Automaty a formálne jazyky

Podľa prednášok prof. RNDr. Viliama Gefferta, DrSc., PrírF UPJŠ
Dňa 8. februára 2005 zostavil Róbert Novotný, r.novotny@szm.sk.
Typeset by LaTeX. Illustrations by jPicEdit.

Úvodné definície

Definícia. *Abeceda*: ľubovoľná konečná množina (zvyčajne neprázdna).
Symboly: prvky abecedy
Tranzitívny uzáver Σ^*: množina všetkých postupností symbolov abecedy konečnej dĺžky. Ak A je abeceda, tak A^* je množina všetkých slov v abecede A.
Slová (reťazce): prvky tranzitívneho uzáveru nad danou abecedou Σ, resp. ľubovoľná konečná postupnosť prvkov abecedy.
Dĺžka slova: počet symbolov, z ktorých sa slovo skladá. Ak w je slovo, $w = a_1a_2 \dots a_n$, potom $|w| = n$.
Operácia $\cdot$ *(zreťazenie*: Nech $\alpha = a_1 \dots a_n$, $\beta = b_1 \dots b_n$. Potom $\alpha \cdot \beta = a_1 \dots a_n b_1 b_n$
Niekedy sa $\cdot$ vynecháva – píšeme $\alpha\beta$.
Platí: $|\alpha\beta| = |\alpha| + |\beta|$, $\alpha \cdot \varepsilon = \varepsilon \cdot \alpha = \alpha$
Jazyk L zo slov danej abecedy Σ je ľubovoľná podmnožina Σ.

$$L \subseteq \Sigma$$

Definícia. Gramatika $G = (N, T, P, S)$ - usporiadaná štvorica, kde:

N	sú neterminály (konečná abeceda).
T	sú terminály (konečná abeceda).
$S \in N$	je počiatočný neterminál, resp. štartovací symbol.
P	je konečná množina gramatických pravidiel, kde $P \subseteq (N \cup T)^* \times (N \cup T)^*$.

Zároveň musí platiť, že: $T \cap N = \emptyset$.

Poznámka. Prvky P by sa mali písať ako usporiadané dvojice $[S, AB]$, $[A, aA] \dots$, ale kvôli názornosti píšeme $S \to AB, A \to aA, \dots$

Definícia. Hovoríme, že reťazec u je *priamo odvoditeľný* z reťazca t v gramatike G, ak sa t, u dajú vyjadriť v tvare

$$\begin{aligned} t &= x\alpha y \\ u &= x\beta y, \end{aligned}$$

kde $x, y, \alpha, \beta \in (NST)^*$, pričom $\alpha \to \beta \in P$.

$$\text{označenie:} \quad t \Rightarrow_G u$$

Definícia. Nech $u, v \in (T \cup N)^*$. Hovoríme, že reťazec u je odvoditeľný z reťazca t v gramatike G, ak existujú také $v_0, v_1, \dots, v_n$, že platí:

$$t = v_0 \Rightarrow_G v_1 \Rightarrow_G \dots \Rightarrow_G v_n = u$$

$$\text{označenie:} \quad t \Rightarrow_G^* u$$

Poznámka. „Priamo odvoditeľný“ znamená vlastne „odvoditeľný na jeden krok“.

Tvrdenie (bez dôkazu). *Nech G je gramatika. Platí:*

1. $\forall w \in (T \cup N)^* : w \Rightarrow_G^* w$
2. $\forall w_1, w_2, w_3 \in (T \cup N)^* : (w_1 \Rightarrow_G^* w_2 \wedge w_2 \Rightarrow_G^* w_3) \longrightarrow (w_1 \Rightarrow_G^* w_3)$

Definícia. Jazyk generovaný gramatikou G je

$$L(G) = \{w \in T^*; S \Rightarrow_G^* w\}$$

Chomského hierarchia gramatík

0. Frázové gramatiky

- pravidlá bez obmedzenia – všeobecné pravidlá $\alpha \to \beta$
- povoľujeme aj $|\alpha| > |\beta|$
- frázovými sa nazývajú preto, lebo povoľujú aj skracovanie: DO NOT $\to$ DON'T
- dá sa simulovať ľubovoľný výpočet ľubovoľného algoritmu s ľubovoľným vstupom (Turingov stroj)

1. Kontextové gramatiky

- ak $\alpha \to \beta \in P$ tak $|\alpha| \leq |\beta|$ – slovo nemožno skracovať. (Výnimka je pravidlo $S \to \varepsilon$, za predpokladu, že S sa nenachádza na pravej strane žiadneho pravidla).
- kontextové sú preto, lebo význam môže závisieť od kontextu (možno podmieniť kontextom), napr.: $abxc \to abDDTc$
- zodpovedá im lineárne ohraničený Turingov stroj, algoritmy s lineárnou pamäťou (veľkosť pamäte je podľa vstupu, nemožno prideliť ďalšiu)

2. Bezkontextové gramatiky

- iba pravidlá typu $A \to \alpha$, kde $A \in N, \alpha \in (N \cup T)^*$. Čiže na ľavej strane pravidla môže byť len jeden neterminálový znak.
- význam neterminálov nezávisí od kontextu
- sú najlepšie preskúmané a najdôležitejšie – programovacie jazyky sú zapísané pomocou bezkontextových gramatík
- zodpovedajú im zariadenia s konečnou pamäťou, zásobníkové automaty (možno používať rekurziu)

3. Regulárne gramatiky

- iba pravidlá typu $A \to aB$, alebo $A \to a$, kde $A, B \in N, a \in T$.
- zodpovedajú im konečnostavové automaty (napr. automat na kávu)

Konečnostavové automaty

Definícia. Konečnostavový automat M je usporiadaná šestica $M = (Q, \Sigma, \Delta, f, g, q_I)$, kde:

Q	konečná množina stavov
Σ	vstupná abeceda
Δ	výstupná abeceda
$f : Q \times \Sigma \to Q$	prechodová funkcia
$g : Q \times \Sigma \to \Delta$	výstupná funkcia
$q_I \in Q$	počiatočný stav

Poznámka. Pre funkcie f, g teda platí:

$$f(q, a) = q' \qquad g(q, a) = b$$

Definícia. Matematickou indukciou si definujeme nasledujúce funkcie:

$$\begin{aligned} f^* &: Q \times \Sigma^* \to Q \\ g^* &: Q \times \Sigma^* \to \Delta^* \end{aligned}$$

$$f^*(q, \varepsilon) = q \qquad \forall q \in Q \tag{1}$$

$$f^*(q, \alpha a) = f(f^*(q, \alpha), a), \quad \forall q \in Q, \forall \alpha \in \Sigma^*, \forall a \in \Sigma \tag{2}$$

$$g^*(q, \varepsilon) = \varepsilon \qquad \forall q \in Q \tag{3}$$

$$g^*(q, \alpha a) = g^*(q, \alpha) \cdot g(f^*(q, \alpha), a), \quad \forall q \in Q, \forall \alpha \in \Sigma^*, \forall a \in \Sigma \tag{4}$$

Poznámka. Na pravej strane rovnosti (3) je medzi funkciami $g^*(q, \alpha)$ a $g(f^*(q, \alpha), a)$ operácia konkatenácie, nie násobenia!

Lema. *Pre každé $q \in Q$ a pre každé $a \in \Sigma$ platí:*

- $f^*(q, a) = f(q, a)$
- $g^*(q, a) = g(q, a)$

Dôkaz:

$$f^*(q, a) = f^*(q, \varepsilon a) \overset{\text{z (2)}}{=\!=} f(f^*(q, \varepsilon), a) \overset{\text{z (1)}}{=\!=} f(q, a)$$

$$g^*(q, a) = g^*(q, \varepsilon a) \overset{\text{z (4)}}{=\!=} g^*(q, \varepsilon) \cdot g(f^*(q, \varepsilon), a) = \overset{\text{z (4)}}{=\!=} \varepsilon \cdot g(q, a) = g(q, a)$$

Lema. *Pre každé $q \in Q$, pre každé $\alpha, \beta \in \Sigma^*$ platí:*

- $f^*(q, \alpha\beta) = f^*(f^*(q, \alpha), \beta)$ *(5)*
- $g^*(q, \alpha\beta) = g^*(q, \alpha) \cdot g^*(f^*(q, \alpha), \beta)$ *(6)*

Dôkaz:
Ad 1.)
Matematickou indukciou vzhľadom na dĺžku slova β:

- nech $|\beta| = 0$, teda $\beta = \varepsilon$

$$\begin{aligned} f^*(q, \alpha\beta) &= f^*(q, \alpha\varepsilon) = f^*(q, \alpha) = q' = \\ &= f^*(q', \varepsilon) = f^*(f^*(q, \alpha), \varepsilon) \overset{\varepsilon = \beta}{=\!=} \\ &\overset{\varepsilon = \beta}{=\!=} f^*(f^*(q, \alpha), \beta) \end{aligned}$$

- nech $\beta = \beta' a$, kde $a \in \Sigma$ a pre β' platí indukčný predpoklad.

$$\begin{aligned} f^*(q, \alpha\beta) &= f^*(q, \alpha\beta' a) = f(\underbrace{f^*(q, \alpha\beta')}_{\text{ind. predp.}}, a) = \\ &= f(f^*(\underbrace{f^*(q, \alpha)}_{q'}, \beta'), a) \\ &= f(f^*(q', \beta'), a) = f^*(q', \beta' a) = \\ &= f^*(f^*(q, \alpha), \beta' a) = f^*(f^*(q, \alpha), \beta), \end{aligned}$$

Ad 2.)
Opäť matematickou indukciou vzhľadom na dĺžku slova β:

- nech $|\beta| = 0$, t.j. $\beta = \varepsilon$
  $g^*(q, \alpha\beta) = g^*(q, \alpha\varepsilon) = g^*(q, \alpha) = g^*(q, \alpha) \cdot \varepsilon = g^*(q, \alpha) \cdot g^*(q', \varepsilon) = g^*(q, \alpha) \cdot g^*(f^*(q, \alpha), \underbrace{\varepsilon}_{\beta})$, pričom sme substituovali $q' = f^*(q, \alpha)$.
- nech $|\beta| > 0$, t.j. $\beta = \beta' b$, kde $b \in \Sigma$ a nech pre β' platí indukčný predpoklad.

$$\begin{aligned} g^*(q, \alpha\beta) &= g^*(q, \alpha\beta' b) \\ &= \underbrace{g^*(q, \alpha\beta')}_{\text{IP}} \cdot g(\underbrace{f^*(q, \alpha\beta')}_{\text{tvrd. pre } f}, b) \\ &= g^*(q, \alpha) \cdot g^*(\underbrace{f^*(q, \alpha)}_{q'}, \beta') \cdot g(f^*(\underbrace{f^*(q, \alpha)}_{q'}, \beta'), b) \\ &= g^*(q, \alpha) \cdot g^*(q', \beta') \cdot g(f^*(q', \beta'), b) \\ &= g^*(q, \alpha) \cdot g^*(q', \beta' b) \\ &= g^*(q, \alpha) \cdot g^*(f^*(q, \alpha), \beta), \end{aligned}$$

Definícia. Automat $M = (Q, \Sigma, \Delta, f, g, q_I)$ je *súvislý* práve vtedy, ak:

$$\forall q \in Q \ \exists w \in \Sigma^* : f^*(q_I, w) = q.$$

Definícia. Dva stavy q_1, q_2 sú *ekvivalentné* práve vtedy, ak

$$\forall w \in \Sigma^* : \ g^*(q_1, w) = g^*(q_2, w).$$

Označujeme: $q_1 \sim q_2$

Definícia. Automat M je *redukovaný* práve vtedy, ak je súvislý a nemá žiadnu dvojicu ekvivalentných stavov.

