Pumping Lemma

Důkazy Matematickou indukcí (10.

- $\overline{L}$ je doplněk jazyka

Př $L = \{w \mid w$ obs. "aba" a neobs "aa"$\}$

$\overline{L} = \{w \mid w$ neobs "aba" a obs. "aa"$\}$

- regulární jazyk – Jazyk L, pro který existuje DFA M, že $L = L(M)$
- Nerodova věta (důkaz, že jazyk ~~je~~ není regulární

L je regulární iff existuje jedna ekvivalence T na Σ^*, že

1) L je sjednocení některých tříd T
2) Pokud uTv, tak $uwTvw$ $\forall w \in \Sigma^*$
3) T má konečný počet tříd ekvivalence

Př $L = \{0^k 1^j 0;\ 0 \leq k \leq j\}$ $\Sigma = \{0,1\}$

Zvolíme nekonečnou posloupnost slov

$0^1, 0^2, 0^3, \ldots, 0^k, \ldots$

z 3) ex. $r < s$, že $0^r T 0^s$. Položíme $w = 1^r$

Pak $0^r 1^r T 0^s 1^r$, ale $0^s 1^r 0 \notin L$, to je spor s podmínkou 1)

$0^r 1^r 0 \in L$

L není sjednocení někt tříd.

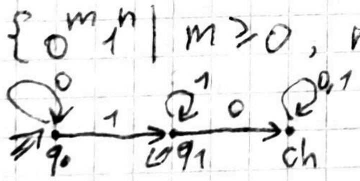
L není regulární

- Nerodova věta (důkaz, že jazyk je regulární)

Použijeme invarianty, t.j. každému stavu přiřadíme popis vš. slov u, pro které $\delta^*(q_0, u) = q$ a to tak, že

1. ε splňuje invariant pro q_0
2. každé u splňuje právě jeden invariant
3. $\forall q \in Q$ s $a \in \Sigma$ platí: Jestliže u odpovídá invariantu q, pak ua odpovídá invariantu $\delta(q, a)$
4. $u \in L$ iff u splňuje invariant pro některý $q \in F$

Př $L = \{0^m 1^n \mid m \geq 0,\ n \geq 1\}$



$\delta^*(q_0, u) = q_0$ 1. třída … $\{\varepsilon, 0, .. 0^n\}$ samé nuly

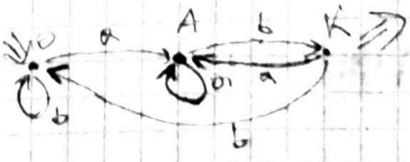
$\delta^*(q_0, u) = q_1$ 2. třída … začínající 0 a konč 1.

$\delta^*(q_0, u) = ch$ … za nulami a aspoň 1 jedničkou má ještě 0.

- Redukce DFA

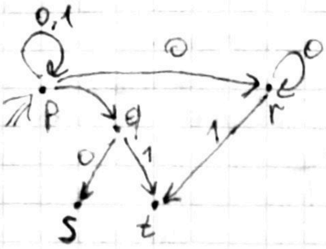
	a	b	V_0	a	b	V_1	a	b	V_2
⇒ 1	2	1	O	O	O	O	A	O	O
2	2	3	O	O	K	A	A	K	A
⇐ 3	4	5	K	O	O	K	A	O	K
4	4	3	O	O	K	A	A	K	A
5	4	5	O	O	O	O	A	O	O

$V_2 = V_1 = V$

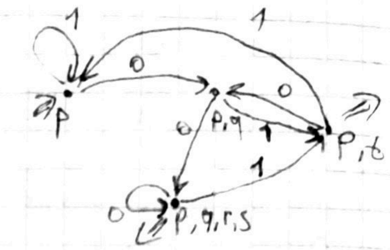


- NFA → DFA   DFA má pouze jeden vstupní stav

	0	1
⇒ p	p,q	p
q	r,s	t
r	p,r	t
⇐ s	∅	∅
⇐ t	∅	∅



	0	1
⇒ p	p,q	p
p,q	p,q,r,s	p,t
⇐ p,q,r,s	p,q,r,s	p,t
⇐ p,t	p,q	p



Když NFA má několik vs. stavů, děláme z nich mn. a přidáváme je do tabulky.

