VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Teoretická informatika 2023/2024

Úkol 1

David Chocholatý (xchoch09)

Brno, 29. října 2023

1 Příklad 1

Mějme následující jazyky:

(a) $L_1 = \{ww^R \mid w \in \{a, b, c\}^*\}$

(b) $L_2 = \{w \mid w \in \{a, b, c\}^* \ \wedge \ \#_a(w)\, mod\, 2 = \#_b(w)\, mod\, 2 = 1\}$

Rozhodněte a dokažte, zda jazyky $L_1 \cap L_2$ a $L_1 \cup L_2$ jsou regulární. Pro důkaz regularity sestrojte příslušný konečný automat, nebo gramatiku. Pro důkaz neregularity použijte Pumping lemma, nebo Myhill-Nerodovu větu.

1.1 Řešení

Uvažujme zadané jazyky L_1 a L_2. Neformálně řečeno, jazyk L_1 představuje dobře známý bezkontextový jazyk, který obsahuje všechny řetězce nad $\Sigma = \{a, b, c\}$, které jsou konkatenací dvou stejně dlouhých řetězců takových, že druhý řetězec je reverzací prvního řetězce. Vyplývá, že $L_1 \notin \mathcal{L}_3$, přičemž by to bylo možné dokázat pomocí Pumping lemmatu podobně, jako je tomu následně pro $L_1 \cup L_2$.

Jazyk L_2 obsahuje všechny řetězce nad $\Sigma = \{a, b, c\}$ takové, že počet symbolů a i b je lichý. Jelikož počet kombinací sudosti a lichosti počtu symbolů a a b je konečný, lze řetězce jazyka L_2 přijímat například úplně definovaným deterministickým konečným automatem, a tudíž $L_1 \in \mathcal{L}_3$. Kompletní důkaz by bylo možné provést pomocí Myhill-Nerodovy věty nebo sestrojením zmiňovaného úplně definovaného deterministického konečného automatu.

- $L_1 \cap L_2$

  Průnikem jazyků L_1 a L_2 vzniká nový jazyk

$$L_1 \cap L_2 = \{ww^R \mid w \in \{a, b, c\}^* \ \wedge \ \#_a(ww^R)\, mod\, 2 = \#_b(ww^R)\, mod\, 2 = 1\}.$$

  Jelikož jazyk $L_1 \cap L_2$ je průnikem regulárního a bezkontextového jazyka, lze z uzávěrových vlastností $\mathcal{L}_2$ vůči průniku s regulárními jazyky usoudit, že výsledný jazyk bude regulární (viz [1]). To může pro nově vzniklý jazyk platit tehdy a jen tehdy, když

$$L_1 \cap L_2 = \emptyset \in \mathcal{L}_3 \iff \nexists ww^R \in \{a, b, c\}^* : \#_a(ww^R)\, mod\, 2 = \#_b(ww^R)\, mod\, 2 = 1.$$

  Lemma 1.1. $\nexists ww^R \in \{a, b, c\}^* : \#_a(ww^R)\, mod\, 2 = \#_b(ww^R)\, mod\, 2 = 1.$

  Důkaz. Důkaz bude proveden sporem. Předpokládejme, že

$$\exists ww^R \in \{a, b, c\}^* : \#_a(ww^R)\, mod\, 2 = \#_b(ww^R)\, mod\, 2 = 1.$$

  Uvažujme libovolné $ww^R \in \{a, b, c\}^*$. Vyplývá, že $\#_a(ww^R) = 2 * \#_a(w)$ a $\#_b(ww^R) = 2 * \#_b(w)$. Nutně ovšem musí pro všechna ww^R platit, že $\#_a(ww^R)\, mod\, 2 = \#_b(ww^R)\, mod\, 2 = 2$. Tudíž docházíme ke sporu s předpokladem, že existuje řetězec ww^R obsahující lichý počet symbolů a i b. □

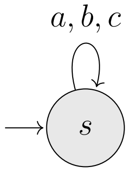
Na základě lemmatu 1.1 platí, že $L_1 \cap L_2 = \emptyset \in \mathcal{L}_3$ (jedná se tedy o regulární jazyk), přičemž takový jazyk je možné přijmout úplně definovaným deterministickým konečným automatem

$$M_1 = (\{s\}, \{a, b, c\}, \delta, s, \emptyset),$$

kde

$$\begin{aligned}\delta : (s, a) &= \{s\},\\ (s, b) &= \{s\},\\ (s, c) &= \{s\}.\end{aligned}$$

Stavový diagram úplně definovaného deterministického konečného automatu M_1 je vyobrazen na obrázku 1.



