Ondrej Vršanský

Materiál na štátnicu

Formálne jazyky a automaty

Verzia 1

07.04.03

Obsah.

8. Formálne jazyky a automaty.

8.1. Úvod do problematiky.

8.1.1. Základné pojmy.

Definícia.

Neprázdnu konečnú množinu symbolov Σ nazývame *abeceda*. Konečná postupnosť symbolov z Σ je *slovo* nad abecedou Σ. Množina slov nad abecedou Σ je *jazyk* nad abecedou Σ. prázdne slovo označujeme symbolom ε. Mohutnosť množiny, vytvárajúcej slovo, sa nazýva *dĺžka slova*.

8.1.2. Elementárne operácie nad slovami.

Definícia.

Zreťazenie slov $v = v_1..v_n$ a $w = w_1..w_m$ je slovo $vw = v_1..v_n w_1..w_m$. *Zrkadlový obraz* slova $v = v_1..v_n$ je slovo $v^R = v_n..v_1$.

Označenie.

Množinu všetkých slov nad abecedou Σ označujeme Σ*. Množinu všetkých neprázdnych slov nad abecedou Σ označujeme Σ^+.

Definícia.

Nech $x,y \in \Sigma^*$. Slovo x je *podslovo* slova y, ak $\exists v,w \in \Sigma^* : vxw = y$.

8.1.3. Elementárne operácie nad jazykmi.

Poznámka.

Nad jazykmi možno vykonávať elementárne množinové operácie, čiže zjednotenie, prienik, rozdiel a doplnok v Σ*.

Definícia.

Zreťazenie jazykov L_1 a L_2 je jazyk $L_1L_2 = \{vw;\ v \in L_1 \text{ a } w \in L_2\}$. n-tá *mocnina* jazyka L je jazyk $L^n = \begin{cases} \{\varepsilon\} \leftarrow n = 0 \\ LL^{n-1} \leftarrow n \geq 1 \end{cases}$. *Uzáver* jazyka L je jazyk $L^* = \bigcup_{n \in N} L^n$. $L^* - \{\varepsilon\}$ je *kladný uzáver* jazyka L.

8.1.4. Gramatiky

Definícia.

Usporiadanú štvoricu $G = (N, T, P, \sigma)$ nazývame *gramatika*. N (ď ť) T ( ) P $(N \cup T)^* N (N \cup T)^* \times (N \cup T)^*$ *pravidiel* na úpravu neterminálnych symbolov a σ je štartovací symbol.

Označenie.

Ak pravidlo P α ť β, tak píšeme $\alpha \to \beta$.

Definícia.

Krok odvodenia gramatiky G je relácia $\underset{G}{\Rightarrow}$ $\forall v,w \in (N \cup T)^* : v \underset{G}{\Rightarrow} w \Leftrightarrow \exists x,y,u,z \in (N \cup T)^* : v = xuy \wedge w = xzy \wedge u \to z \in P$ H í w je *odvoditeľné* z v ($v \underset{G}{\Rightarrow}^* w$), x ť v do w $(N \cup T)$ ľ štartovacieho symbola gramatiky G, nazývame jazyk *generovaný* gramatikou G $L(G)$.

8.1.5. Procedurálna reprezentácia jazyka.

Definícia.

Nech F je program, ktorý na vstup $x \in \Sigma^*$ poskytuje odpovede áno/nie. F je *procedurálna reprezentácia* jazyka L, ak na $x \in \Sigma^*$ odpovie áno práve vtedy ď $x \in L$. Vo všeobecnosti í F *akceptuje* L, ak L F odpovie áno.

8.2. Regulárne gramatiky a konečné automaty.

8.2.1. Regulárne gramatiky.

Definícia.

Gramatika $G = (N, T, P, \sigma)$ je *regulárna* (sprava lineárna), ak sú všetky jej pravidlá tvaru $A \to xB$, kde $A \in N$, $B \in N \cup e$ a $x \in T^*$. Jazyk L je *regulárny*, ak je generovaný regulárnou g T g R.

8.2.2. Konečné automaty.

Definícia.