- NFA s ε-přechody → DFA

	ε	a	b	c
⇒ p	∅	p	q	r
q	p	q	r	∅
⇐ r	q	r	∅	p

ε-UZ(p) = {p}
ε-UZ(q) = {q, p}
ε-UZ(r) = {p, q, r}

→

	a	b	c
p	p	p,q	p,q,r
p,q	p,q	p,q,r	p,q,r
p,q,r	p,q,r	p,q,r	p,q,r

- Každý jazyk reprezentovaný regulárním výrazem je reg.
- $r = (a+b)$, "+" je alternative
  $r = a$ nebo $r = b$

- reg. výraz → DFA
  1) Očíslování
  2) Čím slovo začíná
  3) Čím slovo končí
  4) Která písmena mohou být
  5) ε je slovem v L?
  6) Zkonstruovat NFA
  7) NFA → DFA

Př $(a_1 a_2 + b_3)^* (b_4 a_5)^*$ **Začátek S**

zač: a_1, b_3, b_4 ($=S$)   končí: a_2, b_3, a_5

4) $a_1: a_2$; $a_2: b_3, a_1, b_4$; $b_3: a_1, b_3, b_4$
$b_4: a_5$; $a_5: b_4$

5) $\varepsilon \in L$

6)

	a	b
S	a_1	b_3, b_4
a_1	a_2	∅
a_2	a_1	b_3, b_4
b_3	a_1	b_3, b_4
b_4	a_5	∅
a_5	∅	b_4

← poč. stav **vždy**

- DFA → reg výraz. $S \xrightarrow{\varepsilon} q_0 \dots q_f \xrightarrow{\varepsilon} F$
  1) Ke st. diag. přidáme nový počát. a konc. stav
  2) a) Odstranění paralelních hran
     b) Odstranění smyčky
     c) Odstranění vrcholu

- Kontextové gramatiky: pravidly tvaru $\boxed{\alpha A \beta \to \alpha \gamma \beta}$ a $\boxed{S \to \varepsilon}$ S nemůže být na pravé straně
- Bezkontextové gramatiky (CF): pravidly tvaru $A \to \gamma$, kde $\gamma \in (N \cup \Sigma)^*$
- Regulární gramatiky: pravidly tvaru $A \to wB$, $A \to w$, kde A netermin., w term. slov
- Nevypouštěcí gramatiky – neobsahuje žádné pravidlo $A \to \varepsilon$
- Jednoznačná gramatika – pro každé slovo w existuje jen jeden derivační strom.
- Víceznačná gramatika – existuje slovo w, pro které je několik derivačních stromů.

- Gramatika → Nevypouštěcí gramatika
  1) Uděláme mn s neterminály, kde jsou pravidla $A \to \varepsilon$
  2) Zvětšujeme mn neterminály ve pravidlech kterých jsou prvky množiny
  3) V každém pravidlu postupně odstraňujeme neterminály z mn.

Př $S \to$ [redacted] $aSAbS$
$A \to bAc \mid \varepsilon$

1) $U_1 = \{A\} - \{X \mid X \to \varepsilon \in P\}$
2) $U_{i+1} = U_i \cup \{Y \mid Y \to N, N \in U_i\}$
   $U_2 = \{A, S\} = U$

3) $S \to aSAbS \mid aAbS \mid aSbS \mid aSAb \mid abS \mid aAb \mid aSb \mid ab$
   $A \to bAc \mid bc$

- NFA/DFA → Gramatika
  1) Napíšeme $S \to V_1 \mid V_2 \dots$, kde V jsou vstupní stavy
  2) Před každým terminálem píšeme neterminál

Př

	a	b
⇐⇒ A	A,C	B
B	∅	B,D
⇐ C	∅	∅
⇒ D	A	C,D

P: $S \to A \mid D$
$A \to aA \mid aC \mid bB \mid \varepsilon$
$B \to bB \mid bD$
$C \to \varepsilon$
$D \to aA \mid bD \mid bC$