Obrázek 1: Stavový diagram úplně definovaného deterministického konečného automatu M_1.

- $L_1 \cup L_2$

  Sjednocením jazyků L_1 a L_2 vzniká nový jazyk

$$L_1 \cup L_2 = \{ww^R \,|\, w \in \{a, b, c\}^*\} \cup \{w \,|\, w \in \{a, b, c\}^* \,\wedge\, \#_a(w) \,mod\, 2 = \#_b(w) \,mod\, 2 = 1\}.$$

Na základě skutečnosti, že jazyk $L_1 \cup L_2$ je sjednocením regulárního a neregulárního (bezkontextového) jazyka, je předpokládáno, že $L_1 \cup L_2 \notin \mathcal{L}_3$.

Lemma 1.2. $L_1 \cup L_2 = \{ww^R \,|\, w \in \{a, b, c\}^*\} \cup \{w \,|\, w \in \{a, b, c\}^* \,\wedge\, \#_a(w) \,mod\, 2 = \#_b(w) \,mod\, 2 = 1\} \notin \mathcal{L}_3$.

Důkaz. Důkaz sporem pomocí Pumping lemmatu. Předpokládejme, že $L_1 \cup L_2 \in \mathcal{L}_3$. Pak dle Pumping lemmatu platí, že

$$\begin{aligned}\exists k > 0 :&\forall w \in \Sigma^* : w \in L_1 \cup L_2 \wedge |w| \geq k \implies \\ &\exists x, y, z \in \Sigma^* : w = xyz \wedge y \neq \varepsilon \wedge |xy| \leq k \wedge \forall i \geq 0 : xy^i z \in L_1 \cup L_2.\end{aligned}$$

Uvažujme libovolné $k > 0$ a zvolme $w = a^k bba^k$. Platí, že $w \in L_1 \cup L_2 \wedge |w| = 2 * k + 2 \geq k$. Pro každé rozdělení slova w platí, že $w = xyz \wedge y \neq \varepsilon \wedge |xy| \leq k$. Zřejmě

$$
\begin{aligned}
x &= a^{\alpha_1}, 0 \leq \alpha_1 < k, \\
y &= a^{\alpha_2}, 1 \leq \alpha_2 \leq k \wedge \alpha_1 + \alpha_2 = k, \\
z &= a^{k-\alpha_1-\alpha_2} bba^k.
\end{aligned}
$$

Zvolme $i = 0$. Pak $xy^0z = a^{\alpha_1} a^{k-\alpha_1-\alpha_2} bba^k = a^{k-\alpha_2} bba^k$. Dle Pumping lemmatu by mělo platit, že $a^{k-\alpha_2} bba^k \in L_1 \cup L_2$. To platí tehdy a jen tehdy, když

$$k - \alpha_2 = k \vee ((k - \alpha_2 + k) \, mod \, 2 = 2 \, mod \, 2 = 1).$$

- $k - \alpha_2 = k$:
  Nastane pouze v případě, že $\alpha_2 = 0$, ovšem to je ve sporu s předpokladem, že $\alpha_2 \geq 1$.
- $(k - \alpha_2 + k) \, mod \, 2 = 2 \, mod \, 2 = 1$:
  Nastane pouze v případě, že $(k - \alpha_2 + k) \, mod \, 2 = 1 \wedge 2 \, mod \, 2 = 1$. Druhá část výrazu ovšem nikdy nemůže nastat a vyplývá, že celý výraz neplatí.

Tudíž, $a^{k-\alpha_2} bba^k \notin L_1 \cup L_2$ a dochází ke sporu s předpokladem, že $L_1 \cup L_2 \in \mathcal{L}_3$. $\square$

2 Příklad 2

Uvažujme jazyk $L_3 = \{puvw \,|\, p, v \in \{a, b\}^*, u, w \in \{c, d\}^*, (p = v^R \vee u = w^R)\}$

(a) Sestrojte bezkontextovou gramatiku G_3 takovou, že $L(G_3) = L_3$.