Poznámka. Vo všeobecnosti je problém overiť ekvivalenciu stavov podľa definície, bolo by treba overiť ekvivalenciu pre nekonečne veľa vstupov.

Definícia. Dva stavy q_1, q_2 sú k-ekvivalentné práve vtedy, ak:

$$\forall \alpha \in \Sigma^* : |\alpha| \leq k : g^*(q_1, \alpha) = g^*(q_2, \alpha)$$

Píšeme $q_1 \sim_k q_2$.

Tvrdenie. *Ekvivalencia ($\sim$) a k-ekvivalencia ($\sim_k$) je reláciou ekvivalencie nad množinou stavov Q. Teda platí:*

1. $q \sim q \qquad \forall q$
2. $q_1 \sim q_2 \equiv q_2 \sim q_1 \qquad \forall q_1, q_2$
3. $q_1 \sim q_2 \;\&\; q_2 \sim q_3 \Rightarrow q_1 \sim q_3 \qquad \forall q_1, q_2, q_3$

Dôkaz:
Ad 3.)

$$q_1 \sim q_2 \quad \equiv \quad \forall \alpha : g^*(q_1, \alpha) = g^*(q_2, \alpha)$$

a zároveň

$$q_2 \sim q_3 \quad \equiv \quad \forall \alpha : g^*(q_2, \alpha) = g^*(q_3, \alpha)$$

Potom z tranzitívnosti rovnosti platí:

$$\forall \alpha : g^*(q_1, \alpha) = g^*(q_3, \alpha)$$

Pri k-ekvivalencii stačí ohraničiť kvantifikátory.

Lema.

1. $\forall k \geq 1 : q_1 \sim_k q_2 \quad \Rightarrow \quad q_1 \sim_{k-1} q_2$
2. $q_1 \sim q_2 \quad \equiv \quad \forall k \geq 1 : q_1 \sim_k q_2$

Dôkaz:

- Ad 1.) Ak $g^*(q_1, \alpha) = g^*(q_2, \alpha)$ pre $|\alpha| \leq k$, tým skôr aj pre $|\alpha| \leq k - 1$.
- Ad 2.)
  $\Rightarrow$ triviálne
  $\Leftarrow$ Nech $\forall k : q_1 \sim_k$. Zoberme si ľubovoľné $\alpha \in \Sigma^*$. Označme $k_\alpha = |\alpha| \in \mathbb{N}$.
  Keďže $q_1 \sim_k q_2$ pre každé $k \in \mathbb{N}$, tak aj pre $k = k_\alpha$. Čiže platí: $q_1 \sim_{k_\alpha} q_2$, teda

  $$\forall w \in \Sigma^* : |w| \leq \alpha : g^*(q_1, w) = g^*(q_2, w).$$

  Ak to platí pre každé w také, že $|w| \leq k_\alpha$, musí to platiť aj pre $w = \alpha$, lebo $|w| \leq |\alpha|$.
  Máme teda $g^*(q_1, \alpha) = g^*(q_2, \alpha)$ pre ľubovoľné $\alpha \in \Sigma^*$, a teda $q_1 = q_2$.

Veta.

$$\begin{aligned} \forall k > 1 : q_1 \sim_k q_2 \quad \equiv \quad & q_1 \sim_{k-1} q_2 \\ & \& \\ & \forall a \in \Sigma : f(q_1, a) \sim_{k-1} f(q_2, a) \end{aligned}$$

Dôkaz:
$\Rightarrow$

1. Vlastnosť (1) získame z predchádzajúcej lemy.
2. Vieme, že platí

   $$q_1 \sim_k q_2 \equiv \forall \alpha : |\alpha| \leq k : g^*(q_1, \alpha) = a^*(q_2, \alpha)$$

   Keďže to platí $\forall k$, tak aj pre $k = 2$. Preto ak $\alpha = a\alpha'$, tak:

   $$\forall a \in \Sigma\ \forall \alpha' \in \Sigma^* : g^*(q_1, a\alpha') = g^*(q_2, a\alpha')$$

   t.j. $\forall a \in \Sigma\ \forall \alpha' \in \Sigma : |\alpha'| \leq k - 1$ platí:

   $$\begin{aligned} g^*(q_1, a)g^*(f^*(q_1, a), \alpha') &= g^*(q_2, a)g^*(f^*(q_2, a), \alpha') \\ g(q_1, a)g^*(f^*(q_1, a), \alpha') &= g(q_2, a)g^*(f^*(q_2, a), \alpha') \\ \forall a \in \Sigma : f(q_1, a) &\sim_{k-1} f(q_2, a) \end{aligned}$$

$\Leftarrow$
Zrejme platí:

$$q_1 \sim_{k-1} q_2 \Rightarrow q_1 \sim_1 q_2 \Rightarrow \forall a \in \Sigma : g^*(q_1, a) = g^*(q_2, a)$$

Ďalej, $\forall a \in \Sigma : f(q_1, a) \sim_{k-1} f(q_2, a)$. To je však ekvivalentné tomu, že $\forall \alpha' : |\alpha'| \leq k - 1, \forall a \in \Sigma :$

$$g^*(f(q_1, a), \alpha') = g^*(f(q_2, a), \alpha')$$

Platnosť rovnosti sa nezmení, ak pred výraz na každej strane predradíme 1 symbol:

$$\begin{aligned} g^*(q_1, a)g^*(f(q_1, a), \alpha') &= g^*(q_2, a)g^*(f(q_2, a), \alpha') \\ g(q_1, a)g^*(f^*(q_1, a), \alpha') &= g(q_2, a)g^*(f^*(q_2, a), \alpha') \\ g^*(q_1, a\alpha') &= g^*(q_2, a\alpha') \end{aligned}$$

Posledný riadok je ale ekvivalentný tomu, že

$$\forall \alpha : |\alpha| \leq k, \alpha \neq \varepsilon : g^*(q_1, \alpha) = g^*(q_2, \alpha)$$

Pre prípad $\alpha = \varepsilon$ to platí z definície, takže môžeme písať

$$q_1 \sim_k q_2$$

Algoritmus na konštrukciu tried ekvivalentných stavov

Krok 1. Rozklad Q na triedy 1-ekvivalentných stavov:
ak $q, q' \in A_i^{(1)}$, tak $q \sim_1 q'$.
for q, q' **for** $a \in \Sigma : g(q, a) = g(q', a), q \neq q'$. Ak vyššie uvedená rovnosť platí, vložíme stavy do tej istej množiny v rozklade.

Krok 2. Ak máme $A_1^{(k-1)}, \ldots, A_{n_k}^{(k-1)}$ rozklad, konštruujeme rozklad $A_1^{(k)}, \ldots, A_{n_k}^{(k)}$ tak, že má platiť:

$$\bigcup_{i=1}^{n_k} A_i^{(k)} = Q, \text{ pričom } A_i^{(k)} \cap A_j^{(k)} = \emptyset \text{ pre } i \neq j$$

Zároveň tento rozklad má vlastnosť

$$q \sim_k q' \equiv q, q' \in A_i^{(k)}.$$

for $q, q'; q \neq q'$: ak $q \sim_k q'$. tak q, q' patria do tej istej množiny. Inak patria do rôznych množín.
Overenie, že $q \sim_k q'$ prebehne nasledovne:

- Vykonáme test, či $q \sim_{k-1} q'$ – teda či q, q' patria do tej istej množiny v $k-1$ rozklade.
- Vykonáme test, či $\forall a \in \Sigma : f(q,a) = f(q',a)$, t.j. či $f(q,a), f(q',a)$ patria do tej istej množiny v $k-1$ rozklade.

Krok 3. Opakuj krok 2, kým $(k-1)$-rozklad nie je zhodný s k-rozkladom. Potom posledný rozklad je výstupom algoritmu.

Korektnosť

1. podľa definície 1-ekvivalencie: $q, q' \in A_i^{(1)} \equiv \forall a \in \Sigma : g(q,a) = g(q',a)$
2. s využitím lemy: ak existuje $i : q, q' \in A_i^{(k)} \equiv q, q'$ sú $(k-1)$-ekvivalentné & $\forall a \in \Sigma : f(q,a) \sim_{k-1} f(q',a)$
3. cyklus sa musí raz zastaviť, t.j. musí existovať k : $A_1^{(k)} = A_1^{(k-1)}, \ldots, A_{n_k}^{(k)} = A_{n_k}^{(k-1)}$. Využijúc vlastnosť, že $q_1 \sim_k q_2 \Rightarrow q_1 \sim_{k-1} q_2$ vieme, že ak 2 stavy patria do tej istej množiny v k-rozklade, potom sú v rovnakej množine aj v $(k-1)$-rozklade. Ďalej, ak Q obsahuje n stavov, tak rozklad sa zastaví, lebo prinajhoršom dôjdeme k 1-prvkovým množinám.

Odstraňovanie nedosiahnuteľných stavov

Skonštruujeme množinu nasledovne:

$$\begin{aligned} Q_0 &= \{Q_i\} \\ Q_i &= Q_{i-1} \cup \{f(q,a) : q \in Q_{i-1}, a \in \Sigma\} \\ &\vdots \end{aligned}$$

kým nenastane rovnosť $Q_i = Q_{i-1}$. Potom množina Q_i je množina všetkých dosiahnuteľných stavov.

