

Cvičení 7

Příklad 1: Pro každý z následujících jazyků uveďte nějakých 5 slov, která do něj patří, a nějakých 5 slov, která do něj nepatří.

a) $L_1 = \{w \in \{0, 1\}^* \mid \text{délka slova } w \text{ je menší než } 5\}$

Řešení: Slova z jazyka L_1 jsou např. $\varepsilon, 0, 1, 00, 01$, atd. Do jazyka L_1 nepatří např. $00000, 00001, 00010, 000000, 1111111$.

b) $L_2 = \{w \in \{a, b\}^* \mid \text{počet výskytů symbolu } b \text{ ve slově } w \text{ je sudý}\}$

Řešení: Slova z jazyka L_2 jsou např. $\varepsilon, a, aa, bb, aaa, abb$, atd. Do jazyka L_2 nepatří např. b, ab, ba, aab, aba .

c) $L_3 = \{w \in \{0, 1\}^* \mid \text{ve } w \text{ je každá } 0 \text{ (přímo) následována } 1\}$

Řešení: Slova z jazyka L_3 jsou např. $\varepsilon, 1, 01, 11, 101101$, atd. Do jazyka L_3 nepatří např. $0, 10, 001, 010, 1010$.

d) $L_4 = \{w \in \{0, 1\}^* \mid w \text{ začíná a končí stejným symbolem}\}$

Řešení: Slova z jazyka L_4 jsou např. $0, 1, 00, 11, 000, 010$, atd. Do jazyka L_4 nepatří např. $\varepsilon, 01, 10, 001, 011$.

e) $L_5 = \{w \in \{a, b\}^* \mid w \text{ obsahuje jako podslovo sekvenci } abb\}$

Řešení: Slova z jazyka L_5 jsou např. $abb, aabb, abba, abbb, babb$, atd. Do jazyka L_5 nepatří např. $\varepsilon, a, b, aa, ab, ba, bb, aaa$.

Příklad 2: Předpokládejme, že $\Sigma = \{a, b\}$ a $n \in \mathbb{N}$.

a) Kolik existuje slov ze Σ^* délky n ?

Řešení: 2^n

b) Kolik existuje slov ze Σ^* délky nejvýše n ?

Řešení:

$$2^0 + 2^1 + \dots + 2^n = \sum_{i=0}^n 2^i = \frac{2^{n+1} - 1}{2 - 1} = 2^{n+1} - 1$$

Příklad 3: Uvažujme následující jazyky:

$L_1 = \{w \in \{0, 1\}^* \mid \text{ve } w \text{ je každá } 0 \text{ (přímo) následována } 1\}$

$L_2 = \{w \in \{0, 1\}^* \mid w = w^R\}$

a) Vyjmenujte prvních 5 slov z každého z jazyků L_1, L_2 (nejmenších vzhledem k uspořádání $<_L$).

Řešení:

L_1 : $\varepsilon, 1, 01, 11, 011$

L_2 : $\varepsilon, 0, 1, 00, 11$

b) Vyjmenujte prvních 5 slov z každého z jazyků $\overline{L_1}$, $\overline{L_2}$.

Řešení:

$\overline{L_1}$: 0, 00, 10, 000, 001

$\overline{L_2}$: 01, 10, 001, 011, 100

c) Vyjmenujte prvních 5 slov z jazyka $L_1 \cap L_2$.

Řešení:

$L_1 \cap L_2$: ε , 1, 11, 101, 111

d) Vyjmenujte prvních 5 slov z jazyka $L_1 \cup L_2$.

Řešení:

$L_1 \cup L_2$: ε , 0, 1, 00, 01

Příklad 4: Uvažujme jazyky nad abecedou $\{a, b\}$. Vypište všechna slova ve zřetězení jazyků $L_1 = \{\varepsilon, abb, bba\}$ a $L_2 = \{a, b, abba\}$.

Řešení: $L_1 \cdot L_2 = \{a, b, abba, abbb, abbabba, bbaa, bbab, bbaabba\}$

Příklad 5: Uvažujme jazyky nad abecedou $\{0, 1\}$. Vypište všechna slova ve zřetězení

$$\{0, 001, 111\} \cdot \{\varepsilon, 01, 0101\}$$

Řešení: $\{0, 001, 00101, 0010101, 111, 11101, 1110101\}$

Příklad 6: Uvažujme jazyky nad abecedou $\{0, 1\}$. Popište (slovně) jazyk vzniklý iterací $\{00, 111\}^*$ a vyjmenujte prvních 10 slov z tohoto jazyka.

