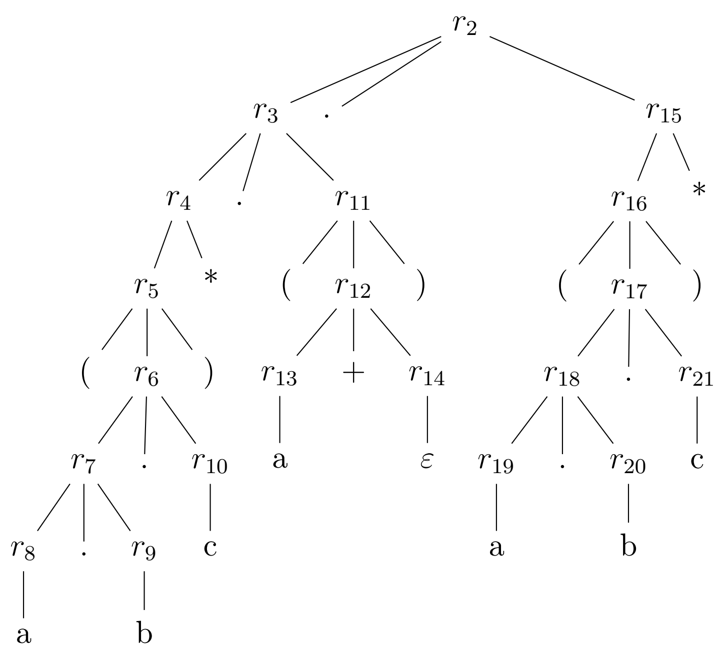
(a) Gramatika G_1 je čtveřice $G_1 = (N, \Sigma, P, S)$, kde

- $N = \{S, A, B, D\}$ je konečná množina neterminálních symbolů,
- $\Sigma = \{a, b, c, d\}$ je konečná množina terminálních symbolů,
- $P = \{S \to AD, A \to aAc, A \to B, B \to bB, B \to \varepsilon, D \to dD, D \to \varepsilon\}$ je konečná množina přepisovacích pravidel a
- $S \in N$ je výchozí symbol gramatiky G_1.

(b) Gramatika G_1 je typu 2 (není typu 3, protože obsahuje pravidlo se dvěma neterminálními symboly na pravé straně). Jazyk L_1 je jazyk typu 2 (bezkontextový jazyk). Jazyk L_1 není jazyk typu 3, protože neexistuje konečný automat, který by jej přijímal. Tyto typy se mohou obecně lišit, protože gramatika daného typu T může generovat jazyky typu T a vyšších (za jazyk nejvyššího typu považuji jazyk $\mathcal{L}_3$). Např. gramatika typu 0 může generovat jazyky typu 1. Toto plyne z faktu, že $\mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \mathcal{L}_2 \subseteq \mathcal{L}_3$.

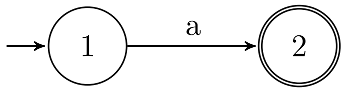
2. Příklad Část (a):

(a) Rozklad regulárního výrazu r_2 si vyjádříme stromem.

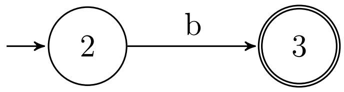


(b) Převod stromu reprezentujícího regulární výraz r_2 na konečný automat.