Usporiadanú päticu $A = (K, \Sigma, \delta, q_0, F)$, kde K tavov, Σ je abeceda, δ je prechodová funkcia $K \times \Sigma \to K$, $q_0 \in K$ a $F \subseteq K$, nazývame *konečný automat*. Usporiadanú dvojicu (q, w), kde $q \in K$ a $w \in \Sigma^*$, nazývame *konfigurácia* automatu A, ak w je prípona vstupného í *Krok odvodenia* ($\underset{A}{\rightarrow}$) f g $(q,v) \underset{A}{\rightarrow} (p,w)$ ď $\exists a \in \Sigma : v = aw$ a $\delta(q,a) = p$ P ť í í í upné slovo v sa nazýva *výpočet* $(q_0,v) \underset{A}{\rightarrow}^* (q,\varepsilon)$. Ak výsledný stav $q \in F$ $L(A) = \{v \in \Sigma^*;\ \sigma \underset{A}{\rightarrow}^* v\}$ nazývame jazyk *rozpoznávaný* automatom A.

8.2.3. Nedeterministické automaty.

Definícia.

Automat A sa nazýva *deterministický*, ak $\forall q \in K\ \forall a \in \Sigma\ \exists! p \in K : \delta(q,a) = p$ V prípade je A *nedeterministický*.

Veta 8.2.3.1.

Nech L A. Potom existuje ý automat A', ktorý rozpoznáva L.

Dôkaz.

Nech $A = (K, \Sigma, \delta, q_0, F)$. Potom $K' = \mathbf{P}(K)$ (í $K \Rightarrow |K'| = 2^{|K|}$). $q_0' = \{q_0\}$, $F' = \{M \subseteq \mathbf{P}(K);\ \exists q \in M : q \in F\}$ a $\delta' : \mathbf{P}(K) \times \Sigma \to \mathbf{P}(K)$ je daná predpisom $\delta'(M, a) = \bigcup_{q \in M} \delta(q,a)$ Ľ ť tento automat akceptuje rovnaký jazyk ako pôvodný.

Lema 8.2.3.1.

Nech L A. Potom existuje A' bez ε-prechodov, ktorý rozpoznáva L.

Dôkaz.

Vo dvoch krokoch. Najpr í f $\delta'(q,a) = \{p \in K;\ (q,a) \underset{A}{\rightarrow}^* (p.\varepsilon)\}$ í í ε-prechodov, rozpoznávajúci L P ľ 8.2.3.1 k nemu existuje deterministický automat A' L. Z koštrukcie tohoto automatu, ktorá ô 8 2 3 1 ε-prechodov.

Veta 8.2.3.2.

Nech L A. Potom L je regulárny jazyk.

Dôkaz.

Pre $A = (K, \Sigma, \delta, q_0, F)$ zostrojíme regulárnu gramatiku $G = (N, T, P, \sigma)$, kde $N = K$, $T = \Sigma$, $P = \{q \to ap;\ q,p \in N,\ a \in \Sigma,\ \delta(q,a) = p\} \cup \{q \to \varepsilon;\ q \in F\}$ a $\sigma = q_0$ Ľ ť g

g í T $L(G) = L(A)$.

Lema 8.2.3.2.

Nech L je regulárny jazyk. Potom existuje regulárna gramatika G s pravidlami tvaru $X \to aY$, $X \to Y$, $X \to \varepsilon$, ktorá generuje L.

Dôkaz.

Jednoduché.

Veta 8.2.3.3.

Nech L g P x A $L = L(A)$.

Dôkaz.

Nech $G = (N, T, P, \sigma)$ je „jednokroková" regulárna gramatika, generujúca L x lema 8.2.3.2. Zostrojíme $A = (K, \Sigma, \delta, q_0, F)$, kde $K = N$, $\Sigma = T$, $q_0 = \sigma$, $F = \{X \in N;\ X \to \varepsilon \in P\}$ a $\delta(X,a) = \{Y \in N;\ X \to aY \in P\}$ a $\delta(X,\varepsilon) = \{Y \in N;\ X \to Y \in P\}$ Ľ ť „ í " í g g í ď g amatika vygeneruje terminálne slovo.