(b) Sestrojte zásobníkový automat Z_3 takový, že $L(Z_3) = L_3$.

2.1 Řešení

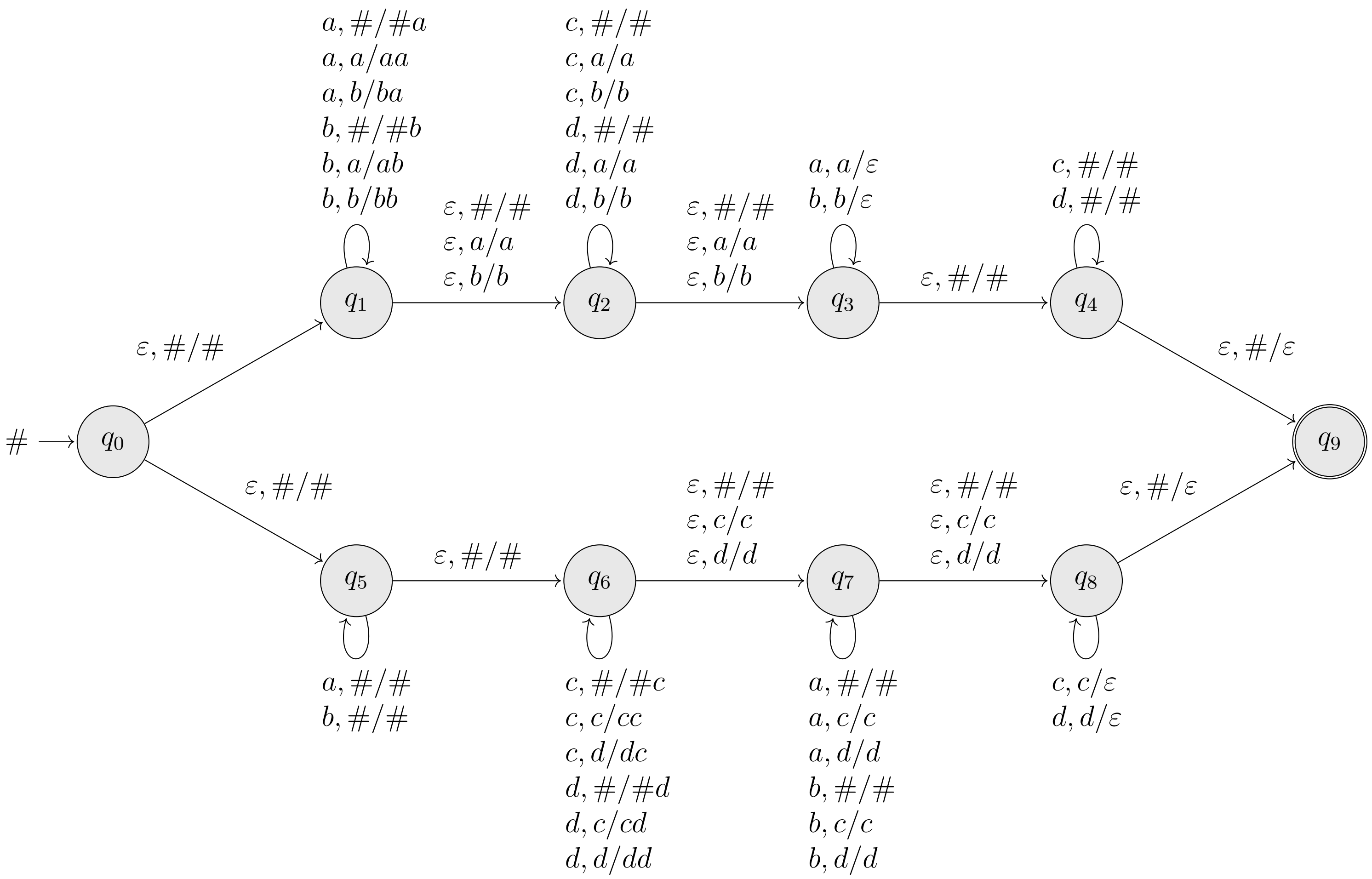
(a) Řešením je bezkontextová gramatika

$$G_3 = (N, \Sigma, P, S),$$

kde

$$
\begin{aligned}
N = \{&S, A, B, C, D\},\\
\Sigma = \{&a, b, c, d\},\\
P = \{&S \to AB \mid CD,\\
&A \to B \mid aAa \mid bAb,\\
&B \to \varepsilon \mid cB \mid dB,\\
&C \to \varepsilon \mid aC \mid bC,\\
&D \to C \mid cDc \mid dDd\}.
\end{aligned}
$$