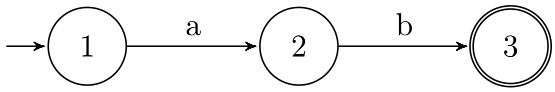
i. Regulárnímu výrazu $r_8 = a$ odpovídá automat N_1:



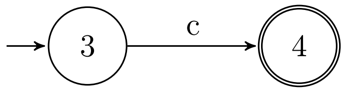
ii. Regulárnímu výrazu $r_9 = b$ odpovídá automat N_2:



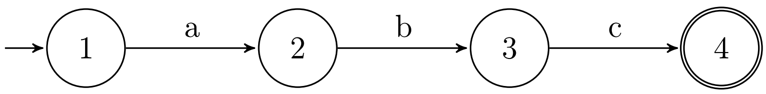
iii. Regulárnímu výrazu $r_7 = r_8 r_9 = ab$ odpovídá automat N_3:



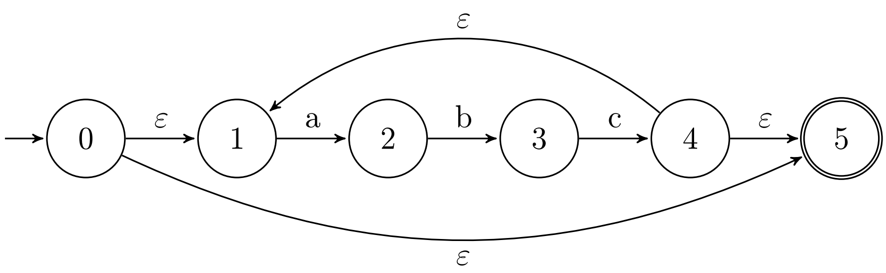
iv. Regulárnímu výrazu $r_{10} = c$ odpovídá automat N_4:



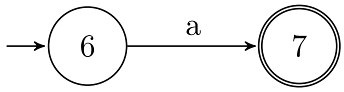
v. Regulárnímu výrazu $r_6 = r_7 r_{10} = abc$ odpovídá automat N_5:



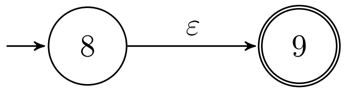
vi. Automat pro regulární výraz $r_5 = (r_6)$ je stejný jako automat N_5. Zkonstruujeme tedy rovnou automat N_6 pro výraz $r_4 = r_5^* = (abc)^*$.



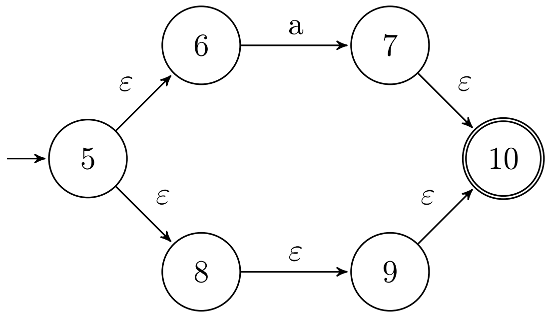
vii. Regulárnímu výrazu $r_{13} = a$ odpovídá automat N_7:



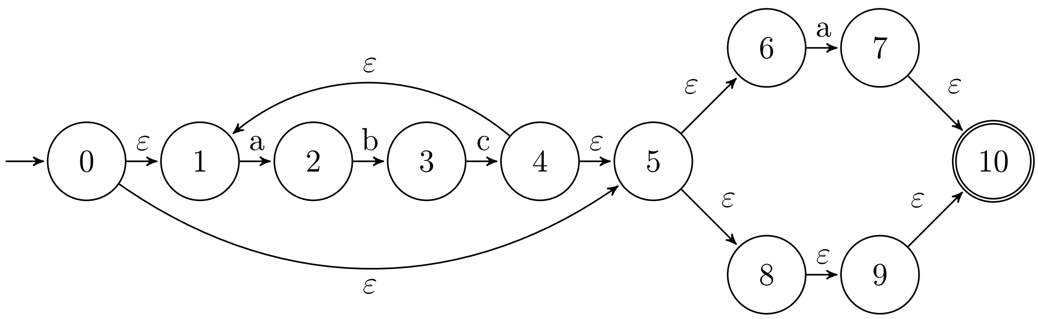
viii. Regulárnímu výrazu $r_{14} = \varepsilon$ odpovídá automat N_8:



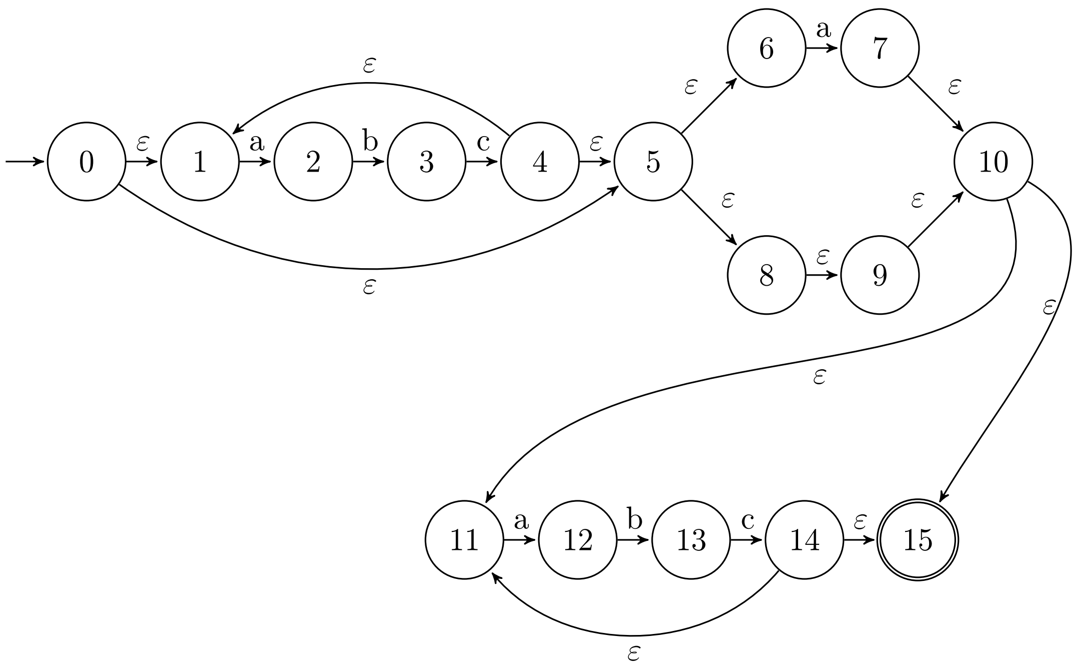
ix. Regulárnímu výrazu $r_{12} = r_{13} + r_{14} = a + \varepsilon$ odpovídá automat N_9:



x. Automat pro regulární výraz $r_{11} = (r_{12})$ je stejný jako automat N_9. Zkonstruujeme tedy rovnou automat N_{10} pro výraz $r_4 = r_4 r_{11} = (abc)^*(a+\varepsilon)$.



xi. Automat pro regulární výraz $r_{15} = (abc)^*$ je ekvivalentní automatu N_6. Uděláme tedy rovnou automat N_{11} pro výraz $r_2 = r_3 r_{15}$.