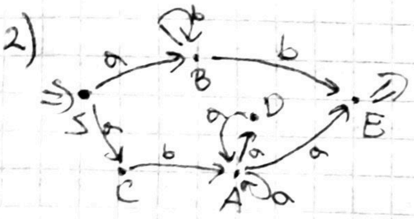
• Gramatika → NFA

1) Převedeme gramatiku k tvaru $N_1 \to tN_2 | \dots$
2) Uděláme NFA

Př $S \to abA | aB$
$A \to aA | aaA | a$
$B \to bB | b$

1) $S \to aC | aB$ $\quad C \to bA$
$A \to aA | aD | aE$ $\quad D \to aA$
$B \to bB | bE$ $\quad E \to \varepsilon$

2)



• Redukce gramatiky

1) Najdeme množinu neterminálů, které generujou slovo $w \in \Sigma^*$
$V_1 = \{X | X \to w \in P, w \in \Sigma^*\}$
$V_{i+1} = V_i \cup \{X | X \to \alpha \in P, \alpha \in (\Sigma \cup V_i)^*\}$
Pak vytvoříme gramatiku $G' = (V, \Sigma, S, P')$
2) Vytvoříme množinu dosažitelných terminálů (indukcí)
$U_0 = \{S\}$
$U_1 = \{X | S \to \alpha X \beta \in P'\}$ $\quad \alpha, \beta \in (\Sigma \cup V)^*$ $\quad$ Pak $G_2 = (U, \Sigma, S, P'')$
$U_{i+1} = U_i \cup \{X | Y \to \alpha X \beta \in P', Y \in U_i; \alpha, \beta \in (\Sigma \cup V)^*\}$

Př $G = (N, \Sigma, S, P)$ $\quad N = \{S, A, B, C, D\}$, $\Sigma = \{a, b\}$

$S \to SB | CD | aSBb | aSa$
$A \to aAa | BA | bS$
$B \to AB | BC | Bbb$
$C \to ACB | aC | bD$
$D \to BCD | aD | a$

a) Najdeme množinu neterminálů, které generují w
$V_1 = \{X | X \to w \in P, w \in \Sigma^*\}$
$V_{i+1} = V_i \cup \{X | X \to \alpha \in P, \alpha \in (\Sigma \cup V_i)^*\}$

$V_1 = \{D\}$, $V_2 = \{D, C\}$, $V_3 = \{D, C, S\}$
$V_4 = \{D, C, S, A\}$, $V_5 = \{D, C, S, A\} = V$
Vytvoříme gramatiku $G' = (V, \Sigma, S, P')$
P': $S \to CD | aSa$
$A \to aAa | bS$
$C \to aC | bD$
$D \to aD | a$

b) Vytvoříme mn. dosažitelných terminálů (indukcí)
$U_0 = \{S\}$
$U_1 = \{X | S \to \alpha X \beta \in P'\}$
$U_{i+1} = U_i \cup \{X | Y \to \alpha X \beta \in P'; Y \in U_i; \alpha, \beta \in (\Sigma \cup V)^*\}$

$U_0 = \{S\}$, $U_1 = \{S, C, D\} = U$. $\quad G'' = (U, \Sigma, S, P'')$
P'': $S \to CD | aSa$
$C \to aC | bD$

• Návrh gramatiky, která generuje L

1) Vygenerovat gramatiku.
Pravidlo $S \to \dots$ má radši být ve tvaru $S \to AB$ (bez neterminál… aby rozložili na části)
2) Každé slovo vygenerujú (popsat derivace)
3) Nevygenerujeme nic navíc (levá derivace)

Př $L = \{0^i 1^j ; 0 \le i \le j\}$
1) $S \to AB$
$A \to 0A1 | \varepsilon$
$B \to 1B | \varepsilon$
2) Každé slovo vygeneruju