Dôsledok 8.2.3.1.

T g g í

Dôkaz.

Tvrdenie vyplýva priamo z predchádzajúcich viet.

8.2.4. Myhill-Nerodova veta.

Definícia.

Nech R P tried ekvivalencie relácie R nazývame *index* R. Triedu ekvivalencie, do ktorej patrí prvok w $[w]$.

Poznámka.

$L \subseteq \Sigma$ $\equiv_L$, kde $\forall v,w \in \Sigma^* : v \equiv_L w \Leftrightarrow \forall x \in \Sigma^* : vx \in L \Leftrightarrow wx \in L$. Pritom ak $\alpha_1..\alpha_n..$ je lexikografické usporiadanie Σ f ľ $T_2[\alpha_1..\alpha_n..,\ \alpha_1..\alpha_n..]$ predpisom $T_2[\alpha_i, \alpha_j] = 1 \Leftrightarrow \alpha_i\alpha_j \in L$, potom $\forall v,w \in \Sigma^* : v \equiv_L w \Leftrightarrow$ ď riadky, prislúchajúce v a w sú rovnaké.

Definícia.

Nech $A = (K, \Sigma, \delta, q_0, F)$ Ak δ je definovaná $\forall q \in K$ $\forall a \in \Sigma$, í δ je *úplne* definovaná.

Označenie.

P $DKA_{\Sigma,n} = \{L \in R;\ \exists$ A o najviac n stavoch s úplne definovanou prechodovou funkciou, akceptujúci $L\}$.

Veta 8.2.4.1 (Myhill-Nerode).

1. Jazyk L je reg ď $\equiv_L$ x
2. $L \in DKA_{\Sigma,n} \Leftrightarrow$ relácia $\equiv_L$ má index $\leq n$.

Dôkaz.

Nech L g $A = (K, \Sigma, \delta, q_0, F)$. Nech index $\equiv_L > |K|$. Potom $\exists w_1..w_{|K|+1}\ \forall i \neq j \in 1..|K|+1 : \neg w_i \equiv_L w_j$. Al A má iba $|K|$ $\exists i \neq j \in 1..|K|+1 :$ w_i a w_j í

Obrátene nech index $\equiv_L = n$. Definujme $A' = (K', \Sigma, \delta', q_0', F')$, kde $K' = \{[w];\ w \in \Sigma^*\}$, $q_0 = [\varepsilon]$, $F = \{[w];\ w \in L\}$ a $\delta'([w],a) = [wa]$. A' má zjavne n „ "

í ť f í F' ď í
do L (práve vtedy patrí trieda, ktorú reprezentuje, do F') T $L \in DKA_{\Sigma,n}$.

Označenie.

$mindet_{\Sigma}(L) = \min\{n;\ L \in DKA_{\Sigma,n}\}$.

Dôsledok 8.2.4.1.

$mindet_{\Sigma}(L) =$ index relácie $\equiv_L$.

Dôkaz.

Tvrdenie vyplýva priamo z Myhill-Nerodovej vety (veta 8.2.4.1).

8.2.5. Aplikácie Myhill-Nerodovej vety.

Poznámka.

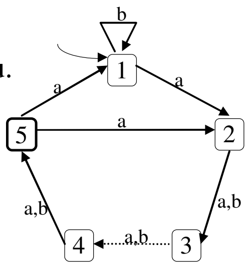
Myhill- ť ô g P e daný jazyk L a odhad indexu relácie $\equiv_L$ í Ď
ô
nedeterministickému, ako ukazuje nasledujúca veta.

Veta 8.2.5.1.

$\forall n \geq 1\ \exists L \subseteq \{a,b\}^* : mindet_{\{a,b\}^*}(L) = 2^n$.

Dôkaz.

P $L(A)$, kde automat A je zobrazený na obrázku.
í ô ť ď



8.2.6. Minimalizácia deterministického konečného automatu.

Definícia.