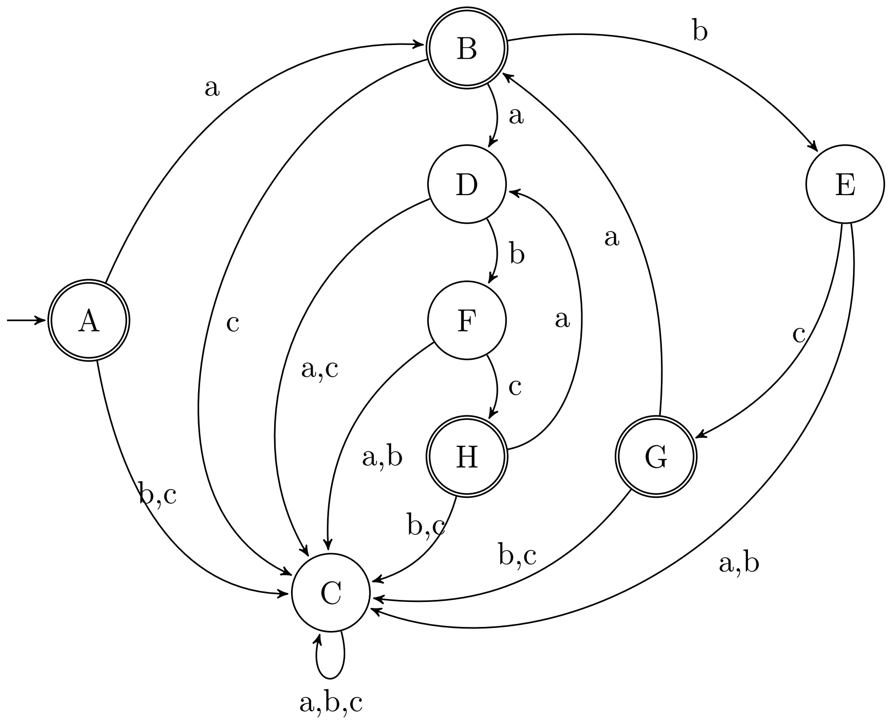


(c) Převod RKA N_{11} na DKA M_1 pomocí algoritmu 3.6 z opory k předmětu TIN. Přechodová funkce DKA M_1 je označena jako δ'.

- Počáteční stav A je $A = \varepsilon$-uzávěr$(\{0\}) = \{0, 1, 5, 6, 8, 9, 10, 11, 15\}$
- $\delta'(A, a) = \varepsilon$-uzávěr$(\{2, 7, 12\}) = \{2, 7, 10, 11, 12, 15\} = B$
- $\delta'(A, b) = \varepsilon$-uzávěr$(\emptyset) = C$
- $\delta'(A, c) = \varepsilon$-uzávěr$(\emptyset) = C$
- $\delta'(B, a) = \varepsilon$-uzávěr$(\{12\}) = \{12\} = D$
- $\delta'(B, b) = \varepsilon$-uzávěr$(\{3, 13\}) = \{3, 13\} = E$
- $\delta'(B, c) = \varepsilon$-uzávěr$(\emptyset) = C$
- $\delta'(C, a) = \varepsilon$-uzávěr$(\emptyset) = C$

- $\delta'(C, b) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(C, c) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(D, a) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(D, b) = \varepsilon\text{-uzávěr}(\{13\}) = \{13\} = F$
- $\delta'(D, c) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(E, a) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(E, b) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(E, c) = \varepsilon\text{-uzávěr}(\{4, 14\}) = \{1, 4, 5, 6, 8, 9, 10, 11, 14, 15\} = G$
- $\delta'(F, a) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(F, b) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(F, c) = \varepsilon\text{-uzávěr}(\{14\}) = \{11, 14, 15\} = H$
- $\delta'(G, a) = \varepsilon\text{-uzávěr}(\{2, 7, 12\}) = \{2, 7, 10, 11, 12, 15\} = B$
- $\delta'(G, b) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(G, c) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(H, a) = \varepsilon\text{-uzávěr}(\{12\}) = \{12\} = D$
- $\delta'(H, b) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- $\delta'(H, c) = \varepsilon\text{-uzávěr}(\emptyset) = C$
- Množina koncových stavů $F = \{A, B, G, H\}$

Grafické znozornění DKA M_1:



(d) Převod DKA M_1 na odpovídající redukovaný DKA M_2. Pro převod je použit algoritmus 3.5 z opory k předmětu TIN.

- Automat M_1 neobsahuje žádné nedosažitelné stavy.

– Automat M_1 je již úplný.

– Můžeme tedy provést přímo převod na redukovaný KA.

