1. príklad

(a) Majme abecedu $\Sigma = \{0, 1\}$. Vymyslite aspoň tri príklady jazykov nad touto abecedou. Sú vaše jazyky konečné alebo nekonečné?

(b) Majme abecedu $\Sigma = \{a, b, c, d\}$. Vymyslite dva konečné jazyky L_1, L_2 nad abecedou Σ.

(i) $L_1 \cup L_2 = ?$

(ii) $L_1 \cap L_2 = ?$

(iii) $L_1.L_2 = ?$

(iv) $L_2.L_1 = ?$

(c) Čo z bežného (informatického) života môže byť jazykom? Nad akou abecedou?

(d)* Majme abecedu $\Sigma = \{a, b, c, d, e\}$ a jazyk $L = \cup_{n=1}^{6} \Sigma^n$. Koľko slov z jazyka má vlastný prefix ab?

(e) Máme abecedu Σ_{bool} a jazyky $A, B, C \subseteq \Sigma^*$ také, že:

- $A = \{0, 1, 00, 11, 000, 111, 0000, 1111\}$
- $B = \{w \in \Sigma_{bool}^* | |w| \geq 2\}$
- $C = \{w \in \Sigma_{bool}^* | |w| \leq 2\}$

Určite jazyky:

(i) $A \cup B$

(ii) $A \cap B$

(iii) $A - B$

(iv) $A \cap C$

(v) $B \cup C$

(vi) $B \cap C$

Riešenie 1.e

Najprv si slovne popíšme, aké slová obsahujú jazyky B a C. V jazyku B sú slová nad Σ_{bool} dĺžky aspoň 2. Do jazyka C patria všetky slová Σ_{bool} dĺžky maximálne 2, čiže $\{\lambda, 0, 1, 00, 01, 10, 11\}$.

(a) $A \cup B = \{\ w \in \Sigma_{bool}^* | |w| \geq 1\} = \Sigma_{bool}^+$

(b) $A \cap B = \{00, 11, 000, 111, 0000, 1111\}$

(c) $A - B = \{0, 1\}$

(d) $A \cap C = \{0, 1, 00, 11\}$

(e) $B \cup C = \Sigma^*_{bool}$

(f) $B \cap C = \{w \in \Sigma^*_{bool} | |w| = 2\} = \Sigma^2_{bool}$

2. príklad

Majme abecedu Σ, nech $A, B, C \subseteq \Sigma^*$. Dokážte, alebo vyvráťte že:

(a)* $A(B \cup C) = AB \cup AC$

(b) Ak $A \subseteq B$, potom pre každé $n \in Z^+$ platí, že $A^n \subseteq B^n$.

(c)* Platí nasledovná rovnosť $A(B \cap C) = AB \cap AC$, pre jazyky nad abecedou $\Sigma = \{0\}$?

(d)* Nech $A \subseteq \Sigma_1^*$ a $B, C \subset \Sigma_2^*$ pre abecedy Σ_1, Σ_2 také, že $\Sigma_1 \cap \Sigma_2 = \emptyset$. Platí $A(B \cap C) = AB \cap AC$?

(e) Ak $A^2 = B^2$, potom $A = B$.

Riešenie 2.b

Túto implikáciu dokážeme pomocou matematickej indukcie.

- Báza: Ukážeme, že tvrdenie platí pre triviálny prípad, čiže v našom prípade pre $n = 1$. Ak $A \subseteq B$, potom $A^1 \subseteq B^1$. Prvá mocnina jazyka je jazyk samotný, preto to samozrejme platí.

- Indukčný predpoklad: Predpokladáme, že tvrdenie platí pre $n = k$. Čiže Ak $A \subseteq B$, potom $A^k \subseteq B^k$, pre $k \in Z^+$.

- Indukčný krok: Ideme dokázať, že ak je splnený IP, tvrdenie platí aj pre $n = k + 1$. Nech $x \in A^{k+1}$ potom (z def. mocniny jazyka) $x \in \{x_1x_2 \mid x_1 \in A, x_2 \in A^k\}$. Vychádzame z predpokladu, ak $A \subseteq B$ a $x_1 \in A$, potom aj $x_1 \in B$. Tiež z IP, ak $A^k \subseteq B^k$ a $x_2 \in A^k$, potom aj $x_2 \in B^k$. Čiže, $x \in \{x_1x_2 \mid x_1 \in B, x_2 \in B^k\} = B^{k+1}$.

3. príklad

Majme abecedu $\Sigma = \{0, 1\}$.

(a) Nech $A = \Sigma^2$. Čo je potom A^*?

(b) Nech A je rovnaké ako v časti (a), a $B = \{0, 1\}$. Aké slová obsahuje jazyk BA^*?

(c) Aký je rozdiel medzi jazykmi $\{0\}\{0,1\}^*$ a $\{0\}\{0,1\}^+$? Čo obsahujú?

(d) Ako zapíšete jazyk všetkých slov nad abecedou Σ, ktoré obsahujú sufix 11?