Konštrukcia redukovaného automatu

- zlikvidujeme nedosiahnuteľné stavy
- z množiny ekvivalentných stavov vyberieme reprezentanta (ľubovoľného, je však vhodné ponechať si ekvivalentný stav)

Algoritmus na konštrukciu redukovaného automatu

Krok 1. M nahradíme automatom $M' = (Q', \Sigma, \Delta, f', g', q_I)$, ktorý neobsahuje nedosiahnuteľné stavy.

$$\begin{aligned} Q' &\ldots \text{ dosiahnuteľné stavy z } Q,\ Q' \subseteq Q \\ f', g' &\ldots \text{ majú zúžený definičný obor z } Q \text{ na } Q' \\ f', g' &: Q' \times \Sigma \to \Box. \end{aligned}$$

Krok 2. Zostrojíme rozklad Q' na triedy ekvivalentných stavov $A_1, \ldots, A_n$. Nech $M'' = (Q'', \Sigma, \Delta, f'', g'', q_I)$, pričom $Q'' = \{q_1, \ldots, q_n\} \subseteq Q'$ a

- $q_i \in Q''$
- ak $q_i \in A_i$, potom

$$\begin{aligned} g'' &: Q'' \times \Sigma \to \Delta, g''(q_i,a) = g'(q_i,a) \\ f'' &: Q' \times \Sigma \to Q'' \\ f''(q_i,a) &= q_j \text{ také, že } f'(q_i,a) \in A_j \end{aligned}$$

(u f'' v podstate presmerujeme hranu do reprezentanta).

Definícia. Majme dva konečnostavové automaty $M_1 = (Q, \Sigma, \Delta, f_1, g_1, q_{I_1})$ a $M_2 = (Q, \Sigma, \Delta, f_2, g_2, q_{I_2})$. Hovoríme, že automaty M_1, M_2 sú *ekvivalentné*, ak

$$\forall \alpha \in \Sigma^* : g_1^*(q_{I_1}, \alpha) = g_2^*(q_{I_2}, \alpha)$$

Algoritmus na testovanie ekvivalencie 2 automatov

Dané sú M_1, M_2. Zostrojíme automat $M = (Q, \Sigma, \Delta, f, g, q_I)$.

Krok 1. Položíme

$$\begin{aligned} Q &= Q_1 \cup Q_2 \\ f(q,a) &= \begin{cases} f_1(q,a), & \text{ak } q \in Q_1 \\ f_2(q,a), & \text{ak } q \in Q_2 \end{cases} \\ g(q,a) &= \begin{cases} g_1(q,a), & \text{ak } q \in Q_1 \\ g_2(q,a), & \text{ak } q \in Q_2 \end{cases} \end{aligned}$$

Týmito definíciami v podstate „zlepíme" 2 automaty do jedného. Ďalej BUNV položíme $q_I = q_{I_1}$, pretože M_1 je ekvivalentný M_2 ak v M platí:

$$\forall \alpha \in \Sigma : \underbrace{g^*}_{g_1^*}(q_{I_1}, \alpha) = \underbrace{g^*}_{g_2^*}(q_{I_2}, \alpha),$$

čo v podstate znamená, že $q_{I_1} = q_{I_2}$ v automate M.

Krok 2. Na takto zostrojený M použijeme algoritmus, ktorý zostrojí rozklad $Q = Q_1 \cup Q_2$ na triedy ekvivalentných stavov $A_1, \ldots, A_k$.

Krok 3. Ak q_{I_1}, q_{I_2} patria do tej istej množiny v rozklade $A_1, A_2, \ldots$, tak

$$q_{I_1} \sim q_{I_2} \Rightarrow M_1 \text{ je ekvivalentný } M_2$$

Ak patria do rôznych množín, tak

$$q_{I_1} \not\sim q_{I_2} \Rightarrow M_1 \text{ nie je ekvivalentný } M_2$$

Definícia. Dva automaty $M_1 = (Q, \Sigma, \Delta, f_1, g_1, q_{I_1})$ a $M_2 = (Q, \Sigma, \Delta, f_2, g_2, q_{I_2})$ sú *izomorfné*, ak existuje funkcia

$$T : Q_1 \to Q_2$$

taká, že

1. $T(f_1(q_1, S)) = f_2(T(q_1, S)) \quad \forall q \in Q_1, S \in \Sigma$
2. $g_1(q, S)) = g_2(T(q), S) \quad \forall q \in Q_1, S \in \Sigma$
3. $T(q_{I_1}) = q_{I_2}$
4. $T(q) = T(q') \Rightarrow q = q'$
5. $\forall q_2 \in Q_2 \; \exists q_1 \in Q_1 : T(q_1) = q_2$

Veta. *Ak sú dva automaty izomorfné, potom sú aj ekvivalentné.*

Poznámka. Naopak táto veta neplatí, kontrapríkladom je redukovaný a neredukovaný automat.

Veta. *Ak sú 2 automaty M_1, M_2 oba redukované a ekvivalentné, potom sú izomorfné.*

Dôkaz:
Nech $M_1 = (Q_1, \Sigma, \Delta, f_1, g_1, q_{I_1})$ a $M_2 = (Q_2, \Sigma, \Delta, f_2, g_2, q_{I_2})$ sú redukované a izomorfné. Zostrojíme $M = (Q_1 \cup Q_2, \Sigma, \Delta, f, g, q_{I_1})$.

$$f(q, a) = \begin{cases} f_1(q, a), & \text{ak } q \in Q_1 \\ f_2(q, a), & \text{ak } q \in Q_2 \end{cases}$$

$$g(q, a) = \begin{cases} g_1(q, a), & \text{ak } q \in Q_1 \\ g_2(q, a), & \text{ak } q \in Q_2 \end{cases}$$

Na automat M aplikujeme algoritmus na tvorbu tried ekvivalentných stavov

$$Q_1 \cup Q_2 = A_1 \cup A_2 \cup \ldots \cup A_n$$

Potom

- A_i obsahuje navzájom ekvivalentné stavy
- žiadne A_i neobsahuje viac ako 1 stav z Q_1
  - pretože M_1 je redukovaný automat, t.j. neobsahuje ekvivalentné stavy
  - ak by $q, q' \in Q_1$ a $q, q' \in A_i$, potom by $q \sim q'$, čo by bol spor
- žiadne A_i neobsahuje 2 stavy z Q_2, lebo M_2 je redukovaný

Potom A_i je tvaru

$$\{q_1\} \qquad \{q_2\} \qquad \{q_1, q_2\} \qquad q_1 \in Q_1, q_2 \in Q_2$$

Nech A_i má tvar $\{q_1\}, q_1 \in Q_1$. M_1 je redukovaný, t.j. nemá nedosiahnuteľné stavy, teda

$$\exists w_0 \in \Sigma^* : f_1^*(q_{I_1}, w_0) = q_1$$

Zoberme stav $q_2 \in Q_2$ taký, že $f_2^*(q_{I_2}, w_0) = q_2$ a zoberme ľubovoľné $w \in \Sigma^*$. Ako sa budú chovať M_1, M_2 pri vstupe w? Keďže sú ekvivalentné, tak na vstup $w_0 w$ musia dať rovnaký výstup, teda

$$\begin{aligned} g_1^*(q_{I_1}, w_0 w) &= g_2^*(q_{I_2}, w_0 w) \\ g_1^*(q_{I_1}, w_0) \cdot g_1^*(f_1^*(q_{I_1}, w_0), w) &= \\ g_2^*(q_{I_2}, w_0) \cdot g_2^*(f_2^*(q_{I_2}, w_0), w) & \end{aligned}$$

Odseknutím jedného symbolu na každej strane rovnice a označením $f_1^*(q_{I_1}, w_0) = q_1$, $f_2^*(q_{I_2}, w_0) = q_2$ máme

$$\begin{aligned} g_1^*(q_1, w) &= g_2^*(q_2, w) \\ q_1 &\sim q_2 \end{aligned}$$

čiže q_2 musí v rozklade $A_1, \ldots, A_n$ patriť do tej istej množiny A_i ako q_1. To je ale spor, lebo sme predpokladali, že množina A_i je tvaru $\{q_1\}$
Analogicky ukážeme, že A_i nie je tvaru $\{q_2\}$
Definujme teraz zobrazenie $T : Q_1 \to Q_2$. Ak $A_i = \{q_1, q_2\}$, tak $T(q_1) = q_2$. Toto zobrazenie spĺňa vlastnosti 4. a 5. automaticky.

3. M_1, M_2 sú ekvivalentné, teda

   $$\begin{aligned} g_1^*(q_{I_1}, w) &= g_2^*(q_{I_2}, w) \quad \forall w \in \Sigma^* \\ q_{I_1} &\sim q_{I_2} \end{aligned}$$

   tzn. majú ekvivalentné počiatočné stavy. Z definície T potom dostávame:

   $$T(q_{I_1}) = q_{I_2}$$

2. stavy $q \in Q_1, T(q) \in Q_2$ tvoria v rozklade $A_1, \ldots, A_n$ jednu množinu, čiže

   $$\exists i_0 : A_{i_0} = \{q, T(q)\} \Rightarrow q \sim T(q),$$

   čo v M znamená

   $$g^*(q, w) = g^*(T(q), w) \quad \forall w \in \Sigma^*$$

   Ak to platí $\forall w \in \Sigma^*$, tak aj pre $b \in \Sigma$ (jeden znak)

   $$\begin{aligned} g^*(q, b) &= g^*(T(q), b) \\ \underbrace{g(q, b)}_{\in Q_1} &= \underbrace{g(T(q), b)}_{\in Q_2} \end{aligned}$$

1. nech $q \sim T(q) \quad \forall w \in \Sigma^*$. Zoberme ľubovoľné $w' \in \Sigma^*$ a položme $w' = ws$. Keďže zo stavov $q, T(q)$ musíme dostať rovnaký výstup, platí:

   $$\begin{aligned} g^*(q, sw) &= g^*(T(q), sw) \\ g^*(q, s) g^*(f^*(q, s), w) &= g^*(T(q), s) g^*(f^*(T(q), s), w) \end{aligned}$$