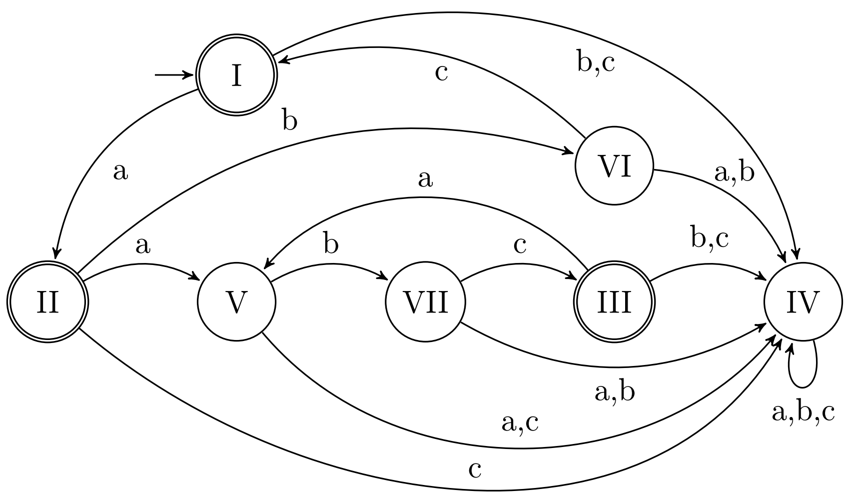
	$\overset{0}{\equiv}$	a	b	c
⟦	I: A	B(I)	C(II)	C(II)
	B	D(II)	E(II)	C(II)
	G	B(I)	C(II)	C(II)
	H	D(I)	C(II)	C(II)
	II: C	C(II)	C(II)	C(II)
	E	C(II)	C(II)	G(I)
	D	C(II)	F(II)	C(II)
	F	C(II)	C(II)	H(I)

	$\overset{1}{\equiv}$	a	b	c
⟦	I: A	B(II)	C(III)	C(III)
	G	B(II)	C(III)	C(III)
⟦	II: B	D(III)	E(IV)	C(III)
	H	D(III)	C(III)	C(III)
	III: C	C(III)	C(III)	C(III)
	D	C(III)	F(IV)	C(III)
	IV: E	C(III)	C(III)	G(I)
	F	C(III)	C(III)	H(II)

	$\overset{2}{\equiv}$	a	b	c
⟦	I: A	B(II)	C(IV)	C(IV)
	G	B(II)	C(IV)	C(IV)
⟦	II: B	D(V)	E(VI)	C(IV)
⟦	III: H	D(V)	C(IV)	C(IV)
	IV: C	C(IV)	C(IV)	C(IV)
	V: D	C(IV)	F(VII)	C(IV)
	VI: E	C(IV)	C(IV)	G(I)
	VII: F	C(IV)	C(IV)	H(III)

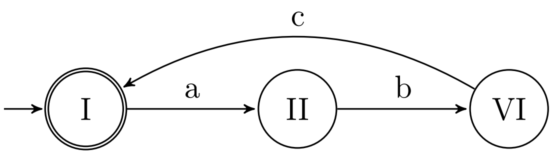
– Počáteční stav automatu: I.

Grafické znozornění výsledného redukovaného DKA M_2:

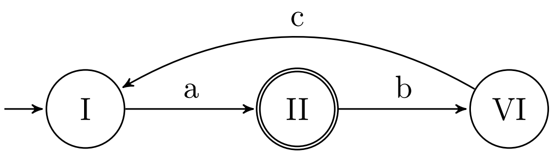


Část (b): Počet tříd ekvivalence relace $\sim_L$ pro jazyk $L(M_2)$ odpovídá počtu stavů redukovaného DKA M_2. Celkem je tedy 7 tříd ekvivalence.