(e) Ako zapíšete jazyk všetkých slov nad abecedou Σ, ktoré obsahujú podslovo 010?

(f) Čo budeme musieť zmeniť v riešení predchádzajúcich dvoch úloh, aby jazyky obsahovali slová s vlastným sufixom resp. vlastným podslovom?

(g) Do ktorých z nasledovných jazykov nad Σ^* patrí slovo 00010?

(i) $\{0, 1\}^*$
(ii) $\{00\}\{0\}^*\{10\}$
(iii) $\{000, 101\}\{10, 11\}$
(iv) $\{000\}^*\{1\}^*\{0\}$
(v) $\{00\}^*\{10\}^*$
(vi) $\{0\}^*\{1\}^*\{0\}^*$

Riešenie 3. príkladu

(a) $A = \Sigma^2 = \Sigma.\Sigma = \{00, 01, 10, 11\}$. Čiže A obsahuje slová dĺžky 2 nad abecedou Σ. Ak spravíme Kleeneho uzáver tohto jazyka, dostaneme zjednotenie všetkých mocnín jazyka A. Pozrime sa aspoň na prvé štyri: $A^0 = \{\lambda\}$, $A^1 = A$, $A^2 = AA = \{0000, 0001, 0010, 0011...\}$, $A^3 = AAA = \{000000, 000001, 000010...\}$. Vidíme, že dĺžka slov v každej ďalšej mocnine bude o dva dlhšia. A^* tak obsahuje všetky slová nad Σ párnej dĺžky.

(b) Vieme, že A^* sú slová párnej dĺžky, keď ich prireťazíme k slovám z B, ktoré sú dĺžky 1, dostaneme všetky slová nepárnej dĺžky.

(c) Jazyk $\{0,1\}^*$ obsahuje aj λ, kým $\{0,1\}^+$ nie. Čiže jazyk $\{0\}\{0,1\}^*$ bude obsahovať všetky slová nad Σ s prefixom 0, a jazyk $\{0\}\{0,1\}^+$ všetky slová s vlastným prefixom 0 nad Σ (tj. dĺžky aspoň 2).

(d) Podobne ako v predchádzajúcom príklade: $\{0,1\}^*\{11\}$.

(e) $\{0,1\}^*\{010\}\{0,1\}^*$

(f) Vo výsledku z (d) stačí namiesto $*$ dať $+$. Ale stačí to aj v (e)? Aké najkratšie slová by patrili do jazyka $\{0,1\}^+\{010\}\{0,1\}^+$? Aspoň dĺžky 5, pretože aj na začiatku, aj na konci slova by musel byť aspoň jeden znak. Ale napr. aj slovo 1010 by malo patriť do tohto jazyka, keďže má vlastné poslovo 010.

Preto to skúsme zapísať aj zjednotenie dvoch jazykov: $\{0,1\}^*\{010\}\{0,1\}^+\cup \{0,1\}^+\{010\}\{0,1\}^*$. Tak máme zaručené, že buď na konci alebo na začiatku slova bude ešte aspoň jeden znak okrem poslova 010.

(g) Slovo 00010 patrí do jazykov (i), (ii), (iii), (iv), (vi). Nepatrí iba do jazyka (v), keďže zápis $\{00\}^*$ znamená, že počet núl na začiatku slova musí byť párny.

4. príklad

Zapíšte jazyk $L = \{0^{m+4}\, 1^{2m}\, 2^{m+3} \mid m \geq 1\}$ ako prienik dvoch jazykov L_1 a L_2, ktoré spĺňajú súčasne nasledujúce podmienky:

(a) $L_1 \nsubseteq L_2$, $L_2 \nsubseteq L_1$

(b) $L_1, L_2 \subseteq \{0\}^+\{1\}^+\{2\}^+$

(c) $L_1 - L$ a $L_2 - L$ sú nekonečné jazyky.

Odôvodnite.

Riešenie 4. príkladu

(a) Hľadáme jazyky, ktoré sú rôzne, ale ich druhé mocniny sú rovnaké. Môžeme skúšať hľadať takéto konečné jazyky, uvažovať, že potrebujeme nejaké slovo vyskladať iným spôsobom, ako vznikne v druhom jazyku. Mohli by ste navrhnúť napr. takúto dvojicu jazykov: $A = \{a, aaa\}, B = \{\lambda, aa, aaaa\}$.

Pozrime sa na ich druhé mocniny. $A^2 = \{aa, aaaa, aaaaaa\} = \{a^2, a^4, a^6\}$. $B^2 = \{\lambda, aa, aaaa, aaaaaa\} = \{\lambda, a^2, a^4, a^6\}$. Vidíme, že sú takmer rovnaké, presnejšie $A^2 \subset B^2$, ale nie sú rovnaké, lebo v B^2 je aj prázdne slovo λ. Je častou chybou, že si neuvedomíte, že $\lambda.\lambda = \lambda$.