(b) Řešením je zásobníkový automat Z_3, jehož stavový diagram je vyobrazen na obrázku 2. Notace pravidel je zapisována ve tvaru, kdy jako vrchol zásobníku je uvažován symbol nejvíce napravo.



Obrázek 2: Stavový diagram zásobníkového automatu Z_3.

3 Příklad 3

Rozhodněte a dokažte následující tvrzení:

(a) $\exists L_1 \in \mathcal{L}_2 \setminus \mathcal{L}_3$ takový, že jeho doplněk $\overline{L_1}$ je konečný jazyk.

(b) $\exists L_1 \in \mathcal{L}_2 \setminus \mathcal{L}_3$ takový, že $\forall L_2 \in \mathcal{L}_3 : L_1 \cap L_2 \in \mathcal{L}_2 \setminus \mathcal{L}_3$

(c) $\exists L_1 \in \mathcal{L}_3$ takový, že $\forall L_2 \in \mathcal{L}_2 \setminus \mathcal{L}_3 : L_1 \cap L_2 \in \mathcal{L}_2 \setminus \mathcal{L}_3$

3.1 Řešení

(a) Tvrzení neplatí. Uvažujme libovolný jazyk $L_1 \in \mathcal{L}_2 \setminus \mathcal{L}_3$. Zároveň takový jazyk není konečný, jelikož by náležel do třídy konečných jazyků, značené $\mathcal{L}_{FIN}$, $\mathcal{L}_{FIN} \subset \mathcal{L}_3$.

Dále bude proveden důkaz sporem. Stanovme, že $\overline{L_1} \in \mathcal{L}_{FIN}$ a předpokládejme, že $L_1 \in \mathcal{L}_2 \setminus \mathcal{L}_3$. Jelikož $\overline{L_1} \in \mathcal{L}_{FIN}$, potom platí, že $\overline{L_1} \in \mathcal{L}_3$. Dle uzavřenosti třídy regulárních jazyků vůči operaci doplňku vyplývá, že $L_1 \in \mathcal{L}_3$, přičemž dochází ke sporu.

(b) Tvrzení neplatí. Důkaz bude proveden sporem. Uvažujme libovolný jazyk $L_1 \in \mathcal{L}_2 \setminus \mathcal{L}_3$. Dále stanovme, že $L_2 = \emptyset$, přičemž $L_2 \in \mathcal{L}_3$. Potom nutně platí, že $L_1 \cap L_2 = \emptyset \in \mathcal{L}_3$ a dochází ke sporu.

(c) Tvrzení platí. Stanovme, že $L_1 = \Sigma^* \in \mathcal{L}_3$. Dále uvažujme libovolný jazyk $L_2 \in \mathcal{L}_2 \setminus \mathcal{L}_3$, přičemž $L_2 \subset L_1$. Vyplývá, že $L_1 \cap L_2 = L_2 \in \mathcal{L}_2 \setminus \mathcal{L}_3$.