Řešení: Je to jazyk všech těch slov, která mají úseky nul sudé délky a úseky jedniček délky dělitelné třemi.

Prvních deset slov je: $\varepsilon, 00, 111, 0000, 00111, 11100, 000000, 111111, 0000111, 0011100$

Příklad 7: Uvažujme následující jazyky:

$$L_1 = \{w \in \{0, 1\}^* \mid |w|_1 \leq 1\}$$

$$L_2 = \{w \in \{0, 1\}^* \mid w = w^R\}$$

Popište, jak vypadají slova v jazyce $L_1 \cap L_2$.

Řešení: Slova ze samých nul nebo ta slova, která mají jediný znak 1 právě uprostřed, tj. $\varepsilon, 0, 00, 000, \dots, 1, 010, 00100, \dots$

Příklad 8: Napište regulární výrazy pro následující jazyky:

a) Jazyk $\{ab, ba, abb, bab, abbb, babb\}$

Řešení: $ab + ba + abb + bab + abbb + babb$ nebo $(ab + ba)(\varepsilon + b + bb)$

b) Jazyk nad abecedou $\{a, b, c\}$ obsahující právě ta slova, která obsahují podslovo abb .

Řešení: $(a + b + c)^*abb(a + b + c)^*$

c) Jazyk nad abecedou $\{a, b, c\}$ obsahující právě ta slova, která začínají prefixem bca nebo končí sufixem $ccab$.

Řešení: $bca(a + b + c)^* + (a + b + c)^*ccab$

d) Jazyk $\{w \in \{0, 1\}^* \mid |w|_0 \bmod 2 = 0\}$.

Řešení: $1^*(01^*01^*)^*$

e) Jazyk $\{w \in \{0, 1\}^* \mid |w|_0 \bmod 3 = 1\}$.

Řešení: $1^*01^*(01^*01^*01^*)^*$

f) Jazyk $\{w \in \{0, 1\}^* \mid w \text{ obsahuje podslova } 010 \text{ a } 111\}$

Řešení: $(0 + 1)^*010(0 + 1)^*111(0 + 1)^* + (0 + 1)^*111(0 + 1)^*010(0 + 1)^*$

g) Jazyk $\{w \in \{a, b\}^* \mid w \text{ obsahuje podslovo } bab \text{ nebo } |w|_b \leq 3\}$

Řešení: $(a + b)^*bab(a + b)^* + a^*(ba^* + \varepsilon)(ba^* + \varepsilon)(ba^* + \varepsilon)$

h) Jazyk $\{w \in \{a, b\}^* \mid w \text{ obsahuje podslovo } bab \text{ a } |w|_b \leq 3\}$

Řešení: $a^*ba^*baba^* + a^*baba^*ba^* + a^*baba^*$ nebo $(\varepsilon + a^*b)a^*baba^* + a^*baba^*ba^*$

i) Jazyk všech slov nad abecedou $\{a, b, c\}$, ve kterých se nikde nevyskytují dva znaky a hned za sebou.

Řešení: $((b + c + a(b + c))^*(\varepsilon + a))$

Příklad 9: Mějme dva jazyky L_1 a L_2 popsané regulárními výrazy

$$L_1 = [0^*1^*0^*1^*0^*], \quad L_2 = [(01 + 10)^*].$$

a) Jaké je nejkratší a nejdelší slovo v průniku $L_1 \cap L_2$?

Řešení: Nejkratší je ε a nejdelší 01100110 , neboť jazyk L_2 nedovoluje opakovat stejný znak za sebou více než dvakrát.

b) Proč žádný z těchto jazyků L_1 a L_2 není podmnožinou toho druhého?

Řešení: Protože $1 \in L_1 - L_2$ a $010101 \in L_2 - L_1$.

c) Jaké je nejkratší slovo, které nepatří do sjednocení $L_1 \cup L_2$? Je to jednoznačné?

Řešení: 10101 , jednoznačně.

Příklad 10: Řekněme, že bychom chtěli navrhnout syntaxi pro zápis jednoduchých aritmetických výrazů pomocí slov nad abecedou

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

- a) Navrhněte, jak budou vypadat identifikátory, a popište to pomocí regulárního výrazu.
- b) 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, oktalové, binární).