Úkol 2 (wrzecond)

pátek 13. listopadu 2020 15:16

Rozhodněte, zdali jsou následující jazyky L_1 a L_2 regulární, či nikoliv. Pokud jazyk není regulární, formálně toto tvrzení dokažte.

Pokud jazyk regulární je, popište jej jedním z formalismů pro popis regulárních jazyků (tj. KA, RG, nebo RV). Dále pro něj nalezněte konstantu pumping lemma p (jejíž existence je pro regulární jazyky v pumping lemma zaručena) a zdůvodněte, proč je vaše konstanta „správná". (Pozn. $\mathbb{N} = \{1, 2, 3, \ldots\}$)

$$L_1 = \{2^m 1^n 2^k : m, n, k \in \mathbb{N} \wedge k \neq m + n \wedge m > 1\}$$

$$L_2 = \{w : w \in \{a, b\}^* \wedge |w|_a \bmod 3 = |w|_b\} \cap \{a^m b^k b^i : m, k, i \in \mathbb{N}_0 \wedge k < m\}$$

$L_1 = \{2^m 1^n 2^k : m, n, k \in \mathbb{N} \wedge k \neq m+n \wedge m > 1\}$

Úvaha: Tento jazyk bude nejspíše neregulární, jelikož si musíme pamatovat počet dvojek a jedniček, který není nijak shora omezen.

Důkaz neregularity (sporem):

Nechť L_1 je regulární. Pak i jeho doplněk $\overline{L_1}$ je regulární (viz. algoritmus na konstrukci automatu pro doplněk jazyka).

Pokud je $\overline{L_1}$ regulární, pak má dle Pumping Lemma pumpující vlastnost, tedy:

$(\exists p \geq 1)(\forall w \in \overline{L_1})[|w| \geq p \Rightarrow (\exists x,y,z \in \Sigma^*)(w = xyz \wedge |xy| \leq p \wedge y \neq \varepsilon$
$\wedge (\forall k \geq 0)\ xy^k z \in \overline{L_1})]$

Pre spor dokažme:

$(\forall p \geq 1)(\exists w \in \overline{L_1})[|w| \geq p \wedge (\forall x,y,z \in \Sigma^*)(w \neq xyz$
$\vee |xy| > p \vee y = \varepsilon \vee (\exists k \geq 0)\ xy^k z \notin \overline{L_1})]$

$\hookrightarrow xy^k z \in L_1$

Volme jako slovo $w = 2^p 1^p 2^{2p}$.

O tomto slovu určitě víme, že $w \notin L_1$,
a tedy $w \in \overline{L_1}$, jelikož $k = m+n$ $(2p = p+p)$

Taktéž triviálně platí, že $|w| = 4p \geq p$.

Uvažujme nyní všechna možná rozdělení w.
Z předpokladu $|xy| \leq p$ určitě x a y
nemohou obsahovat nic jiného než dvojky.

Tedy:

$x = 2^r, \quad r \geq 0$

$y = 2^s, \quad s \geq 1, \quad r+s \leq p$

$y \neq \varepsilon \qquad |xy| \leq p$

$z = 2^t 1^p 2^{2p}, \quad r+s+t = p$

(Volit $k=0$ nemůžeme, jelikož napr. pro $p=2$, $w = \underset{x}{2}\underset{y}{2}\underbrace{112222}_{z}$ bychom získali jako xy^kz 2112222, což $\notin L_1$, jelikož $m \leq 1$) $\boxed{\text{VOLBA } k}$

Volme $k=2$. Pak xy^kz lze popsat jako

$$2^r 2^{2s} 2^t 1^p 2^{2p} =$$

$$= 2^{r+2s+t} 1^p 2^{2p} \overset{r+s+t=p}{=} \underline{2^{p+s} 1^p 2^p}$$

$$2^{\overset{m}{p+s}} 1^{\overset{n}{p}} 2^{\overset{k}{p}} \in L_1 \ (\notin \overline{L_1}) \iff \overset{m}{p+s} > 1 \wedge \overset{m+n}{p+s+p} \neq \overset{k}{2p}$$

což platí, jelikož ① $p \geq 1, s \geq 1 \Rightarrow p+s \geq 2$

② $s \geq 1 \Rightarrow 2p+s > 2p$

$\Rightarrow 2p+s \neq 2p$

Našli jsme tedy pro každé $p \geq 1$ slovo w takové, že žádné jeho rozdělení nesplňuje $|xy| \leq p \wedge y \neq \varepsilon \wedge (\forall k \geq 0)(xy^kz \in \overline{L_1})$

$\Rightarrow$ jazyk <u>$\overline{L_1}$ není regulární</u>.