   Keďže $q \in q'$, môžeme odrezať prvé symboly, ktoré sú rovnaké

   $$\begin{aligned} g^*(f^*(q, s), w) &= g^*(f^*(T(q), s), w) \; \forall w \in \Sigma^* \\ f(q, s) &\sim f(T(q), s) \text{ použijeme definíciu } f \\ f_1(q, s) &\sim f_2(T(q), s) \end{aligned} \tag{5}$$

Keďže $q \in T(q)$ pre $\forall q \in Q$, tak aj pre $\tilde{q} = f_1(q,s)$, čiže potom aj

$$\begin{aligned} \tilde{q} &\sim T(\tilde{q}) \\ f_1(q,s) &\sim T(f_1(q,s)) \\ \text{zo vzťahu (5)} & \\ \underbrace{T(f_1(q,s))}_{\in Q_2} &\sim \underbrace{f_2(T(q),s)}_{\in Q_2} \end{aligned}$$

Máme ekvivalenciu dvoch stavov, ale M je z predpokladu redukovaný, teda nastane len rovnosť

$$T(f_1(q,s)) = f_2(T(q),s)$$

Veta. *Pre každý automat M je automat M', ktorý vznikne z M odstránením nedosiahnuteľných a ekvivalentných stavov automatom s minimálnym počtom stavov spomedzi všetkých automatov, ktoré sú ekvivalentné s M.*

Dôkaz:
Nech M je ľubovoľný automat a nech M_R je ekvivalentný s M, pričom M_R nech vznikol štandardnou redukciou (odstránením nedosiahnuteľných a ekvivalentných stavov). Nech M' je ekvivalentný s M_R a má menej stavov ako M_R (čiže nech M' je automat s minimálnym počtom stavov). Je zrejmé, že:

- M' nemá nedosiahnuteľné stavy (ak by ich mal, vieme ich odstrániť)
- M' nemá 2 stavy, ktoré by boli ekvivalentné

Máme M, M_R navzájom ekvivalentné a ani M ani M_R nemá nedosiahnuteľné ani ekvivalentné stavy. Teda M, M_R sú izomorfné. Čiže, existuje

$$T : Q' \to Q_R$$

Potom z vlastnosti 4.) sa rôzne stavy zobrazujú do rôznych (čiže T je prosté) a z vlastnosti 5.)

$$\forall q_R \in Q_R \; \exists q' \in Q' : T(q') = q_R$$

máme zaistené, že T je surjekcia. Z týchto dvoch vlastností máme

$$\|Q'\| = \|Q_R\|,$$

čiže redukovaný automat je minimálny.

Konečnostavové akceptory

Definícia. Konečnostavový akceptor je usporiadaná pätica $M = (Q, \Sigma, f, q_I, F)$, kde

Q	konečná množina stavov
Σ	vstupná abeceda
$f : Q \times \Sigma \to Q$	prechodová funkcia
$q_I \in Q$	počiatočný stav
$F \subseteq Q$	koncové stavy

Funkciu f^* definujeme podobne ako u konečnostavových automatov:

$$\begin{aligned} f^*(q,\varepsilon) &= q \\ f^*(q,\alpha a) &= f(f^*(q,\alpha),a) \end{aligned}$$

Jazyk akceptovaný akceptorom M je

$$L(M) = \{w \in \Sigma : f(q_i, w) \in F\}$$

Definícia. *Nedeterministický konečnostavový akceptor* je usporiadaná pätica

$$M = (Q, \Sigma, \delta, I, F).$$

Q	konečná množina stavov
Σ	vstupná abeceda
$\delta : Q \times \Sigma \to 2^Q$	prechodová funkcia
$I \subseteq Q$	počiatočné stavy
$F \subseteq Q$	koncové stavy

Potom funkciu $\delta^* : 2^Q \times \Sigma^* \to 2^Q$ definujeme takto:

$$\begin{aligned} \delta^*(A,\varepsilon) &= A \\ \delta^*(A,\alpha a) &= \bigcup_{q \in \delta(A,\alpha)} \delta(q,a) \quad \forall A \subseteq Q, \forall \alpha \in \Sigma^*, \forall \alpha \in \Sigma \end{aligned}$$

Veta. *Pre každý nedeterministický konečnostavový akceptor M existuje deterministický konečnostavový akceptor M' taký, že $L(M) = L(M')$.*

<u>*Konštrukcia.*</u> Daný je $M = (Q, \Sigma, I, \delta, F)$. Definujme $M' = (Q, \Sigma, q'_I, f, F')$.

$$\begin{aligned} Q' &= 2^Q \\ q'_I &= I \\ F' &= \{A \subseteq Q : A \cap F = \emptyset\} \\ f &: Q' \times \Sigma \to Q' \\ f(A,\alpha) &= \bigcup_{q \in A} \delta(q,a) \quad \forall A \subseteq Q, \forall a \in \Sigma \end{aligned}$$

Potom

$$\begin{aligned} L(M) &= \{w \in \Sigma^* : \delta^*(I,w) \cap F \neq \emptyset\} \\ L(M') &= \{w \in \Sigma^* : f^*(q'_I,w) \in F'\} \end{aligned}$$

Podmienka v $L(M)$ znamená, že aspoň 1 zo stavov je koncový, podmienka v $L(M')$ zase hovorí, že na konci to musí zastať v koncovom stave.

Lema (intermission).

$$\delta^*(I,w) = f^*(q_I,w)$$

Dôkaz:
Prebehne ÚMI podľa $|w|$.

1. nech $w = \varepsilon$.
   $\delta^*(I,w) = \delta^*(I,\varepsilon) = I = q'_I = f^*(q'_I,\varepsilon) = f^*(q'_I,w)$

2. nech $w = \alpha a$ a nech platí IP.

$$\begin{aligned}
\delta(I, \alpha a) &= \bigcup_{q\in\delta^*(I,\alpha)} \delta(q,a)\\
&\overset{\text{IP}}{=\!=} \bigcup_{q\in f^*(I,\alpha)} \delta(q,a)\\
&= \bigcup_{q\in A} \delta(q,a)\\
&\overset{\text{def. } f \text{ v } M}{=\!=} f(f^*(q'_I,\alpha),a) = f^*(q'_I,\alpha a)\\
&= f^*(q'_I, w)
\end{aligned}$$

Dôkaz:

$$\begin{aligned}
w \in L(M) &\equiv \delta^*(I,w)\cap F = \emptyset\\
&\equiv f^*(q_I,w)\cap F = \emptyset\\
&\equiv A \cap F = \emptyset \equiv f^*(q'_I,w) \in F'\\
&\equiv w \in L(M')
\end{aligned}$$

Súvis automatov a gramatík

Veta. *Pre každý deterministický konečnostavový akceptor M existuje regulárna gramatika G taká, že $L(G) = L(M)$.*

Konštrukcia. Daný je $M = (Q, \Sigma, f, q_I, F)$. Definujeme $G = (N, T, P, S)$ nasledovne:

$$\begin{array}{lcllcl}
N & = & Q \cup \{S\} & & & \\
T & = & \Sigma & & & \\
P & : & \text{ak} & f(q,a) = q' & \Rightarrow & q \to aq' \in P\\
 & & \text{ak} & f(q,a) \in F & \Rightarrow & q \to a \in P\\
 & & \text{ak} & f(q_I,a) = q' & \Rightarrow & S \to aq' \in P\\
 & & \text{ak} & f(q_I,a) \in F & \Rightarrow & S \to aq \in P\\
 & & \text{ak} & q_I \in F & \Rightarrow & S \to \varepsilon \in P
\end{array}$$

Lema (yeah, intermission).

$$f^*(q_I,\alpha) = q \equiv S \Rightarrow \ldots \Rightarrow \alpha q \quad \forall\alpha : |\alpha| > 0$$

Dôkaz:
ÚMI k $|\alpha|$

1. nech $|\alpha| = 1$, t.j. $\alpha = a \in \Sigma$

$$\begin{array}{rcl}
f^*(q_I,a) & = & q \equiv\\
f(q_I,a) & = & q \equiv\\
S \to aq & \in & P \equiv\\
S & \Rightarrow^* & aq
\end{array}$$

2. nech $\alpha = \alpha' a$, $a \in \Sigma$ a nech platí IP

$$\begin{array}{rcl}
f^*(q_I,\alpha' a) = f(f^*(q_I,\alpha),a) & = & q \equiv\\
(\exists\bar q)\bar q = f^*(q_I,\alpha') & \& & f(\bar q,a) = q\\
S \Rightarrow^* \alpha'\bar q & \& & \bar q \to aq\\
S & \Rightarrow & \alpha' a q\\
S & \Rightarrow & \alpha q
\end{array}$$

Dôkaz:
Máme 3 možnosti:

1. nech $|w| \geq 2$, t.j. $w = \alpha a, |\alpha| > 0, a \in \Sigma$.

$$\begin{array}{rcl}
w \in L(M) & \equiv & f(q_I,\alpha a) \in F\\
 & \equiv & f(f^*(q_I,\alpha),a) \in F\\
 & \equiv & \exists\bar q : \bar q = f^*(q_I,\alpha) \;\&\; f(\bar q,a) \in F\\
S \Rightarrow^* \alpha\bar q & \& & \bar q \to a\\
S \Rightarrow^* \alpha a & \in & L(G)
\end{array}$$

2. $|w| = 1, w = a \in \Sigma$. Potom

$$w \in L(M) \equiv f(q_i,a) \in F \equiv S \to a \in P \equiv w \in L(G)$$

3. $|w| = 0$, teda $w = \varepsilon$. Potom

$$\begin{array}{rcl}
\varepsilon \in L(M) & \equiv & f(q_I,\varepsilon) = q_I \in F\\
 & \equiv & S \to \varepsilon \in P\\
 & \equiv & w \in L(G)
\end{array}$$

Veta. *Pre každú regulárnu gramatiku G existuje nedeterministický konečnostavový akceptor taký, že*

$$L(M) = L(G')$$

Konštrukcia. Nech je daná $G = (N, T, P, S)$. Definujeme $M = (Q, \Sigma, \delta, I, P)$ nasledovne:

$$\begin{array}{lcl}
Q & = & N \cup \{q_I\}\\
I & = & \begin{cases} \{S\} & \text{ak} \quad S \to \varepsilon \notin P\\ \{S, q_F\} & \text{ak} \quad S \to \varepsilon \in P \end{cases}\\
\Sigma & = & T\\
\delta & : & \text{ak } A \to aB \in P, \text{ potom } \delta(A,a) \ni B\\
 & & \text{ak } A \to a \in P, \text{ potom } \delta(A,a) \ni q_f
\end{array}$$

Dôkaz vety neuvádzame.