D $A_1 = (K_1, \Sigma, \delta_1, q_1, F_1)$ a $A_2 = (K_2, \Sigma, \delta_2, q_2, F_2)$ sú *izomorfné*, ak existuje bijekcia $h : K_1 \to K_2$ $h(q_1) = q_2$, $F_2 = \{h(q);\ q \in F_1\}$ a $\forall q \in K_1$ $\forall a \in \Sigma : \delta_2(h(q),a) = h(\delta_1(q,a))$.

Veta 8.2.6.1.

Jediné minimálne determini L sú automaty izomorfné s automatom, definovaným v dôkaze Myhill-Nerodovej vety.

Dôkaz.

Nedokazujeme.

8.2.7. Regulárne výrazy.

Definícia.

Nech Σ $\oplus$, $\otimes$ a $*$ $RV(\Sigma)$ *regulárnych výrazov* na Σ definujeme rekurzívne

1. $\varnothing \in RV(\Sigma)$, $\varepsilon \in RV(\Sigma)$
2. $\forall a \in \Sigma : a \in RV(\Sigma)$
3. $\forall \alpha, \beta \in RV(\Sigma) : \alpha \oplus \beta \in RV(\Sigma)$, $\alpha \otimes \beta \in RV(\Sigma)$, $\alpha^* \in RV(\Sigma)$.

Definícia.

Jazyk $L[\alpha]$, reprezentovaný regulárnym výrazom α definujeme nasledovne :

1. $[\varnothing] = \varnothing$, $[\varepsilon] = \{\varepsilon\}$
2. $\forall a \in \Sigma : [a] = \{a\}$
3. $\forall \alpha, \beta \in RV(\Sigma)$: $L[\alpha] = L_\alpha \wedge L[\beta] = L_\beta \Rightarrow L[\alpha \oplus \beta] = L_\alpha \cup L_\beta \wedge L[\alpha \otimes \beta] = L_\alpha L_\beta \wedge L[\alpha^*] = L_\alpha^*$.

Označenie.

H í ť P

ť í O $Min(\Sigma)$

ť P a $\mathbf{T} = \{[\alpha];\ \alpha \in RV(\Sigma)\}$.

Veta 8.2.7.1.

$Min(\Sigma) = \mathbf{T}$.

Dôkaz.

Tvrdenie vyplýva z definície regulárnych výrazov.

a) ť P $\mathbf{T}$

b) $\mathbf{T}$ ť P.

8.2.8. Uzáverové vlastnosti.

Veta 8.2.8.1.

R je uzavretá na zjednotenie a prienik.

Dôkaz.

Nech L_1 a L_2 sú regulárne jazyky; L_1 $A_1 = (K_1, \Sigma, \delta_1, q_1, F_1)$, L_2 $A_2 = (K_2, \Sigma, \delta_2, q_2, F_2)$. definujme automat $A = (K_1 \times K_2, \Sigma, \delta, (q_1, q_2), F)$, kde $(p,q) \in F \Leftrightarrow p \in F_1 \vee q \in F_2$ a $\delta((p,q), a) = (\delta_1(p,a), \delta_2(q,a))$. Tento automat vykonáva paralelne L_1 a L_2, í ď í ň Ľ ť $L(A) = L_1 \cup L_2$.

O A predpisom $(p,q) \in F \Leftrightarrow p \in F_1 \wedge q \in F_2$ dostaneme automat, akceptujúci $L_1 \cap L_2$.

Veta 8.2.8.2.

R je uzavretá na doplnok.

Dôkaz.

Nech L je regulárny jazyk, akceptovaný automatom $A = (K, \Sigma, \delta, q_0, F)$. Zámenou F a K-F dostávame automat pre doplnok.

Veta 8.2.8.3.

R ť

Dôkaz.

Nech L_1 a L_2 sú regulárne jazyky; L_1 $A_1 = (K_1, \Sigma, \delta_1, q_1, F_1)$, L_2 $A_2 = (K_2, \Sigma, \delta_2, q_2, F_2)$. definujme automat $A = (K_1 \cup K_2, \Sigma, \delta, q_1, F_2)$, kde $\delta = \delta_1 \cup \delta_2$ plus ε- $q \in F_1$ do q_2. Automat A ť $L_1 L_2$.