(obměna implikace) $\Rightarrow$ L_1 není regulární. $\square$

$$L_2 = \left[\{ w : w \in \{a,b\}^* \wedge |w|_a \bmod 3 = |w|_b \}\right]$$

$$\cap$$

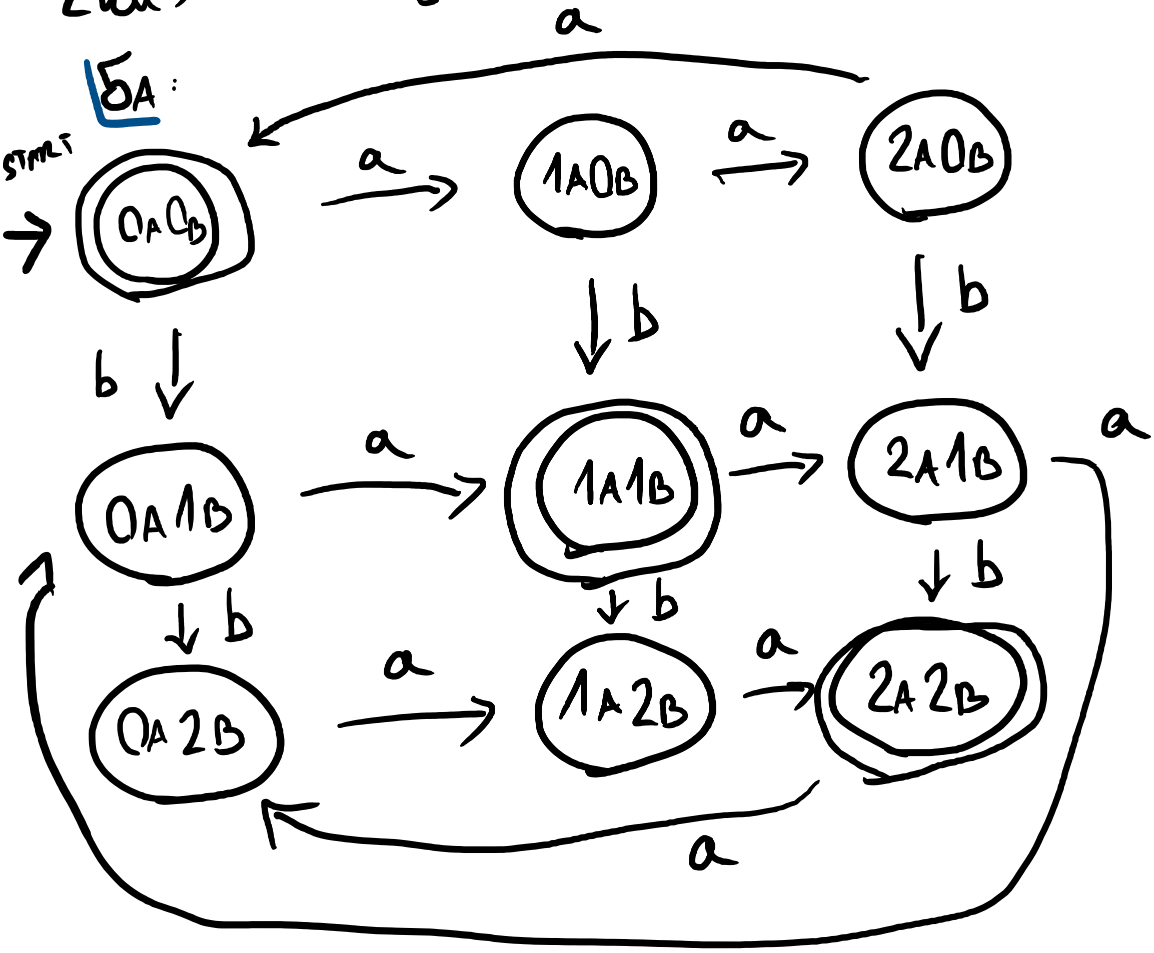
$L_{2a} =$ $L_{2b} =$ $\left[\{ a^m b^k b^i : m,k,i \in \mathbb{N}_0 \wedge k < m \}\right]$

Úvaha: Tento jazyk je průnikem dvou jazyků, nazvěme si je L_{2a} a L_{2b}. Ty jsou regulární $\Leftrightarrow$ jsou přijímány konečnými automaty. Nazvěme si je M_A, $L(M_A) = L_{2a}$, M_B, $L(M_B) = L_{2b}$. Pokud jsou regulární, pak dle algoritmu na konstrukci kon. automatu pro průnik 2 regulárních jazyků by i L_2 byl regulární.