ε-akceptory

Povoľujú sa aj ε-hrany – možno sa preklápať do iných stavov, aj keď nič nie je na vstupe.

Definícia. ε-akceptor $M = (Q, \Sigma, \delta, q_I, q_F)$, kde
$Q\ldots$ stavy — $q_F\ldots$ koncový stav
$\Sigma\ldots$ vstupné symboly — $\delta : Q \times \{\Sigma \cup \{\varepsilon\}\} \to 2^Q$
$q_I\ldots$ počiatočný stav — $\delta(q,\varepsilon) \ni q \quad \forall q$

Veta. *Pre každý M ε-akceptor existuje M' nedeterministický konečnostavový akceptor, že:*

$$L(M) = L(M').$$

Dôkaz:
Ukážeme len konštrukciu. Majme $M = (Q, \Sigma, \delta, q_I, q_F)$. Definujme $M' = (Q, \Sigma, \delta', q_I, F)$. Potom:

$$\begin{array}{rcl} \delta' & : & \text{ak } \delta^*(q,\varepsilon) \ni q' \;\&\; \delta(q',a) = q'', \text{ tak } \delta'(q,a) \ni q'' \\ F' & : & \text{ak } \delta^*(q,\varepsilon) \ni q_F, \text{ potom } F' \ni q \end{array}$$

Dôkaz neuvádzame.

Poznámka. Nedeterministický automat je špeciálny prípad ε-akceptora.

$$\forall q : \delta(q,\varepsilon) = \{q\}$$

Regulárne výrazy

Definícia. Nech Σ je konečná abeceda. Potom *regulárny výraz* zo symbolov (reťazcov) abecedy Σ je

- $\emptyset, \varepsilon$ sú regulárne výrazy
- $a \in \Sigma$ sú regulárne výrazy
- ak α, β sú regulárne výrazy, potom aj $\alpha \cdot \beta, \alpha + \beta, \alpha^*$ sú regulárne výrazy.

Definícia. *Hodnota* regulárneho výrazu α je množina $\underline{\alpha}$, pre $\underline{\alpha} \subseteq \Sigma^*$ definovaná nasledovne:

- $\underline{\emptyset} = \emptyset$, $\underline{\varepsilon} = \varepsilon$, $\underline{a} = \{a\}$
- $\underline{\alpha \cdot \beta} = \{xy \in \Sigma^* : x \in \underline{\alpha} \;\&\; y \in \underline{\beta}\}$
- $\underline{\alpha + \beta} = \{x \in \Sigma : x \in \underline{\alpha} \vee y \in \underline{\beta}\} = \underline{\alpha} \cup \underline{\beta}$
- $\underline{\alpha^*} = \{x_1, \ldots, x_n \in \Sigma^* : n \geq 0, x_1, \ldots, x_n \in \underline{\alpha}\}$

Veta.

$$\begin{array}{rcl} \alpha + \beta & = & \beta + \alpha \\ \alpha + \emptyset & = & \emptyset + \alpha \\ \alpha \cdot \varepsilon & = & \varepsilon \cdot \alpha \\ (\alpha + \beta) + \gamma & = & \alpha + (\beta + \gamma) \\ \alpha \cdot \emptyset & = & \alpha = \emptyset \cdot \alpha \\ \emptyset^* = \varepsilon^* & = & \varepsilon \\ \alpha^* & = & (\alpha^*)^* = \alpha^*\alpha^* \\ (\alpha + \beta)^* & = & (\alpha^*\beta^*)^* \end{array}$$

Veta. *Nech $\alpha x + \beta = x$ je rovnica, kde α, β sú regulárne výrazy, x je neznáma. Potom jej jediným riešením je $\alpha^*\beta$.*

Dôkaz:

- Existencia riešenia: Dosadíme za x do rovnice:

$$\begin{array}{rcl} \alpha \cdot (\alpha^*\beta) + \beta & = & \alpha^*\beta \\ (\alpha\alpha^* + \varepsilon)\beta & = & \alpha^*\beta \\ (\alpha\alpha^* + \varepsilon) & = & \alpha^* \\ (\alpha\alpha^* + \varepsilon)^* & = & (\alpha^*)^* \\ ((\alpha\alpha^*)^* \cdot \varepsilon^*)^* & = & \alpha^* \\ ((\alpha\alpha^*)^*)^* & = & \alpha^* \\ (\alpha\alpha^*) & = & \alpha^* \\ \alpha & = & \alpha \end{array}$$

- Jednoznačnosť riešenia: Nech y vyhovuje rovnici. Overme, či $y = x$.

$$\begin{array}{rcl} y & = & \alpha y + \beta = \alpha(\alpha y + \beta) + \beta \\ & = & \alpha\alpha y + \alpha\beta + \beta \supseteq \alpha^2 y \\ \alpha^2 y & = & \alpha^2(\alpha y + \beta) = \alpha^3 y + \alpha^2\beta \supseteq \alpha^3 y \\ & \vdots & \end{array}$$

Potom $\forall k \in \mathbb{N} : y \supseteq \alpha^k\beta$, ale aj $y \supseteq \alpha^{k-1}\beta$.
Nech $w \in \alpha^*\beta = x, w \in \Sigma^*$. Potom w je tvaru

$$w = \underbrace{u_1}_{\alpha}, \ldots, \underbrace{u_n}_{\alpha}, \underbrace{v}_{\beta},$$

čiže $\exists k : w \in \alpha^k\beta \Rightarrow w \in y$. Z toho máme $w \in x \Rightarrow w \in y$, pre ľubovoľné $w \in \Sigma^*$, teda $x \subseteq y$.

Poznámka (Sústavy rovníc).

$$\begin{array}{rcl} x & = & \alpha x + \beta y + \gamma \\ y & = & \varphi x + \psi y + \tau \end{array}$$

Z prvej rovnice vyjadríme:

$$\begin{array}{rcl} y & = & \psi y + (\varphi x + \tau) \text{ (použijeme vetu)} \\ y & = & \psi^* y \cdot (\varphi x + \tau) \end{array}$$

Dosadením do prvej rovnice:

$$\begin{array}{rcl} x & = & \alpha x + \beta(\psi^* y \cdot (\varphi x + \tau)) + \gamma \\ x & = & \alpha x + \beta\psi^* y\varphi x + \beta\psi^* y\tau + \gamma \end{array}$$

Na túto rovnicu použijeme vetu. Jej riešením bude:

$$x = x(\alpha + \beta\psi^* y\varphi) + \beta\psi^* y + \gamma$$

Vetu použijeme ešte raz, čím máme:

$$\begin{array}{rcl} x & = & (((\alpha + \beta\psi^* y\varphi) + \beta\psi^* y + \gamma)(\alpha + \beta\psi^* y\varphi)) + \\ & & + \beta\psi^* y + \gamma \\ x & = & (\alpha + \beta\psi^* y\varphi)^*(\beta\psi\tau + \gamma) \end{array}$$

Riešenie pre n rovníc:

$$\begin{array}{rcl} x_1 & = & \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n + \beta_1 \\ & \vdots & \\ x_n & = & \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n + \beta_n \end{array}$$

- vyjadríme x_n ako funkciu premenných $x_1, \ldots, x_{n-1}$
- x_n dosadíme do prvých $n-1$ rovníc, čím získame sústavu n rovníc o n neznámych
- predchádzajúci krok opakujeme pre všetkých n premenných. Získané hodnoty premenných $x_1, \ldots, x_n - 1$ dosadíme do poslednej rovnice.

Veta. *Pre každý regulárny výraz α existuje ε-akceptor M taký, že*

$$L(M) = \underline{\alpha}$$

Konštrukcia. Cez skladanie regulárnych výrazov.

Veta. *Pre každý deterministický akceptor M existuje regulárny výraz α taký, že*

$$L(M) = \underline{\alpha}$$

<u>*Konštrukcia.*</u> Majme $M = (Q, \Sigma, f, q_I, F)$, pričom

$$Q = \{q_1, \ldots, q_n\}, \qquad \Sigma = \{a_1, \ldots, a_k\}$$

Definujeme $\overline{q_1}, \ldots, \overline{q_n} \subseteq \Sigma^*; \overline{q_i} = \{w \in \Sigma^* : f^*(\overline{q_i}, w) \in F\}$. $\overline{q_i}$ sú vlastne všetky slová, na ktoré sa dostaneme zo stavu q_i do koncového stavu.
Nech $w \in \overline{q_i}$. Potom

- buď $w = \varepsilon$, potom $q_i \in F$
- alebo $w = a_j\beta$, kde $a_j \in \Sigma$ a $\beta \in \overline{f(q_i, a_j)}$
  Potom
  $$\overline{q_i} = a_1\overline{f(q_i, a_1)} + a_2\overline{f(q_i, a_2)} + \ldots + a_k\overline{f(q_i, a_k)} + \varphi_i,$$
  kde $i \in [1, n], \varphi_i = \begin{cases} \varepsilon, & \text{ak} \quad q_i \in F \\ \emptyset, & \text{ak} \quad q_i \notin F \end{cases}$
  Získavame sústavu n rovníc o n neznámych, jej vyriešením získame regulárne výrazy pre $q_1, \ldots, q_n$. Hľadaný regulárny výraz α bude potom
  $$\alpha = \underbrace{\overline{q_1}}_{q_I} = \{w \in \Sigma^* : f^*(q_1, w) \in F\}$$

Uzáverové vlastnosti regulárnych jazykov

Veta. *Trieda regulárnych jazykov je uzavretá na komplement.*
Ak L je regulárny jazyk, $L \subseteq \Sigma^*$, tak aj

$$L^{\complement} = \Sigma^* \setminus L = \{w \in \Sigma : w \notin L\}$$

je regulárny jazyk.

Dôkaz:
Majme $L = L(M)$, kde $M = (Q, \Sigma, f, q_I, F)$ je deterministický konečnostavový akceptor. Definujme $M' = (Q, \Sigma, f, q_I, F')$, kde $F' = Q \setminus F$.
Potom $w \in L(M) \equiv w \notin L(M')$, lebo

$$\begin{aligned} w \in L(M) &\equiv f^*(q_I, w) \in F \\ w \notin L(M) &\equiv f^*(q_I, w) \notin F' \end{aligned}$$

ale platí $F = Q \setminus F'$.

Veta. *Trieda regulárnych jazykov je uzavretá na prienik.*
Ak L_1, L_2 sú regulárne jazyky, tak ak $L_1 \cap L_2$ je regulárny jazyk.