Veta 8.2.8.4.

R je uzavretá na uzáver.

Dôkaz.

K automatu pre L pridáme ε-prechody z F do q_0.

Veta 8.2.8.5 (Kleene).

$R = Min(\Sigma)$.

Dôkaz.

R ť P (z predchádzajúcich viet) $\Rightarrow Min(\Sigma) \subseteq R$. Obrátene nech L je regulárny jazyk, akceptovaný automatom $A = (K, \Sigma, \delta, q_0, F)$ D f $R_{i,j} = \{w \in \Sigma^*;\ \delta(q_i, w) \rightarrow q_j\}$. Zrejme $L = \bigcup_{i,j=1}^{|K|} R_{i,j}$ O $R_{i,j}^{k} = \{w \in R_{i,j};$ $q_1..q_k\}$. $\forall i,j \in 1..|K| : R_{i,j}^{0} \subseteq \Sigma \cup \{\varepsilon\}$ a to ť g A $\forall i,j \in 1..|K|\ \forall k \geq 1 : R_{i,j}^{k} = R_{i,j}^{k-1} \cup R_{i,k}^{k-1}.(R_{k,k}^{k-1})^*.R_{k,j}^{k-1}$ a ď RV ť $R \subseteq Min(\Sigma)$.

8.3. Bezkontextové gramatiky a zásobníkové automaty.

8.3.1. Bezkontextové gramatiky.

Definícia.

Gramatika $G = (N, T, P, \sigma)$ je *bezkontextová*, ak sú všetky jej pravidlá tvaru $A \to x$, kde $A \in N$, a $x \in (N \cup T)^*$. Jazyk L je *bezkontextový*, ak je generovaný bezkontextovou gramatikou. T x L_{CF}.

Definícia.

Bezkontextová gramatika G je *nevypúšťajúca*, ak neobsahuje pravidlá tvaru $X \to \varepsilon$.

Veta 8.3.1.1.

Nech L je bezkontextový jazyk, generovaný gramatikou G a neobsahujúci ε. Potom existuje ť g G', generujúca L.

Dôkaz.

Ak $\varepsilon \notin L$ $\sigma = \alpha_0 \underset{G}{\Rightarrow} \alpha_1 \underset{G}{\Rightarrow}^* \alpha_n = w \in T^*$ $\exists i \in 1..n : |\sigma| \leq |\alpha_i|$. Ak v G zameníme všetky pravidlá $\alpha_k \underset{G}{\Rightarrow} \alpha_{k+1}$, kde $|\alpha_k| > |\alpha_{k+1}|$ pravidlom $\alpha_k \underset{G}{\Rightarrow} \alpha_{k+j}$, kde $j = \min\{i \in k..n;\ |\alpha_k| \leq |\alpha_{k+i}|\}$, dostávame ť g g L.

Veta 8.3.1.2.

Nech L je bezkontextový jazyk, generovaný gramatikou G. Potom existuje bezkontextová gramatika G' bez pravidiel tvaru $X \Rightarrow Y$, generujúca L.

Dôkaz.

Triviálne.

8.3.2. Redukované gramatiky.

Definícia.

Bezkontextová gramatika G je *redukovaná* ď erminálu ť ť Proste G ť „ “

Veta 8.3.2.1.

x L existuje redukovaná gramatika, ktorá ho generuje.

Dôkaz.

Triviálne.

8.3.3. Najľavejšie odvodenie.

Definícia.

Odvodenie v bezkontextovej gramatike G je *najľavejšie* ľ í

Veta 8.3.3.1.

Nech L je bezkontextový jazyk generovaný gramatikou G a nech $w \in L$. Potom v G x ľ odvodenie w.

Dôkaz.

Triviálne.

8.3.4. Chomského a Greibachovej normálny tvar.

Definícia.

Bezkontextová gramatika G je v *chomského normálnom tvare*, ak všetky jej pravidlá sú tvaru $A \Rightarrow a$ a $A \Rightarrow BC$, kde $a \in T$ a $B, C \in N$.