$S \Rightarrow AB \xRightarrow{A \to 0A1\,(i)} 0^i A 1^i B \xRightarrow{A \to \varepsilon} 0^i 1^i B$
$\xRightarrow{B \to 1B\,(j-i)} 0^i 1^i 1^{(j-i)} B \xRightarrow{B \to \varepsilon} 0^i 1^j$
$0 \le i \le j$

3) Nevygenerujeme nic navíc.
levá derivace
$S \Rightarrow AB \Rightarrow^* u$
$A \Rightarrow^* 0^i 1^i$ $\quad i \ge 0$
$B \Rightarrow^* 1^k$ $\quad k \ge 0$
Tedy $S \Rightarrow^* u$ pro $u = 0^i 1^i 1^k = 0^i 1^{i+k}$
$i \le i+k$

• Gramatika v Chomského normálním tvaru — má pouze pravidla tvaru
$A \to BC$, $A \to a$ $\quad A, B, C \in N$, $a \in \Sigma$

• Převod do Chomského normálního tvaru

1) Gramatika → nevypouštějící gramatika
2) Odstraníme pravidla $A \to B$ ($A \to B, B \to \alpha \Rightarrow A \to \alpha$)
3) Odstraníme terminály $a \in \Sigma$, které se vyskytují na pravé straně pravidla délky aspoň 2. ($X \to aY, Z \to a \Rightarrow X \to ZY, Z \to a$)
4) Zkrátíme pravé strany pravidel, které jsou delší než 2. ($X \to YZT \Rightarrow X \to YX_1, X_1 \to ZT$)

Př $G = (N, \Sigma, S, P)$ $\quad N = \{S, A, C\}$, $\Sigma = \{a, b\}$
P: $S \to SCA | a$
$A \to aCb | \varepsilon$
$C \to AA | b$

! redukce na závěr!

1) P: $S \to SCA | SA | SC | a$
$A \to aCb | ab$
$C \to AA | A | b$

2) P: $S \to SCA | SA | SC | a$
$A \to aCb | ab$
$C \to AA | aCb | ab | b$

spojení pravidel
$S \to AB | CD$
$C \to a | S$
↓
$S \to AB | CD$
$C \to AB | CD$

3) P: $X \to a$, $Y \to b$
$S \to SCA | SA | SC | a$
$A \to XCY | XY$
$C \to AA | XCY | XY | b$

4) P: $X \to a$ $\quad Y \to b$
$D \to SC$ $\quad Z \to XC$
$S \to DA | SA | SC | a$
$A \to ZY | XY$

• CYK (slovo w generováno G?)

Př: P: $S \to AB | CS | AD$ $\quad C \to DS | SC | a$

$A \to AC | AD | a$ $\quad D \to BA | b$

$B \to BC | b$

$w = abbaa$

1) Přepsat pravidla ve tv. $BC \leftarrow A$

$AB \leftarrow S$ $\quad BA \leftarrow D$ $\quad DS \leftarrow C$

$AC \leftarrow A$ $\quad BC \leftarrow B$ $\quad SC \leftarrow C$

$AD \leftarrow S, A$ $\quad CS \leftarrow S$

2) Přepsat pravidla neterm. $a \leftarrow B$

$a \leftarrow A, C$ $\quad b \leftarrow B, D$

3) Udělat tabulku

5	S A				
4	S A C	∅			
3	S A	∅	D B		
2	S A	∅	D B	A	
1	A C	B D	B D	A C	A C
	1 a	2 b	3 b	4 a	5 a

Chomského normální tvar!

4) Na konci CYK máme $S \in X_{1,5}$, takže touto gramatikou jde vygenerovat slovo $w = abbaa$

• Jazyk přijímaný koncovým stavem: úspěch, když po přečtení slova dostaneme do koncového stavu $\boxed{L(A)}$

• Jazyk přijímaný prázdným zásobníkem: úspěch, když zásobník je prázdný $\boxed{N(A)}$

