Σ - konečná abeceda

P.L. pro reg. jazyky

je regulární

$$\forall L \subseteq \Sigma^* : \underbrace{L \in \mathcal{L}_3}_{A} \Rightarrow \underbrace{\exists k > 0 : \forall w \in L \wedge |w| \geq k : \exists x,y,z \in \Sigma^* \wedge y \neq \varepsilon \wedge |xy| \leq k : \forall i \geq 0 : xy^iz \in L}_{B}$$

$A \Rightarrow B \Leftrightarrow \neg B \Rightarrow \neg A$ ⇝ použití P.L. k důkazu neregularity daného jazyka

$$\boxed{\neg B : \forall k>0 : \exists w \in L \wedge |w| \geq k : \forall x,y,z \in \Sigma^* \wedge y \neq \varepsilon \wedge |xy| \leq k : \exists i \geq 0 : xy^iz \notin L}$$

Důkaz $\neg B$ lze vidět jako hru 2 hráčů - my hrajeme $\exists$

Př.

$L = \{ w \in \{a,b\}^* \mid \#_a(w) = \#_b(w) \}$

- Nechť $k > 0$
- Volíme slovo $w = a^k b^k$ → $w \in L \wedge |w| \geq k$ pro libovolné k
- Uvážíme všechna rozdělení (použije jejich symbolickou reprezentaci)

$$\boxed{\begin{array}{ll} x = a^m & m \geq 0 \\ y = a^l & \boxed{l > 0} \\ z = a^{k-l-m} b^k & \end{array}}$$

→ popisuje všechna validní rozdělení

Volíme $i=0$ (zde je možné vše kromě $i=1$) → $l > 0$

Pak $xy^iz = a^m (a^l)^0 a^{k-l-m} b^k = a^{k-l} b^k \notin L$ □

~~[illegible]~~ Ukážeme si nějaké špatné ~~[illegible]~~ volby slova (2)

1) $w = a^5 b^5$ ⇝ neplatí, že $|w| \geq k$ pro lib. k

2) $w = aba^k b^k$ ⇝ existuje rozdělení, kde důkaz neprojde, tj. nemůžeme 'vypumpovat' slovo z jazyka

$x = \varepsilon$
$y = ab$ ⇝ pro lib. i
$z = a^k b^k$ $\quad xy^iz \in L$

$$L = \{ a^n w \mid n \geq 0 \wedge w \in \{b,c\}^* \wedge \#_b(w) = \#_c(w) \}$$

Nechť $k > 0$

Volím $w = b^k c^k$ ⇝ zbytek důkazu identický k předchozímu příkladu

Ukážeme si špatnou volbu např. $w = a^k b^k c^k$ ⇝ pak pro rozdělení např.

$x = \varepsilon$
$y = a$
$z = a^{k-1} b^k c^k$

nemůžeme slovo vypumpovat

podobně pro slovo $w = ab^k c^k$

Uvažme variantu, kde $n > 0$ → volba $w = b^k c^k$ je špatná

Musíme volit $w = ab^k c^k$

Všechna rozdělíme do dvou případů

1. $x = \varepsilon$
   $y = ab^m$ $\quad m \geq 0$
   $z = b^{k-m} c^k$

pak pro $i = 0$ dostáváme $xy^iz = b^{k-m} c^k \notin L$ (neobsahuje a)

(3)

2. $x = ab^m$ $m \geq 0$ $m + l < k$
$y = b^l$ $l > 0$
$z = b^{k-l-m} c^k$

pro $i = 0$ dostáváme $l > 0$
$xy^iz = ab^{k-l}c^k \notin L$

Pro některé jazyky je použití P.L. dost náročné např.

$L = \{ w \in \{a,b\}^* \mid \#_a(w) \neq \#_b(w) \}$

Nechť $k > 0$

Zkusme zvolit $w = a^k b^{k+1}$

Uvažme všechna rozdělení kde
$x = a^m$ $m \geq 0$
$y = a^l$ $l > 0$
$z = a^{k-l-m} b^k$

Pak potřebujeme $i \geq 0$
t.ž. $a^{k+mi} b^{k+1} \in L$

tudíž $k + m \cdot i = k + 1$

pro lib m
vidíme, že w je
špatně zvolené

Místo hledání vhodného w (doporučuji z DÚ)
použijeme M.N. věty

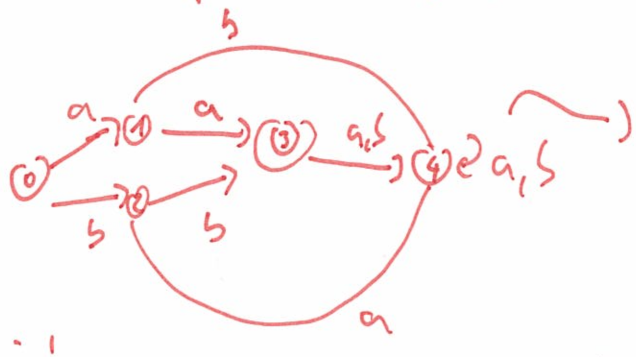
M.N. → opakování teorie viz poznámky od prof. Vojnara