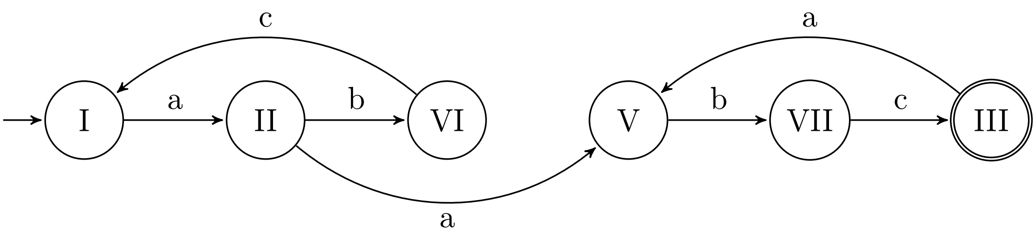
(a) Ekvivalenční třída $L^{-1}(I)$:



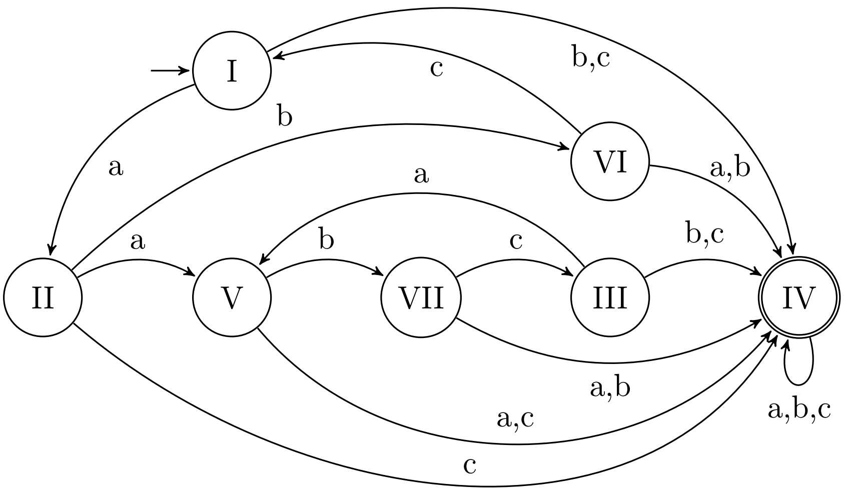
(b) Ekvivalenční třída $L^{-1}(II)$:



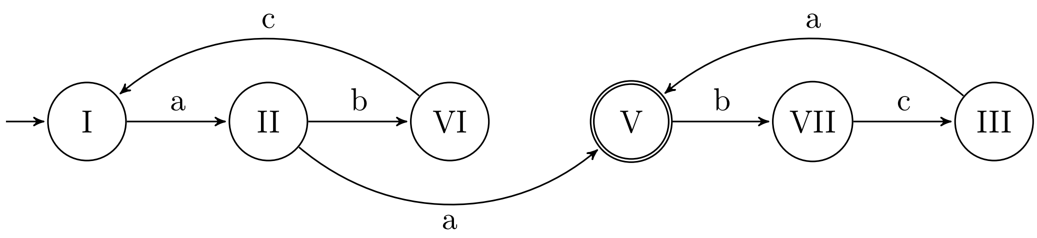
(c) Ekvivalenční třída $L^{-1}(III)$:



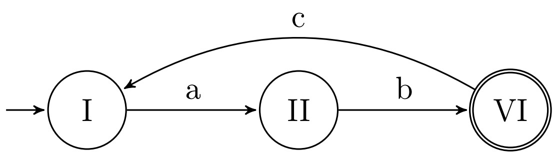
(d) Ekvivalenční třída $L^{-1}(IV)$:



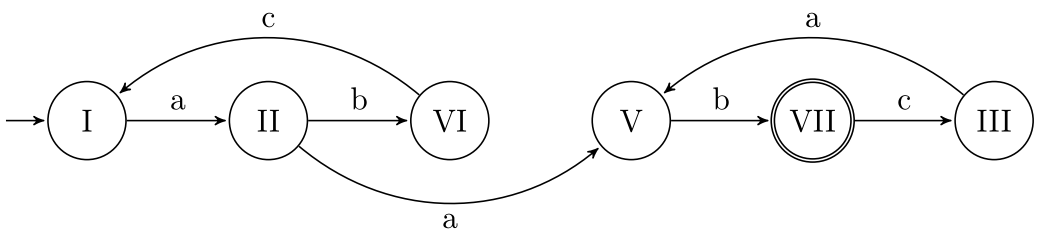
(e) Ekvivalenční třída $L^{-1}(V)$:



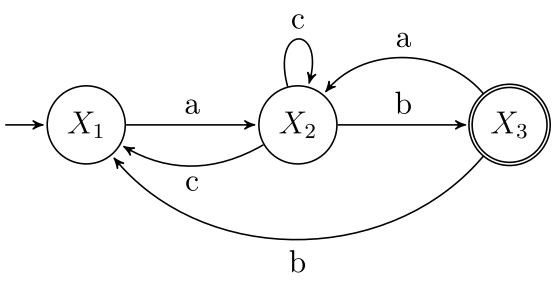
(f) Ekvivalenční třída $L^{-1}(VI)$:



(g) Ekvivalenční třída $L^{-1}(VII)$:



3. Příklad



Soustava rovnic pro konečný automat M_3:

$$X_1 = aX_2 \tag{1}$$
$$X_2 = cX_1 + cX_2 + bX_3 \tag{2}$$
$$X_3 = \varepsilon + bX_1 + aX_2 \tag{3}$$

Řešení soustavy rovnic: Výraz pro X_1 dosadíme z (1) do (2) a z (1) do (3) a dostaneme soustavu

$$X_2 = caX_2 + cX_2 + bX_3 \tag{4}$$
$$X_3 = \varepsilon + baX_2 + aX_2 \tag{5}$$

Výraz X_3 dosadíme z (5) do (4) a dostaneme rovnici

$$X_2 = caX_2 + cX_2 + b(\varepsilon + baX_2 + aX_2) \tag{6}$$

Vzhledem k tomu, že operace $\cdot$ a $+$ jsou distributivní, z rovnice (6) dostáváme

$$X_2 = caX_2 + cX_2 + b\varepsilon + bbaX_2 + baX_2 \tag{7}$$

S využitím identity ε a po částečném vytknutí X_2 získáme tvar

$$X_2 = (ca + c + bba + ba)X_2 + b \tag{8}$$

Podle věty 3.14 (opora k předmětu TIN) je řešení rovnice (8) dáno následujícím předpisem

$$X_2 = (ca + c + bba + ba)^* b \tag{9}$$

Abychom získali regulární výraz, který je ekvivalentní automatu M_3, musíme vyjádřit výraz X_1.

$$X_1 = aX_2 = a(ca + c + bba + ba)^* b \tag{10}$$

Ekvivalentní regulární výraz k automatu M_3 je tedy $a(ca + c + bba + ba)^* b$.

4. Příklad

- Vstup algoritmu: Dva konečné automaty $M_1 = (Q_1, \Sigma_1, \delta_1, q_1^0, F_1)$ a $M_2 = (Q_2, \Sigma_2, \delta_2, q_2^0, F_2)$.
- Výstup: Konečný automat $M_{restrict}$ takový, že $L(M_{restrict}) = restrict(L(M_1), L(M_2))$, při čemž operace $restrict$ je dána následovně

$$restrict(L_1, L_2) = \{w | w \in L_1 \land \exists w' \in L_2 : |w| = |w'|\}.$$