<u>*Konštrukcia.*</u>

Nech $L_1 = L(M_1); M_1 = (Q_1, \Sigma, f_1, q_{I_1}, F_1)$.
$L_2 = L(M_2); M_2 = (Q_2, \Sigma, f_2, q_{I_2}, F_2)$.

Definujme $M = (Q, \Sigma, f, q_I, F)$.

$$\begin{aligned} Q &= Q_1 \times Q_2,\ q = [q_1, q_2] \\ q_I &= [q_{I_1}, q_{I_2}] \\ f([q_1, q_2]) &= [f_1(q_1, a), f_2(q_2, a)] \\ F &= F_1 \times F_2 \end{aligned}$$

Lema.

$$\forall w : f^*(q_I, w) = [f_1^*(q_I, w), f_2^*(q_I, w)]$$

Dôkaz:
ÚMI vzhľadom na $|w|$:

1. $f^*(q_I, \varepsilon) = q_I = [q_{I_1}, q_{I_2}] = [f^*(q_{I_1}, \varepsilon), f^*(q_{I_2}, \varepsilon)]$
2. nech $w = \alpha a$ a nech platí IP.
   $$\begin{aligned} f^*(q_I, \alpha a) &= f(\underbrace{f^*(q_I, \alpha)}_{\text{IP}}, a) = \\ &= f([f_1^*(q_I, \alpha), f_2^*(q_I, \alpha), a) \\ &= [f_1(f_1^*(q_I, \alpha), a), f_2(f_2^*(q_I, \alpha), a)] \\ &= [f_1^*(q_I, \alpha a), f_2^*(q_I, \alpha a)] \end{aligned}$$

Dôkaz:
Pokračujeme v dôkaze vety:

$$\begin{aligned} w \in L(M) &\equiv f^*(q_I, w) \in F \\ &\equiv [f_1^*(q_I, w), f_2^*(q_I, w)] \in F = F_1 \times F_2 \\ &\equiv f_1^*(q_I, w) \in F_1\ \&\ f_2^*(q_I, w) \in F_2 \\ &\equiv w \in L(M_1)\ \&\ w \in L(M_2) \end{aligned}$$

Teda:

$$L(M) = L(M_1) \cap L(M_2)$$

Veta. *Trieda regulárnych jazykov je uzavretá na zjednotenie.*

Dôkaz:
Automat ako v predchádzajúcej vete, len

$$F = (Q_1 \times F_2) \cup (F_1 \times Q_2)$$

Potom $w \in L(M) \equiv w \in L(M_1) \vee w \in L(M_2)$.

Veta. *Trieda regulárnych jazykov je uzavretá na rozdiel.*

Dôkaz:
Stačí si uvedomiť, že $L_1 \setminus L_2 = L_1 \cap L_2^{\complement}$ a využiť uzavretosť regulárnych jazykov na prienik a komplement.

Veta. *Ak L_1, L_2 sú regulárne jazyky, potom aj $L_1 \times L_2 = \{w : w = xy : x \in L_1\ \&\ y \in L_2\}$, $L_1^* = \{w : w = x_1, \ldots, x_n, n \leq 0, x_i \in L\}$ sú regulárne jazyky.*

Dôkaz:
Ak α_1, α_2 sú hľadané regulárne výrazy pre L_1, L_2, potom $\alpha_1\alpha_2, \alpha_1^*$ sú regulárne výrazy pre $L_1 \times L_2$, L_1^*.

Bezkontextové jazyky

Pravidlá sú tvaru

$$A \to \beta, \beta \in \Sigma^*.$$

Bezkontextové jazyky sa dajú generovať bezkontextovými gramatikami.

Veta. *Neexistuje konečnostavový akceptor, ktorý akceptuje*

$$L(M) = L = \{a^n b^n; n \in \mathbb{N}\}$$

Dôkaz:
Dôkaz vykonáme sporom. Majme $M = (Q, \{a, b\}, f, q_I, F)$, kde Q je konečná množina, teda $Q = \{q_I = q_1, \ldots, q_k\}$ a platí $k = \|Q\|$.

$$\begin{aligned} \text{Nech } r_1 &= f^*(q_I, a) \\ r_2 &= f^*(q_I, aa) \\ &\vdots \\ r_i &= f^*(q_I, a^i), \text{ pričom položíme } i = k+1 \end{aligned}$$

Máme viac množín ako stavov, teda

$$\exists j_1, j_2 : r_{j_1} = r_{j_2}.$$

Pri vstupe $a^{j_2}b^{j_2}$ má M akceptovať, lebo $a^{j_2}b^{j_2} \in L$, t.j. $f^*(q_I, a^{j_2}b^{j_2}) \in F$.
Zoberme teraz j_1, j_2 tak, že BUNV $j_1 < j_2$. Potom

$$\begin{aligned} f^*(q_I, a^{j_1}b^{j_2}) &= f^*(f^*(q_I, a^{j_1}), b^{j_2}) = f^*(r_{j_1}, b^{j_2}) \\ &\overset{r_1 = r_2}{=} f^*(r_{j_2}, b^{j_2}) = f^*(f^*(q_I, a^{j_2}), b^{j_2}) \\ &= f^*(q_I, a^{j_2}b^{j_2}) \in F \\ &\Rightarrow a^{j_2}b^{j_2} \in L, \text{teda je akceptované} \end{aligned}$$

Získavame však spor, keďže sme predpokladali, že $j_1 < j_2$.

Dôsledok.

- *jazyk L je bezkontextový a nie je regulárny*
- *pre triedu regulárnych jazykov $\mathcal{RJ}$ a bezkontextových jazykov $\mathcal{BKJ}$ platí*

$$\mathcal{RJ} \subset \mathcal{BKJ}.$$

Veta. *Pre každú bezkontextovú gramatiku G existuje bezkontextová gramatika G', ktorá nemá pravidlá typu $A \to \varepsilon$* a

$$L(G') = L(G) \setminus \varepsilon$$

Dôkaz:
Majme $G = (N, T, P, S)$. Skonštruujme $E = \{A \in N : A \Rightarrow^* \varepsilon\}$ nasledovne:

- $E_1 = \{A \in N : A \to \varepsilon \in P\}$
- $E_{i+1} = E_i \cup \{A \in N : A \to B_1, \ldots, B_l \in P; B_1, \ldots, B_l \in E_i\}$
- ak $E_i = E_{i+1}$, končíme.

G' vytvoríme nasledovne: Každé pravidlo $A \to x_1, \ldots, x_n$ nahradíme pravidlom $A \to y_1, \ldots, y_n$, kde

$$y_i = \begin{cases} x_i, & \text{ak} \quad x_i \in T \\ x_i, & \text{ak} \quad x_i \in N \setminus E \\ \in \{x_i, \varepsilon\}, & \text{ak} \quad x_i \in E \end{cases}$$

pričom $A \to \varepsilon \notin P'$.[1]

[1] Pre istotu treba skontrolovať druhú možnosť vo „vidličke.“ Je správna? — pozn. sadzač

Veta. *Pre každú bezkontextovú gramatiku G existuje bezkontextová gramatika G' taká, že*

- $L(G) = L(G')$
- *G' nemá pravidlá typu $A \to B$, kde $A, B \in N$*

Dôkaz:
Množinu pravidiel P' skonštruujeme nasledovne:
Ak v G máme odvodenie $A \Rightarrow B_1 \Rightarrow B_2 \Rightarrow \ldots \Rightarrow B_k \Rightarrow \alpha$, tak do P' zaradíme pravidlo $A \to \alpha$. Špeciálne: ak $A \Rightarrow \alpha$, tak zaradíme pravidlo $A \to \alpha$.

Definícia. Bezkontextová gramatika je v *Chomského normálnom tvare*, ak má pravidlá tvaru

- $X \to YZ$, pre $X, Y, Z \in N$
- $X \to a$, pre $a \in T$

Veta. *Pre každú bezkontextovú gramatiku G existuje bezkontextová gramatika G' v Chomského normálnom tvare, kde*

$$L(G') = L(G) \setminus \varepsilon.$$

Dôkaz:

- pre danú G zostrojíme G_1, ktorá neobsauje pravidlá $A \to \varepsilon$.
- potom ku G_1 zostrojíme G_2 takú, že G_2 neobsahuje pravidlá tvaru $A \to B$.
- z G_2 vytvoríme G'

Pravidlá:

- $A \to \varepsilon$ – pravidlá tohoto typu už v G_2 nie sú
- $A \to x, x \in \Sigma$, v tomto prípade musí $x \in T$ (neterminál nenastane)
- $A \to x_1, \ldots, x_m; m \geq 2$. Namiesto toho dáme do G' pravidlá:

$$\begin{aligned} A &\to x_1X_1; X_1 = x_2, \ldots, x_m \\ X_1 &\to x_2X_2; X_2 = x_3, \ldots, x_m \\ &\vdots \\ X_{m-2} &\to x_m; X_{m-2} = x_{m-1}x_m \end{aligned}$$

Ďalej pre X_i platí: $\begin{cases} \text{ak } X_i \in N & \text{, potom} \quad X_i = x_i \\ \text{ak } X_i \in T \end{cases}$ Ak X_i je nový neterminál, pridáme $X_i \to x_i$.

Veta (Pumping lema). *Nech L je bezkontextový jazyk, pričom $L = L(G)$. Potom*

1. *buď L je konečná množina*

2. *alebo existujú reťazce u, v, x, y, z také, že*

$$uv^ixy^iz \in L, \forall i \geq 0$$

Naviac, aspoň jeden z reťazcov v, y je rôzny od ε.

Dôkaz:
Ad 1.)
Nech L je konečný jazyk. Potom zrejme niet čo dokazovať.
Ad 2.)
Nech L je nekonečný jazyk. Ak máme gramatiku G, potom ju nahradíme gramatikou G', ktorá bude v Chomského normálnom tvare.
To, že $L(G)$ je nekonečný, znamená, že

$$\forall k\ \exists w \in L(G) : |w| \geq k \tag{6}$$

Ďalej platí, že všetkých reťazcov dĺžky k je najviac $\sum_{i=0}^{k-1} \|T\|^i$, teda konečne veľa. Ak vzťah (6) platí pre každé k, tak tým skôr aj pre

$$k = 2^{\|N\|+2}$$

Ak $w \in L(G)$, musí existovať odvodenie $S \to^* w$. Zakreslime toto odvodenie ako binárny strom. Výška stromu bude aspoň $\log \|N\|+2$; pri výške stromu l bude počet listov najviac 2^l. Dĺžka cesty v strome bude aspoň $\|N\| + 2$, to ale znamená, že máme aspoň $\|N\| + 1$ neterminálov. Z definície gramatiky však máme, že počet neterminálov je $\|N\|$, teda niekde sa musel neterminál opakovať. Označme tento neterminál ako A.