Připomenutí, že relace $\sim$ je nezávislá na L, ale $\sim_L$ je daná (jednoznačná) pro uvažovaný jazyk L

Dále jsme se bavili o vztahu $\sim$ ($\sim_L$) a slovech pod kterými nějaký automat přejde s počátečního do koncového stavu

Příklad

$L = \{aa, bb\}$



je minimální pro L

$\sim = \sim_L$

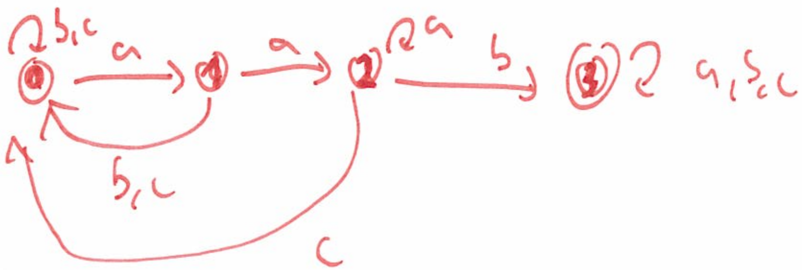
- třídy rozkladu odpovídají jazykům pro jednotlivé stavy

$L^{-1}(q_0) = \varepsilon$ $\quad L^{-1}(q_1) = \{a\}$ $\quad L^{-1}(q_2) = \{b\}$ $\quad L^{-1}(q_3) = \{aa, bb\} = L$

$L^{-1}(q_4) = \Sigma^* \setminus (L \cup L^{-1}(q_0) \cup L^{-1}(q_1) \cup L^{-1}(q_2))$

Vidíme, že index $\sim_L$ ($|\sim_L|$) je 4 tudíž L je reg.

$L = \{ w \in \{a,b,c\}^* \mid$ ~~ah~~ w obsahuje podslovo 'aab' $\}$



$L^{-1}(q_0)$ = slova, která nepatří do L a nekončí na 'a'

$L^{-1}(q_1)$ = —— " —— a nekončí na 'aa' a končí na 'a'

$L^{-1}(q_2)$ = —— " —— a končí na aa

$L^{-1}(q_3) = L$

$L = \{ w \in \{a,b\}^* \mid \#_a(w) \geq 2 \}$

$u \sim_L v \iff \#_a(u) = \#_a(v) \lor (\#_a(u) \geq 2 \land \#_a(v) \geq 2)$

přímo definujeme $\sim_L$

Použití M.N. k důkazu neregularity

Připomeňme jazyk

$$L = \{ w \in \{a,b\}^* \mid \#_a(w) \neq \#_b(w) \}$$

Předpokládejme, že L je regulární. Pak dle M.N. relace $\sim_L$ má konečný index tj. $\exists n \in \mathbb{N} : |\sim_L| = n$

Zvolíme (vhodně) $n+1$ slov

$$a^0, a^1, a^2, \dots, a^n$$

Dle předpokladu ($|\sim_L| = n$) aspoň dvě slova musí být v relaci.

Uvažme libovolnou dvojici slov u, v: $u = a^i$, $v = a^j$, $i \neq j$

když teď „připojím" $w = b^i$ dostávám

$u.w = a^i b^i \notin L$

$v.w = a^j b^i \in L$ $\to$ tudíž $u \not\sim_L v$. Vzhledem k tomu, že to platí pro lib. a^i, a^j dostáváme, že $|\sim_L| > n$

spor $\Rightarrow$ L není reg.

Na závěr si ukážeme špatnou volbu $n+1$ slov:

Např.

$ab^1, a^2b^2, a^3b^3, \dots, a^{n+1}b^{n+1}$

$(i \neq j)$

když teď vezmu li b $u = a^i b^i$ a přiřetězím například doplněk

$v = a^j b^j$

mají stejný počet 'a' a 'b'

$uw = a^i b^i b^i = a^i b^{2i} \in L$

$vw = a^j b^j b^i = a^j b^{i+j} \in L$

---

* Zkusme najít $\sim_L$ pro náš L * [náročnější]

$$u \sim_L v \iff \#_a - \#_b(v) = \#_a - \#_b(u)$$

tj. 'u' a 'v' mají stejný rozdíl mezi počty 'a' a 'b'

je vidět, že $|\sim_L| = \infty$ → což jsme ukázali výše