• Pro každé A, B platí $L(A) = N(B)$

$N(A) = L(B)$

• CF $G \to ZA$ pro $L(G)$

1) Uděláme automat

$A = (\{q\}, \Sigma, \Gamma, \delta, q, S)$

$\Gamma = \Sigma \cup \{\text{Neterminály}\}$

2) Definujeme δ pro pravidla

$\delta(q, \varepsilon, X) = \{(q, \alpha) \mid X \to \alpha \in P\}$

3) Definujeme δ pro terminály

$\delta(q, a, a) = (q, \varepsilon) \quad \forall a \in \Sigma$

Př

$S \to aAc$ $\quad A \to aAc | B$

$B \to bBc | B$

1) $A = (\{q\}, \{a, b\}, \{a, b, S, A, B\}, \delta, q, S)$

2) $\delta(q, \varepsilon, X) = \{(q, \alpha) \mid X \to \alpha \in P\}$

$\delta(q, \varepsilon, S) = \{(q, aAc)\}$

$\delta(q, \varepsilon, A) = \{(q, aAc), (q, B)\}$

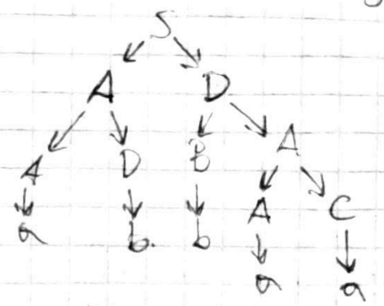
$\delta(q, \varepsilon, B) = \{(q, bBc), (q, bc)\}$

3) $\delta(q, a, a) = (q, \varepsilon) \quad \forall a \in \Sigma$

$\delta(q, a, a) = (q, \varepsilon)$

$\delta(q, b, b) = (q, \varepsilon)$ $\quad$ [S v zásobníku

• Derivační strom (z tabulky CYK zadaným vlevo)



• Zásobníkový automaty

$ZA = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$

stavy; počát. stav; začáteční zás. symbol; symboly ve zásobníku

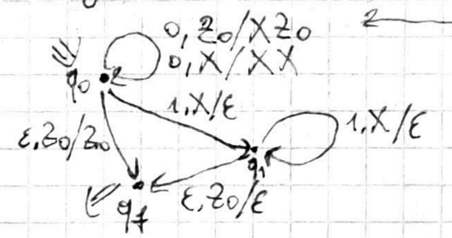
Př NFZA

	$(0, Z_0)$	$(0, X)$	$(1, Z_0)$	$(1, X)$	(ε, Z_0)	(ε, X)
$\Rightarrow q_0$	(q_0, XZ_0)	(q_0, XX)	–	(q_1, ε)	(q_f, Z_0)	–
q_1	–		–	(q_1, ε)	(q_f, ε)	–
$\Leftarrow q_f$	–		–	–	–	–

Např. jsme ve stavu q_0 a na zásobníku máme Z_0. Jestli máme na vstupu 0, tak máme zvolit $(0, Z_0) \to (q_0, XZ_0)$

stav, do kterého jdeme; jak se změní zásobník.

• ε ve změně zásobníku znamená, že mažeme symbol ze zásobníku.



Z_0 / XZ_0

starý zásobník; nový zásobník (update)

Př práce ZA

a) $w = 0010$

$(q_0, 0010, Z_0) \vdash (q_0, 010, XZ_0) \vdash (q_0, 10, XXZ_0)$

$\vdash (q_1, 0, XZ_0)$ neúspěch

b) $w = 0011$

$(q_0, 0011, Z_0) \vdash (q_0, 011, XZ_0) \vdash (q_0, 11, XXZ_0)$

$\vdash (q_1, 1, XZ_0) \vdash (q_1, \varepsilon, Z_0) \vdash (q_f, \varepsilon, \varepsilon)$ ✓

• Vytvoření ZA pro jazyk $L = \{\dots\}$

I. 1) Udělat CF gramatiku pro jazyk L

2) Ověřit jestli gramatika L generuje (vygenerujeme každé slovo. Nevygen. nic. navíc)

3) Udělat ZA z CF gramatiky.

Př Jazyky typu $L = \{0^n 1^m \mid 0 \leq n \leq m \leq 2n\}$

II 1) Udělat diagram pro jazyk a popsat každý stav.