Náš konečný automat M_A si musí držet modulo počtu přijmutých a a počet přijmutých b. Jelikož ale cokoliv mod 3 = $\{0,1,2\}$, pak víme, že žádné slovo z L_{2a} nebude mít více než 2 b.

Zkusme tedy:

δ_A:



$$M_A = (\{0_A 0_B, \dots, 2_A 2_B\}, \{a, b\}, \delta_A, 0_A 0_B, \{0_A 0_B, 1_A 1_B, 2_A 2_B\})$$

Dobrá věc se podařila, máme kon. automat.

Zkusme nyní sestrojit automat M_B:

$$L_{2b} = \{a^m b^k b^i : m, k, i \in \mathbb{N}_0 \wedge k < m\}$$

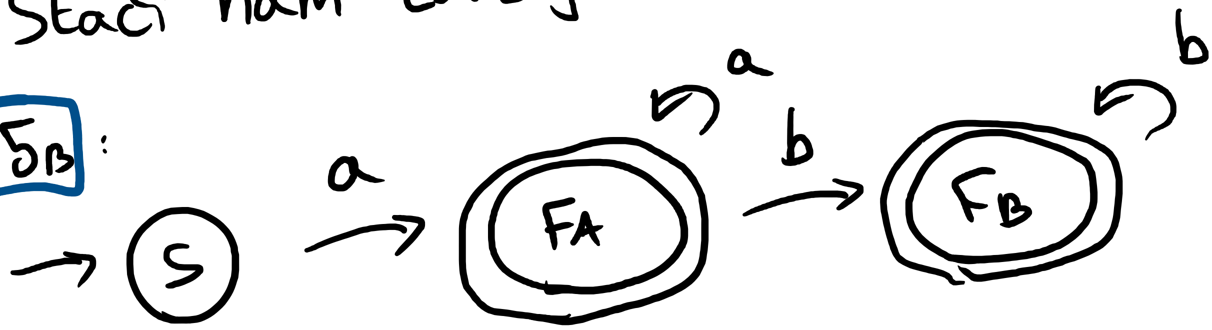
Byť působí zápis komplexně, L_{2b} se dá popsat jednoduše jako

$$L_{2B} = \{ a^m b^i : m, i \in \mathbb{N}_0, (m > 0) \}$$

$m \geq 1$

Nejmírnější podmínku totiž získáme pro $k = 0$. Všimněme si, že opravdu neztrácíme žádná slova, ani žádná nepřidáváme. Takovýto jazyk je ale velmi lehké přijmout. Stačí nám takovýto konečný automat:

$\boxed{\delta_B}$:



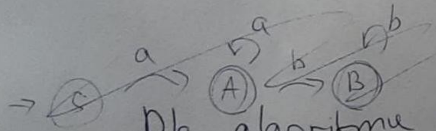
$$M_B = (\{S, F_A\}, \{a, b\}, \delta_B, S, \{F_A, F_B\})$$

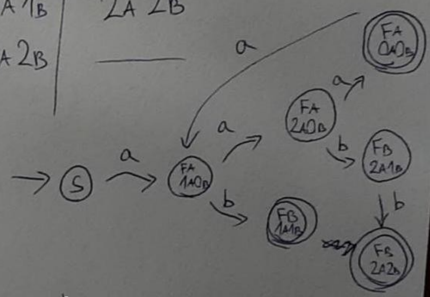
Algoritmem na konstrukci autom. pro průnik jazyků a následnou eliminací zbytečných stavů získáme kýžený automat M_2.

(FOTO ALGORITMU)

L_{2b},	a	b
→ S	F_A	—
← F_A	F_A	F_B
← F_B	—	F_B

L_{2a},	a	b
↔ 0_A0_B	1_A0_B	0_A1_B
0_A1_B	1_A1_B	0_A2_B
0_A2_B	1_A2_B	—
1_A0_B	2_A0_B	1_A1_B
← 1_A1_B	2_A1_B	1_A2_B
1_A2_B	2_A2_B	—
2_A0_B	0_A0_B	2_A1_B
2_A1_B	0_A1_B	2_A2_B
← 2_A2_B	0_A2_B	—