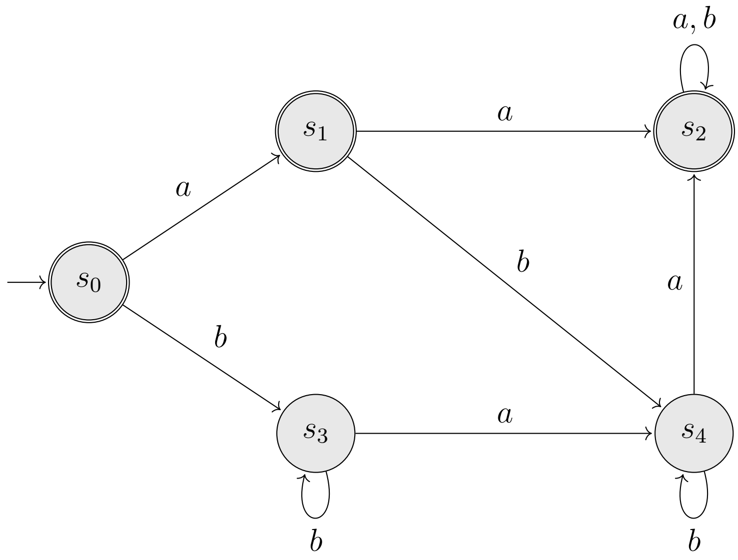
4 Příklad 4

Uvažujme jazyk $L = \{w \in \{a,b\}^* \mid \#_a(w) \geq 2 \vee \#_b(w) = 0\}$. Sestrojte relaci pravé kongruence $\sim$, která splňuje následující dvě podmínky: 1) L je sjednocením některých tříd rozkladu $\Sigma^*/\sim$ a 2) index $\sim$ je o jedna větší než index $\sim_L$.

4.1 Řešení

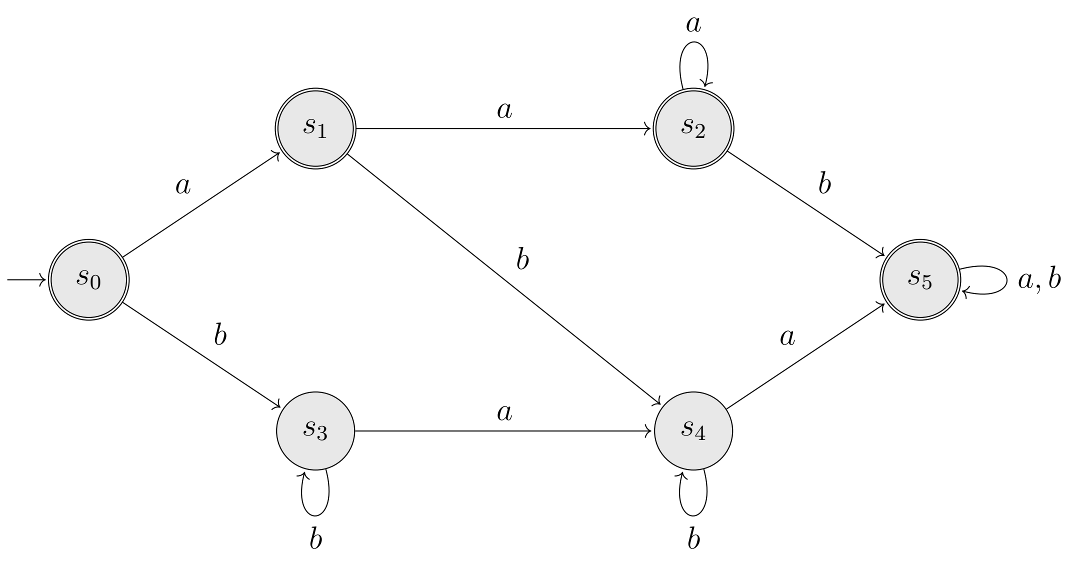
Nejprve sestrojíme úplně definovaný deterministický konečný automat, který neobsahuje nedostupné stavy. Pro to, aby index $\sim$ byl o jedna větší než index $\sim_L$, je nejprve zapotřebí sestrojit minimální úplně definovaný deterministický konečný automat, abychom zjistili počet tříd rozkladu $\Sigma^*/\sim_L$.

Minimální úplně definovaný deterministický konečný automat pro jazyk L je vyobrazen na obrázku 3. Důkaz toho, že se jedná o minimální úplně definovaný deterministický konečný automat, by bylo možné provést s využitím Myhill-Nerodovy věty nebo algoritmu pro převod úplně definovaného deterministického konečného automatu na redukovaný deterministický konečný automat.



Obrázek 3: Stavový diagram minimálního úplně definovaného deterministického konečného automatu pro jazyk L.

Úplně definovaný deterministický konečný automat pro relaci pravé kongruence $\sim$ nesmí být minimální. Tento konečný automat je vyobrazen na obrázku 4. Dále rozklad $\Sigma^*/\sim$ musí obsahovat navíc právě jednu třídu oproti rozkladu $\Sigma^*/\sim_L$. Tudíž, index relace $\sim$ bude o jedna větší oproti indexu relace $\sim_L$.