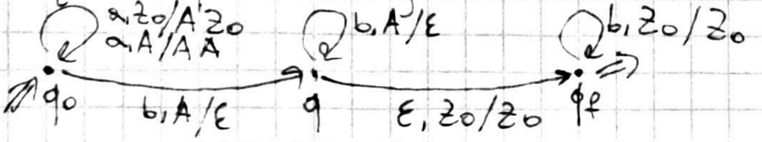
Př Jazyky typu $L = \{(ab)^i b^j a^{j-i} \mid 0 < i < j\}$

$j - i = k \geq 0$ $\quad j = i + k$

$L = \{(ab)^i b^{i+k} a^k \mid i > 0, k > 0\}$

- Deterministický ZA $A = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$
  - $\forall q \in Q,\ a \in \Sigma \cup \{\varepsilon\}\ X \in \Gamma$ je $|\delta(q, a, X)| \leq 1$.
    t.j. z každého stavu maximálně jednu operaci (a, X)
  - if $\delta(q, \varepsilon, X) \neq \emptyset$, then $\delta(q, a, X) = \emptyset\ \forall a \in \Sigma$
    t.j. jestliže je operace (ε, X) (update jen zásobníku), pak operace s dalšími symboly (a stejným updatem zásobníku) nejsou.

Př $L = \{a^n b^m \mid 0 < n \leq m\}$



- Deterministický jazyk – je přijíman DZA koncovým stavem
- Bezprefixový jazyk – je přijíman DZA prázdným zásobníkem.

- Odstranění levé rekurze
1) Najdeme neterminál ve pravidlech kterého je levá rekurze $N \to NAB \mid C$
2) Pro pravidla s levou rekurzí uděláme pravidla typu $Z_N \to AB \mid ABZ_N$
3) Zduplikujeme pravidla tohoto neterm. a přidáme na konec Z_N
   $N \to C \mid CZ_N$
4) Na konci máme
   $N \to C \mid CZ_N$
   $Z_N \to AB \mid ABZ_N$

Př
$A \to ACD \mid BaD \mid aD$
mají levou rek: $A \to ACD$
nemají levou rek: $A \to BaD \mid aD$
pak
$A \to BaD \mid aD \mid BaDZ_A \mid aDZ_A$
$Z_A \to CD \mid CDZ_A$

- Pumping lemma pro CF jazyky
  Pro každý CF jazyk L ex. kladné přirozené číslo n takové, že pro každé slovo $z \in L$ délky více než n jde najít rozklad $z = uvwxy$, že
  1) $|vwx| \leq n$
  2) $vx \neq \varepsilon$
  3) $\forall i \geq 0 \quad uv^iwx^iy \in L$

- Důkaz matematickou indukcí

Př $G = (N, \Sigma, S, P)$
P:
$S \to BSA \mid b \qquad C \to cA \mid \varepsilon$
$A \to CB \mid AB \mid Aa \mid c \qquad D \to ABC \mid SA \mid a$
$B \to Ab \mid bB \mid \varepsilon$

Dokázat $A \Rightarrow^*_G c^i A (b^2 a)^i \quad \forall i \geq 0$

$i = 0$: $c^i A (b^2 a)^i = A$
$A \Rightarrow A$ platí

$i \geq 0$: Indukční předpoklad
IP: $A \Rightarrow^* c^i A (b^2 a)^i$

$A \overset{IP}{\Rightarrow^*} c^i A (b^2 a)^i \overset{A \to Aa}{\Longrightarrow} c^i Aa (b^2 a)^i \Rightarrow$
$\overset{A \to CB}{\Longrightarrow} c^i CBa (b^2 a)^i \overset{C \to cA}{\Longrightarrow}$
$\Longrightarrow c^{i+1} ABa (b^2 a)^i \overset{B \to bB}{\Longrightarrow} c^{i+1} AbBa (b^2 a)^i \Rightarrow$
$\Longrightarrow c^{i+1} Ab^2 Ba (b^2 a)^i \overset{B \to \varepsilon}{\Longrightarrow} c^{i+1} A (b^2 a)^{i+1}$