- Nech z „dolného“ A sa vygeneruje nejaké x, t.j.

$$A^D \Rightarrow^* x$$

- z „horného“ A sa vygeneruje slovo, ktoré obsahuje x, lebo dolné A je pod ním

$$A^H \Rightarrow^* vxy \tag{7}$$

  Ak z tohto odvodenia odstránime všetky kroky, ktoré modifikujú „spodné“ A, máme odvodenie

$$A^H \Rightarrow^* vA^Dy \tag{8}$$

- z S sa vygeneruje slovo, ktoré obsahuje vxy, keďže A^H, z ktorého vygenerujeme vxy je pod ním. Ak zoberieme odvodenie $S \Rightarrow^* uvxyz$, odstránime všetky kroky prepisujúce A^H, získame

$$S \Rightarrow^* uA^Hz \tag{9}$$

  Z odvodení (7)-(9) máme

$$\begin{array}{l} S \Rightarrow\ uAz \Rightarrow\ uxz \\ S \Rightarrow\ uAz \Rightarrow\ uvxyz \\ S \Rightarrow\ uAz \Rightarrow\ uv^ixy^iz \end{array}$$

  čiže slovo $uv^ixy^iz \in L(G')$, kde však $L(G)$ je ekvivalentná $L(G)$.

Ukážme ešte[2], že $v \neq \varepsilon \vee y \neq \varepsilon$.
A^D je v ľavom alebo pravom podstrome A^H. BUNV predpokladajme, že je v ľavom podstrome. Nech A^H má synov BC. Potom cesta z A^D do A^H ide cez B. Nech y_1 je odvodené v ľavom podstrome a nech y_2 je tá časť, ktorá je odvodená z C, čo znamená, že

$$C \Rightarrow^* y_2$$

Ale gramatika v Chomského normálnom tvare nemá pravidlá typu $X \to \varepsilon$. Teda $y_2 \neq \varepsilon$, z čoho $y = y_1y_2 \neq \varepsilon$
Analogicky ukážeme, že ak cesta z A^H do A^D ide cez C, tak $v \neq \varepsilon$.

Dôsledok. *Ak slovo $w \in L(G')$ a $|w| \geq 2^{\|N\|+2}$, tak w sa dá rozdeliť na slovo $uvxyz$ tak, že*

$$uv^ixy^iz \in L \qquad \forall i \geq 0,$$

pričom $v \neq \varepsilon$ alebo $y \neq \varepsilon$ a $|vxy| \leq 2\|N\| + 1$

Lema. *Jazyk $L(G) = \{a^nb^nc^n, n \geq 1\}$ sa nedá vygenerovať pomocou bezkontextovej gramatiky.*

Dôkaz:
L je nekonečná množina, teda pre dostatočne veľké n máme, že

$$a^nb^nc^n = uvxyz$$

Z tohto tvaru slova vyplýva, že vo v musia byť písmená rovnakého typu. Písmená rovnakého typu však musia byť aj v y. Ale u, x, z sú pevné, teda $uvvxyyz \notin G$ (to by bolo možné len vtedy, keď by $v = \varepsilon$ a zároveň $y = \varepsilon$, ale to by nebolo delenie podľa pumping lemy!). Gramatika G je bezkontextová, teda slovo $uvxyz$ musí patriť do $L(G)$, čo je spor.

Uzáverové vlastnosti bezkontextových jazykov

Veta. *Trieda bezkontextových jazykov je uzavretá na zjednotenie.*

Dôkaz:
Pre jazyk L_i majme gramatiku $G_i = (N_i, T_i, P_i, S_i)$, pričom nech BUNV platí $N_1 \cap N_2 = \emptyset$.
Definujme $G = (N, T, P, S)$ pre jazyk $L_1 \cup L_2$.

$$\begin{array}{rcl} N &=& N_1 \cup N_2 \\ T &=& T_1 \cup T_2 \\ P &=& T_1 \cup T_2 \cup \{S \to S_1, S \to S_2\} \end{array}$$

[2] Inak povedané, A^D sa nachádza v ľavom alebo v pravom podstrome A^H. Ak je to napríklad v ľavom, tak sa musí v pravom podstrome A^H niečo vygenerovať, lebo naše slovo w má padnúť do $L(G)$, čiže nemôžu v pravom podstrome A^H ostať „visieť“ nejaké neterminály, teda $y \neq \varepsilon$. Podobné je to ak sa A^D nachádza v pravom podstrome A^H, vtedy musí platiť, že $v \neq \varepsilon$.

Z počiatočného symbolu S sa dostaneme do symbolu S_1 a odtiaľ pravidlami P_1 vieme odvodiť $S \Rightarrow^* w \in L(G_1)$ Analogicky, z S sa dostaneme so S_2 a odtiaľ už máme pravidlá P_2, pomocou ktorých odvodíme $S_2 \Rightarrow^* w \in L(G_2)$

Veta. *Trieda bezkontextových jazykov NIE JE uzavretá na prienik.*

Dôkaz:
Stačí uvážiť nasledovné 2 jazyky ako kontrapríklad:

$$\begin{aligned} L_1 &= \{a^k b^n c^n; k, n \geq 1\} \\ L_2 &= \{a^n b^n c^k; k, n \geq 1\} \end{aligned}$$

Zrejme pre jazyk $L = L_1 \cap L_2$ bude platiť

$$L = L_1 \cap L_2 = \{a^n b^n c^n : n \geq 1\}$$

Toto však podľa predchádzajúcej lemy nie je bezkontextový jazyk.

Veta. *Trieda bezkontextových jazykov NIE JE uzavretá na komplement.*

Dôkaz:
Ak by bola táto trieda uzavretá na komplement, potom by to isté platilo aj pre

$$L_1 \cap L_2 = (L_1^\complement \cup L_2^\complement)$$

čo by bol spor.

Zásobníkové automaty

Definícia. Zásobníkový automat je usporiadaná sedmica

$$M = (Q, \Sigma, \Gamma, \delta, q_I, z_I, F),$$

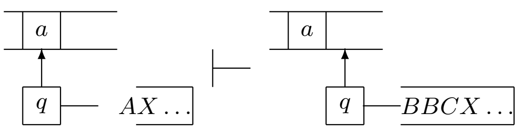
$$\begin{array}{rcl} Q & \ldots & \text{konečná množina stavov} \\ \Sigma & = & \text{vstupná abeceda} \\ \Gamma & \ldots & \text{zásobníková abeceda} \\ Z_i \in \Gamma & \ldots & \text{počiatočný zásobníkový symbol} \\ F \subseteq Q & \ldots & \text{koncové stavy} \\ \delta & : & Q \times (\Sigma \cup \{\varepsilon\}) \to 2^{Q \times \Gamma^*} \end{array}$$

Prechodovú funkciu δ definujeme nasledovne:

$$\delta(q, a, A) = \{q', BBC\}$$

V prípade nedeterministického zásobníkového automatu to bude

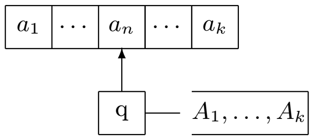
$$\delta(q, a, A) = \{(q', \varepsilon), (q'', AA)\}$$



Definícia. Konfigurácia zásobníkového automatu je usporiadaná trojica

$$(q, \alpha, \gamma) \in Q \times \Sigma^* \times \Gamma^*,$$

pričom q je momentálny stav, α je neprečítaný vstup a γ je obsah zásobníka.



Obr. 1 – Konfigurácia $(q, a_n \ldots a_k, A_1 \ldots A_k)$

Definícia. Jazyk $L(M)$ rozpoznávaný zásobníkovým automatom je

$$L(M) = \{w \in \Sigma^* : (q_i, w, z_I) \vdash \ldots \vdash (q, \varepsilon, \gamma), q \in F\}$$

Ak automat je v koncovom stave, ale vstup nie je prečítaný celý, potom reťazec nie je akceptovaný.

Definícia. Akceptovanie prázdnym zásobníkom definujeme nasledovne:

$$N(M) = \{w \in \Sigma : (q_i, w, Z_i) \vdash \ldots \vdash (q, \varepsilon, \varepsilon)\}$$

Stav q teda nemusí byť koncový, stačí, že je prečítaný celý vstup a zásobník je na konci prázdny.

Lema. *Pre každý zásobníkový automat M existuje zásobníkový automat M' taký, že*

$$N(M) = L(M')$$

Dôkaz:
Nech je daný $M = (Q, \Sigma, \Gamma, \delta, q_I, z_I, \emptyset)$ (na koncových stavoch nezáleží, teda môžu byť aj prázdne). Definujeme $M' = (Q', \Sigma, \Delta', \delta', q'_I, Z'_I, \{q'_F\}$

$$\begin{array}{rcl} Q' & = & Q \cup \{q'_I, q'_F\} \\ \Gamma' & = & \Gamma \cup \{z'_I\} \\ \delta' & : & \delta'(q'_I, \varepsilon, z'_i) = \{(q_I, z_I z'_I)\} \quad \forall a \in Q \\ & & \delta'(q, a, A) = \delta(q, a, A) \quad a \in \Sigma \cup \{\varepsilon\}, A \in \Gamma \\ & & \delta'(q, \varepsilon, z'_I) = \{(q'_F, \varepsilon)\} \end{array}$$

Akceptujúci výpočet v M:

$$\begin{array}{lcl} (q_I, a_1 \ldots a_n, z_i) & \vdash \ldots \vdash & (q_I, a_1 \ldots a_n, A_{i_1} \ldots A_{i_m}) \\ & \vdash \ldots \vdash & (q_{n+1}, \varepsilon, \varepsilon) \end{array}$$

Príslušný výpočet v M':

$$\begin{array}{lcl} (q'_I, a_1 \ldots a_n, z'_i) & \vdash & (q_I, a_1 \ldots a_n, A_{i_1} \ldots A_{i_m}, z_I z'_I) \\ & \vdash \ldots \vdash & (q_{n+1}, \varepsilon, Z'_I) \\ & \vdash \ldots \vdash & (q'_F, \varepsilon, \varepsilon) \end{array}$$