Obrázek 4: Stavový diagram neminimálního úplně definovaného deterministického konečného automatu pro jazyk L.

Definice 4.1. Nechť $M = (Q, \Sigma, \delta, s_0, F)$ je nedeterministický konečný automat a nechť $q \in Q$. Jazyk přístupových řetězců stavu q je definován následovně

$$L^{-1}(q) = \{w \in \Sigma^* \mid (s_0, w) \vdash^* (q, \varepsilon)\}.$$

Dále za použití definice 4.1 najdeme jazyky přístupových řetězců nad $\Sigma = \{a, b\}$ všech stavů neminimálního úplně definovaného deterministického konečného automatu:

$$\begin{aligned}
L^{-1}(s_0) &= \{w \in \Sigma^* \mid \#_a(w) = 0 \land \#_b(w) = 0\}, \\
L^{-1}(s_1) &= \{w \in \Sigma^* \mid \#_a(w) = 1 \land \#_b(w) = 0\}, \\
L^{-1}(s_2) &= \{w \in \Sigma^* \mid \#_a(w) \geq 2 \land \#_b(w) = 0\}, \\
L^{-1}(s_3) &= \{w \in \Sigma^* \mid \#_a(w) = 0 \land \#_b(w) \geq 1\}, \\
L^{-1}(s_4) &= \{w \in \Sigma^* \mid \#_a(w) = 1 \land \#_b(w) \geq 1\}, \\
L^{-1}(s_5) &= \{w \in \Sigma^* \mid \#_a(w) \geq 2 \land \#_b(w) \geq 1\}.
\end{aligned}$$

Najdeme rozklad množiny Σ^*, který je daný jako množina jazyků přístupových řetězců jednotlivých stavů:

$$P = \{L^{-1}(s_0), L^{-1}(s_1), L^{-1}(s_2), L^{-1}(s_3), L^{-1}(s_4), L^{-1}(s_5)\}.$$

Najdeme pravou kongruenci k tomuto rozkladu takovou, že $P = \Sigma^*/\sim$:

$$\begin{aligned}
\forall u, v \in \Sigma^* : u \sim v \iff\ & (\#_a(u) = 0 \land \#_a(v) = 0 \land \#_b(u) = 0 \land \#_b(v) = 0) \lor \\
& (\#_a(u) = 1 \land \#_a(v) = 1 \land \#_b(u) = 0 \land \#_b(v) = 0) \lor \\
& (\#_a(u) \geq 2 \land \#_a(v) \geq 2 \land \#_b(u) = 0 \land \#_b(v) = 0) \lor \\
& (\#_a(u) = 0 \land \#_a(v) = 0 \land \#_b(u) \geq 1 \land \#_b(v) \geq 1) \lor \\
& (\#_a(u) = 1 \land \#_a(v) = 1 \land \#_b(u) \geq 1 \land \#_b(v) \geq 1) \lor \\
& (\#_a(u) \geq 2 \land \#_a(v) \geq 2 \land \#_b(u) \geq 1 \land \#_b(v) \geq 1).
\end{aligned}$$

- Jistě jde o relaci ekvivalence, protože vznikla z rozkladu množiny.
- Jistě jde o pravou kongruenci, protože jde o ekvivalenci, které vznikla pomocí jazyka přístupových řetězců stavů konečného automatu.
- Sjednocením některých tříd rozkladu $\Sigma^*/\sim$ získáme jazyk L:

  $$L = L^{-1}(s_0) \cup L^{-1}(s_1) \cup L^{-1}(s_2) \cup L^{-1}(s_5) = \{w \in \{a, b\}^* \mid \#_a(w) \geq 2 \lor \#_b(w) = 0\}.$$

  Index $\sim$ je o jedna větší oproti indexu prefixové ekvivalence $\sim_L$, neboť $|\Sigma^*/\sim| = 6$ a $|\Sigma^*/\sim_L| = 5$.

Použitá literatura

[1] Rozenberg, G.; Salomaa, A. (editoøi): *Handbook of Formal Languages*. Berlin, Heidelberg: Springer-Verlag, první vydání, 1997, ISBN 978-3-642-63863-3, 873 s.