- Metoda:
  KA $M_{restrict} = (Q_3, \Sigma_3, \delta_3, q_3^0, F_3)$ je zkonstruován následovně:
  - Množina stavů $Q_3 = Q_1 \times Q_2$
  - Vstupní abeceda $\Sigma_3 = \Sigma_1$
  - Přechodová funkce $\delta_3 : Q_3 \times \Sigma_3 \rightarrow 2^{Q_3}$ je dána následovně:
    $\forall a \in \Sigma_1$, $\forall q_1^1, q_1^2 \in Q_1$ a $\forall q_2^1, q_2^2 \in Q_2$:

$$(q_1^2, q_2^2) \in \delta_3((q_1^1, q_2^1), a) \Leftrightarrow q_1^2 \in \delta_1(q_1^1, a) \land \exists b \in \Sigma_2 : q_2^2 \in \delta_2(q_2^1, b)$$

  - Počáteční stav $q_3^0 = (q_1^0, q_2^0)$
  - Množina koncových stavů $F_3 = F_1 \times F_2$

5. Příklad

- Předpokládejme, že jazyk L je regulární. Potom podle Pumping lemma (PL) existuje konstanta $p > 0$ taková, že platí: $w \in L \land |w| \geq p \Rightarrow w = xyz \land y \neq \varepsilon \land |xy| \leq p \land xy^i z \in L$ pro $i \geq 0$.
- Uvažujme libovolné celočíselné $p > 0$. Řetezec w zvolme následovně:
  $w = a^{2p} b^{4p} c^{2p+1} \in L$, přičemž $|w| = 8p + 1 \geq p$.
- Potom tedy z PL plyne, že $\exists x, y, z \in \Sigma^*$: $w = xyz \land y \neq \varepsilon \land |xy| \leq p \land xy^i z \in L$ pro $i \geq 0$.

- Z řetězců $x, y, z \in \Sigma^*$, které splňují předchozí podmínku zvolíme libovolné z nich. Tím pádem $x = a^n$, $y = a^m$, $z = a^{2p-m-n}b^{4p}c^{2p+1}$ pro nějaké $n, m \in \mathbb{N}_0, m \neq 0, m + n \leq p$.
- Potom ale také musí platit, že $xy^iz = a^{2p-m+mi}b^{4p}c^{2p+1} \in L$ pro $i \geq 0$. Zvolíme-li $i = 0$, dostáváme $w = a^{2p-m}b^{4p}c^{2p+1}$, což ale znamená, že $w \notin L$, protože $4p \neq 2(2p - m)$, kde $m > 0$. Což je spor. Jazyk L tedy není regulární.

6. Příklad

(a) Z definice regulárních množin plyne, že každá regulární množina je zároveň i jazyk nad danou abecedou. Je možné tedy používat operace definované pro jazyky. Zejména tedy pro operaci $\cdot$ platí distributivní zákon vzhledem k $\cup$ i pro regulární množiny. Nechť A je libovolná regulární množina. Potom z definice A^* (definice 2.9, opora k předmětu TIN) dostáváme:

$$\{\varepsilon\} \cup A \cdot A^* = \{\varepsilon\} \cup A \cdot (A^0 \cup A^1 \cup A^2 \cup A^3 \cup \cdots)$$

S využitím distributivity operací $\cdot$ a $\cup$ a faktu, že $A^0 = \{\varepsilon\}$, můžeme předchozí rovnost upravit následovně

$$\{\varepsilon\} \cup A \cdot (A^0 \cup A^1 \cup A^2 \cup \cdots) = A^0 \cup A^1 \cup A^2 \cup A^3 \cup \cdots = A^*,$$

čímž je rovnost $\{\varepsilon\} \cup A \cdot A^* = A^*$ pro A_{RM} dokázána.

(b) Nechť Σ je libovolná abeceda. Dále předpokládejme regulární množiny $A = \{a\}$ a $B = \{aa\}$, pro $a \in \Sigma$ (a je libovolný symbol abecedy Σ). Vzhledem k tomu, že abeceda je konečná, neprázdná množina symbolů, takový prvek a určitě existuje. Potom ale $A \not\subseteq B$ a $B \not\subseteq A$. $\subseteq$ tedy v A_{RM} není totální uspořádání.