Dle algoritmu na průnik jazyků

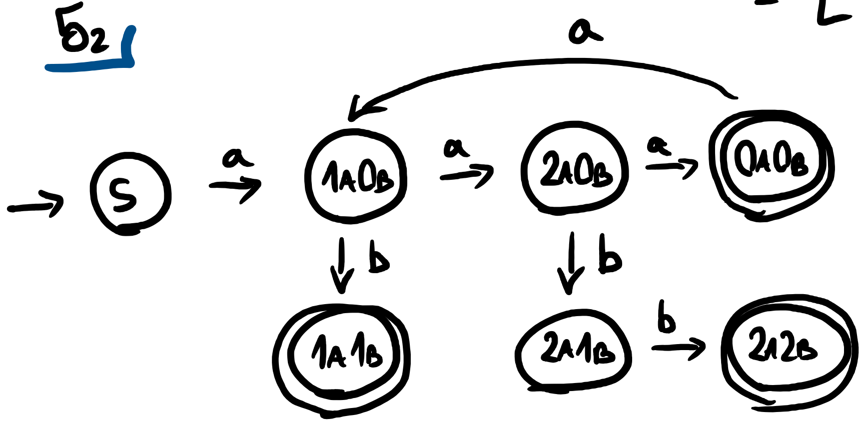
L_2	a	b		a	b
→ S, 0_A0_B	F_A, 1_A0_B	—	← F_A, 0_A0_B	F_A, 1_A0_B	~~F_B, 0_A1_B~~
F_A, 1_A0_B	F_A, 2_A0_B	F_B, 1_A1_B	F_B, 2_A1_B	—	F_B, 2_A2_B
F_A, 2_A0_B	F_A, 0_A0_B	F_B, 2_A1_B	~~F_B, 1_A2_B~~	—	—
← F_B, 1_A1_B	—	~~F_B, 1_A2_B~~	~~F_B, 0_A1_B~~	—	~~F_B, 0_A2_B~~
		—	~~F_B, 0_A2_B~~	—	—
			F_B, 2_A2_B	—	—

(Poznámka: TOTO JE ELIMINACE ZBYTEČNÝCH STAVŮ)

Automat M_2, $L(M_2) = L_2 = L_{2a} \cap L_{2b} =$

$= L(M_A) \cap L(M_B)$

δ_2,



(Pro zjednodušení vynecháno F_A, F_B z názvu stavů)

$$M_2 = (\{S, 0_A0_B, 1_A0_B, 2_A0_B, 1_A1_B, 2_A1_B, 2_A2_B\}, \{a,b\}, \delta_2, S, \{0_A0_B, 1_A1_B, 2_A2_B\})$$

Což lze také popsat regulárním výrazem

$$v = a(aaa)^*(aa+b+abb)$$ $h(v) = L_2$

L_2 lze také popsat jako $\{1,2,3,\dots\}$

$$L_2 = \{ a^m b^{m \bmod 3} : m \in \mathbb{N} \}$$

- **UKÁZKA p, PRO KTERÉ PLATÍ PUMPING LEMMA**

$$(\exists p \geq 1)(\forall w \in L_2)[|w| \geq p \Rightarrow (\exists x,y,z \in \Sigma^*)(w = xyz \wedge |xy| \leq p \wedge y \neq \varepsilon \wedge xy^kz \in L_2)]$$

Volme $p=5$. Pak lze každé slovo $w = a^m b^{m \bmod 3}$ rozdělit následovně:

$x = \varepsilon$ $y = aaa$
$z = a^{m-3} b^{m \bmod 3}$

Všimněme si, že platí $|xy| = 3 \leq 5$, a také $y \neq \varepsilon$.

Víme, že $xy^k z = a^{3k} a^{m-3} b^{m \bmod 3}$.

Platí ale, že $xy^k z \in L_2$?
To platí, protože ① $3k + m - 3 \bmod 3 = m \bmod 3$
(což platí, jelikož násobky 3 modulo výrazu neovlivní)
② $3k + m - 3 \geq 1$ ($k \geq 0 \Rightarrow m \geq 4$*)

Zbývá nám tedy splnit 2 podmínky:
① $m \geq 4$ ② $|w|_a \geq 3$, aby šlo w rozdělit na xyz

Víme, že $p \geq 5 \Rightarrow |w| \geq 5$ a víme, že $|w|_b \leq 2$.
Z toho přímo vyplývá, že $|w|_a = m \geq 3$.

Jelikož ale pro $m = 3$ $|w| = |a^3| = 3 < 5$, platí
$m > 3 \Rightarrow m \geq 4$.

$\square$