Veta. *Pre každý zásobníkový automat M existuje zásobníkový automat M', že*

$$L(M) = N(M')$$

Dôkaz:
Nech je daný $M = (Q, \Sigma, \Gamma, \delta, q_I, z_I, F)$. Definujeme $M' = (Q', \Sigma, \Delta', \delta', q_I', Z_I', \emptyset\}$.

$$\begin{array}{lcl} Q' & = & Q \cup \{q_I', q_F'\} \\ \Gamma' & = & \Gamma \cup \{z_I'\} \\ \delta' & : & \delta'(q_I', \varepsilon, z_I') = \{(q_I, z_I z_I')\} \\ & & \text{ak } \delta(q,a,A) \ni (q', \beta), \text{ potom } \delta'(q,a,A) \ni (q', \beta) \\ & & \delta'(q, \varepsilon, A) \ni (q_F', \varepsilon) \quad \forall q \in F, a \in \Gamma \cup \{z_I'\} \\ & & \delta(q_F', \varepsilon, A) \ni (q_F', \varepsilon) \quad \forall A \in \Gamma \cup \{z_I'\} \end{array}$$

Výpočet v pôvodnom automate M:

$$\begin{array}{lll} q_I, a_1 \dots a_n, z_I) & \vdash \dots \vdash & (q_I, a_1 \dots a_n, A_1 \dots A_m) \\ & \vdash \dots \vdash & (q_{n+1}, \varepsilon, A_{n+1,1} \dots A_{n+1,m}) \\ & & \text{kde } q_{n+1} \text{ musí patriť do } F \end{array}$$

Výpočet v automate M':

$$\begin{array}{lll} (q_I', a_1 \dots a_m, z_I') & \vdash & (q_I, a_1 \dots a_n, z_I, z_I') \\ & \vdash \dots \vdash & (q_i, a_i \dots a_n, A_{i1} \dots A_{im}, z_I') \\ & \vdash \dots \vdash & (q_{n+1}, \varepsilon, A_{n+1,1} \dots A_{n+1,m}, z_I') \\ & \vdash \dots \vdash & (q_F', \varepsilon, A_{n+1,2} \dots A_{n+1,m}, z_I') \\ & \vdash \dots \vdash & (q_F', \varepsilon, \varepsilon) \end{array}$$

Veta. *Pre každú bezkontextovú gramatiku G existuje zásobníkový automat M' taký, že*

$$L(G) = N(M')$$

Dôkaz:
Majme $G = (N, T, P, S)$, ktorú pre jednoduchosť prevedieme do Chomského normálneho tvaru

$$G' = (N', T', P', S')$$

Čiže G' bude mať len pravidlá typu

$$A \to BC \qquad A \to a.$$

Definujeme

$$M = (\{q_I\}, T, N, \delta, q_I, S', \emptyset)$$

Ak pravidlo $A \to BC \in P'$, potom $\delta(q_I, \varepsilon, A) \ni (q_I, BC)$.
Ak pravidlo $A \to a \in P'$, potom $\delta(q_I, a, A) \ni (q_I, \varepsilon)$.
Zoberme najľavejšie odvodenie v G:

$$S' \quad \Rightarrow \dots \Rightarrow \quad a_1 \dots a_k A_1 \dots A_n$$

Máme 2 možnosti ako pokračovať v odvodzovaní:

- $a_1 \dots a_k \underbrace{BCA_2A_3 \dots A_n}_{A_1' \dots A_k'} \Rightarrow \dots \Rightarrow a_1 \dots a_l = w$
- $a_1 \dots a_k a \underbrace{A_2 \dots A_n}_{A_1' \dots A_k'} \Rightarrow \dots \Rightarrow a_1 \dots a_j = w$

Výpočet v M:

$$(q_I, a_1 \dots a_k, S') \vdash \dots \vdash (q_I, a_{i+1} \dots a_k, A_1 \dots A_N)$$

$$\vdash \dots \vdash \begin{cases} (q_I, a_{i+1} \dots a_k, BCA_2 \dots A_n) \\ (q_I, \underbrace{a_{i+2}}_{A_1 \to a_{i+1}} \dots a_m, A_2 \dots A_n) \end{cases}$$

$$\vdash \dots \vdash (q_I, \varepsilon, \varepsilon)$$

Lema. *Pre každý zásobníkový automat M existuje zásobníkový automat M' taký, že výška zásobníka v jednom kroku sa mení maximálne o jedna.*

Dôkaz:
Ak v M máme inštrukciu $\delta(q, a, A) \ni (q', BBCD)$, nahradíme ju nasledovnými inštrukciami:

$$\begin{array}{rcl} \delta(q, a, A) & \ni & (q_1, CD) \\ \delta(q, \varepsilon, C) & \ni & (q_2, BC) \\ \delta(q_2, \varepsilon, B) & \ni & (q_3, BB) \end{array}$$

kde q_1, q_2, q_3 sú nové stavy.
V M' sú potom všetky inštrukcie typu:

$$\delta(q,a,A) \ni \begin{cases} (q', \varepsilon) & \text{výška zásobníka klesne o 1} \\ (q', B) & \text{výška sa nezmení} \\ (q', BC) & \text{výška vzrastie o 1} \end{cases}$$

Definícia. Situáciu, že zásobníkový automat po prečítaní reťazca α vyhodí zo zásobníka symbol A a skončí po nejakom čase v stave q_2 budeme označovať

$$[q_1, A, q_2]$$

Lema. *Ak $(q_1, \alpha, A) \vdash \dots \vdash (q_2, \varepsilon, \varepsilon)$, potom $[q_1, A, q_2] \Rightarrow \dots \Rightarrow \alpha \in T^*$*

Dôkaz:
Dôkaz je príliš dlhý, preto ho neuvádzame[3].

Lema. *Ak $[q_1, A, q_2] \Rightarrow \dots \Rightarrow \alpha \in T^*$, potom $(q_1, \alpha, A) \vdash \dots \vdash (q_2, \varepsilon, \varepsilon)$*

Dôkaz:
Dôkaz vykonáme analogicky ako v predchádzajúcej leme.

Veta. *Pre každý zásobníkový automat existuje bezkontextová gramatika G taký, že*

$$L(G) = N(M)$$

[3] Ale na prednáške bol!

Dôkaz:
$\supseteq$:

$$\begin{aligned} w \in N(M) &\equiv (q_I, w, z_I) \vdash \ldots \vdash (q, \varepsilon, \varepsilon) \\ \text{z lemy} &\equiv [q_I, z_I, q] \Rightarrow \ldots \Rightarrow w \end{aligned} \tag{10}$$

Potom gramatika má pravidlá typu

$$S \to [q_I, z_I, \tilde{q}] \quad \forall \tilde{q} \in Q$$

Toto však platí špeciálne pre $\tilde{q} = q$, čiže

$$S \to [q_I, z_I, q]$$

a teda v gramatike G máme odvodenie

$$S \Rightarrow \ldots \Rightarrow [q_I, z_I, q] \Rightarrow \ldots \Rightarrow w$$

čo vlastne znamená, že $w \in L(G)$.
$\subseteq$:
Nech $w \in L(G)$, potom v G máme odvodenie

$$S \Rightarrow \ldots \Rightarrow w$$

V prvom kroku sa prepisuje symbol S, lenže pre symbol S máme len pravidlá typu

$$S \to [q_I, z_I, \tilde{q}], \text{kde } \tilde{q} \in Q$$

Teda odvodenie $S \Rightarrow \ldots \Rightarrow w$ má tvar podľa ekvivalencie (10):

$$S \Rightarrow [q_I, z_I, \tilde{q}] \Rightarrow \ldots \Rightarrow w$$

Potom ale platí, že $w \in N(M)$.

Dôsledok. *Každý zásobníkový automat M sa dá nahradiť ekvivalentným zásobníkovým automatom M', že*

$$Q' = \{q_I\}$$

Veta. *Ak L_1 je bezkontextový jazyk a L_2 je regulárny jazyk, potom*

$$L_1 \cap L_2$$

je bezkontextový jazyk.
Bezkontextové jazyky sú uzavreté na prienik s regulárnymi jazykmi.

Dôkaz:
Pre L_1 majme zásobníkový automat

$$M_1 = (Q_1, \Sigma, \Gamma, q_{I_1}, z_I, \delta, F_1),$$

ktorý akceptuje L_1 koncovým stavom.
Pre L_2 majme deterministický konečnostavový akceptor

$$M_2 = (Q_2, \Sigma, q_{I_2}, f_2, F_2)$$

Zostrojme zásobníkový automat M, ktorý súčasne simuluje M_1 a M_2 Uvedomme si, že M_1 ako zásobníkový automat nemusí čítať vstup (stačí, že pracuje zo zásobníkom), ale M_2 musí stále niečo čítať zo vstupu. Teda skonštruujeme:

$$M = (Q, \Sigma, \Gamma, q_I, z_I, \delta', F)$$

$$\begin{aligned} Q &= Q_1 \times Q_2 \\ q_I &= [q_{I_1}, q_{I_2}] \\ F &= F_1 \times F_2 \end{aligned}$$

Teda stav automatu M bude tvaru $q = [q_1, q_2]$, kde q_1 je stav M_1 a q_2 je stavom M_2. Tvar množiny F zaisťuje, že M akceptuje $\equiv$, ak M_1 akceptuje a zároveň akceptuje M_2.
Pre prechodovú funkciu δ' bude platiť:

- ak $\delta(q_1, a, A) \ni (q_1', \beta)$ & $f(q_2, a) = q_2'$, potom

$$\delta'([q_1, q_2], a, A) \ni ([q_1', q_2'], \beta) \quad \forall a \in \Sigma, a \neq \varepsilon$$

- ak $\delta(q_1, \varepsilon, A) \ni (q_1', \beta)$, potom

$$\delta'([q_1, q_2], \varepsilon, A) \ni ([q_1, q_2], \beta) \quad \forall q_2 \in Q$$

(druhý stroj čaká, kým prvý pracuje so zásobníkom)

Poznámka. Prienik 2 bezkontextových jazykov nie je bezkontextový aj preto, že by bolo treba uchovávať stav dvoch zásobníkov[4] (pre každý zásobníkový automat prislúchajúci jednému z jazykov zvlášť). Pri použití dvoch zásobníkov v podstate získavame Turingov stroj so všetkými dôsledkami (Halting Problem).

[4] Toto však nie je argument! Treba využiť Pumping lemu.