Z toho vyplýva, že nevieme mať dva jazyky, v ktorých v jednom prázdne slovo je a v druhom nie je. Keby sme dali do oboch prázdne slovo, tak v ich druhých mocninách by bol aj jazyk samotný (jazyk zreťazený s prázdnym slovom), čo by nám to opäť mohlo pokaziť.

Pri konečných jazykoch bude ich druhá mocnina vždy obsahovať viac slov[1] (zamyslite sa, prečo). Tak skúsme nájsť nejaké nekonečné jazyky.

Nech $A = \{\lambda, a, aa...\} = \{a^k \mid k \in N\}$.

$B = \{\lambda, a, aaa, aaaa...\} = \{a^k \mid k \in N, k \neq 2\}$

Ich druhé mocniny budú rovné $A^2 = B^2 = A = \{a^k \mid k \in N\}$, keďže slovo aa v B^2 vznikne zreťazením $a.a$. Vidíme, že tieto jazyky rovnaké nie sú, ale ich druhé mocniny áno.

(b) Jazyk L je nekonečný. Keďže $L_1 - L$, $L_2 - L$ sú nekonečné jazyky, tak aj L_1 a L_2 budú nekonečné. Aby sme splnili podmienku (ii) a zároveň, aby sme jednoznačne vedeli povedať, že $L_1 \cap L_2 = L$, použime jazyk L s miernymi úpravami pre oba jazyky.

Nech: $L_1 = \{0^{m+k}\ 1^{2m}\ 2^{m+3} \mid m, k \geq 1\}$
$L_2 = \{0^{m+4}\ 1^{lm}\ 2^{m+3} \mid m, l \geq 1\}$.

Čo sme zmenili? V oboch jazykoch pribudol jeden parameter navyše. Čo to znamená? Skúsme si obaja jazyky napísať ako zjednotenie jazykov postupnou iteráciou týchto parametrov.

Čiže: $L_1 = \{0^{m+1}\ 1^{2m}\ 2^{m+3} \mid m \geq 1\} \cup \{0^{m+2}\ 1^{2m}\ 2^{m+3} \mid m \geq 1\} \cup \{0^{m+3}\ 1^{2m}\ 2^{m+3} \mid m \geq 1\} \cup \{0^{m+4}\ 1^{2m}\ 2^{m+3} \mid m \geq 1\}...$

$L_2 = \{0^{m+4}\ 1^{m}\ 2^{m+3} \mid m \geq 1\} \cup \{0^{m+4}\ 1^{2m}\ 2^{m+3} \mid m \geq 1\} \cup \{0^{m+4}\ 1^{3m}\ 2^{m+3} \mid m \geq 1\} \cup \{0^{m+4}\ 1^{4m}\ 2^{m+3} \mid m \geq 1\}...$

Vidíme, že jazyk L je podmnožinou oboch týchto jazykov. Je aj ich prienikom? V jazyku L_1 budú rovnaké slová ako v L_2 iba ak v jazyku L_1 bude $k = 4$ a zároveň v jazyku L_2 bude $l = 2$. Pre iné hodnoty k, l slová z jedného jazyka nebudú patriť do druhého (a opačne).

Ako budú vyzerať $L_1 - L$ a $L_2 - L$? Pozrime sa vyššie. Vidíme, že ak z oboch jazykov vynecháme jazyk L, parametre k, l nemôžu nadobudnúť hodnoty 4 resp. 2. To znamená, že:

[1] okrem jazyka $\{\lambda\}$

$L_1 - L = \{0^{m+k}\ 1^{2m}\ 2^{m+3} \mid m, k \geq 1\, k \neq 4\}$.
$L_2 - L = \{0^{m+4}\ 1^{lm}\ 2^{m+3} \mid m, l \geq 1, l \neq 2\}$.

Oba jazyky sú nekonečné (podmienka (iii)), keďže množina celých kladných čísel okrem 4 (resp. okrem 2) je nekonečná. Podmienku (ii) sme splnili, v jazyku L_1 by mohlo byť aj $k = 0$, keďže $m \geq 1$, čiže tam vždy bude aspoň jedna nula. V jazyku L_2 nemôže byť $l = 0$, keďže v takom prípade by v slovách neboli žiadne jednotky.

A čo s podmienkou (i)? Pozrime sa na slovo z jazyka L_1, keď $m = 1 \wedge k = 1$. $w = 0^{1+1}\ 1^2\ 2^{1+3}\ = 00112222$. Toto slovo sa určite nebude nachádzať v jazyku L_2. Prečo? Lebo pre $m = 1$ bude v slovách z jazyka L_2 päť núl, dve nuly sa v žiadnom slove z L_2 nachádzať nebudú. To znamená, že $L_1 \not\subseteq L_2$.

Skúsme to aj opačne. Pre $m = 1 \wedge l = 1$ bude slovo $w = 0^{1+4}\ 1^{1.1}\ 2^{1+3}\ = 0000012222$. Toto slovo nepatrí do jazyka L_1, keďže v ňom budú vždy slová, ktoré majú párny počet jednotiek. Z toho vyplýva, že $L_2 \not\subseteq L_1$.