Veta 8.3.4.1.

Nech L je bezkontextový jazyk bez ε, generovaný gramatikou G. Potom existuje bezkontextová gramatika G' v chomského normálnom tvare, generujúca L.

Dôkaz.

Pravidlo tvaru $A \Rightarrow Xx$, kde $x \in (N \cup T)$ a $X \in (N \cup T)^*$, nahradíme pravidlom $A \Rightarrow YZ$, kde y a Z sú nové neterm Ď $Z \Rightarrow x$ a pravidlo $Y \Rightarrow X$ ď í í istým spôsobom. Výsledkom je gramatika v chomského normálnom tvare.

Definícia.

Bezkontextová gramatika G je v *greibachovej normálnom tvare*, ak všetky jej pravidlá sú tvaru $A \Rightarrow a\alpha$, kde $a \in T$ a $\alpha \in N^*$.

Veta 8.3.4.1.

Nech L je bezkontextový jazyk, generovaný gramatikou G. Potom existuje bezkontextová gramatika G' v greibachovej normálnom tvare, generujúca L.

Dôkaz.

Gramatiku v chomského normálnom tvare upravíme na greibachovej normálny t tvaru ABC nahradíme pravidlom $A \Rightarrow aB_n..B_1C$, kde $aB_n..B_1$ ľ B prerušené v okamihu $B_{n+1} \Rightarrow a$.

8.3.5. Stromy odvodenia.

Definícia.

Nech G x g O ň aný strom s vrcholmi z $(N \cup T)^*$ nazývame *strom odvodenia* v G ň σ a β je synom α ď $\alpha \underset{G}{\Rightarrow} \beta$ ť listov stromu odvodenia T nazývame *výsledok* stromu odvodenia T w_T.

Veta 8.3.5.1.

Nech G je bezkontextová gramatika. $\forall w \in (N \cup T)^* : w \in L(G) \Leftrightarrow \exists T : w_T = w$.

Dôkaz.

Triviálne.

8.3.6. Pumpovacia lema.

Veta 8.3.6.1 (Pumpovacia).

Nech L je bezoknteхtový jazyk. Potom $\exists p,q \in N\ \forall z \in L : |z| > p \Rightarrow \exists u,v,w,x,y \in \Sigma^* : z = uvwxy \wedge |vwx| \le q \wedge vx \neq \varepsilon \wedge \forall i \ge 0 : uv^iwx^iy \in L$.

Dôkaz.

Nech G je bezkontextová gramatika v Greibachovej normálnom tvare, generujúca L-$\{\varepsilon\}$. Nech $|N| = k$. P $p = 2^{k-1}$ a $q = 2^k$ P T v G, ktorého výsledok t je dlhší ako p, obsahuje vetvu dlhšiu ako k. Nech P je najdlhšia cesta v T. Potom v G existuje neterminál, ktorý sa na P nachádza O A. Potom $A \underset{G}{\Rightarrow} vAx$ (predposledný výskyt) a $A \underset{G}{\Rightarrow} w$ (posledný výskyt), kde $v,w,x \in T^*$. O u „ ľ “ P a y výsledok „napravo“ od P. Potom $t = uvwxy$ $|vwx| \le 2^k$, í $\le k+1$, $vx \neq \varepsilon$ a úsek $A \underset{G}{\Rightarrow} vAx$ ô P ť ľ ľ $\forall i \ge 0 :$ $uv^iwx^iy \in L$.

Definícia.

Nech G je bezkontextová gramatika, $A \in N$. Pravidlo tvaru $A \Rightarrow \alpha$, kde $\alpha \in (N \cup T)^*$, nazývam *A-pravidlo*.

Lema 8.3.6.1.

Nech G je bezkontextová gramatika bez ε-pravidiel, nech $A \Rightarrow \alpha\beta\gamma$ je jej A-pravidlo a nech $\{\beta \Rightarrow \beta_i\}_{i=1..n}$ sú všetky β-pravidlá. Potom nahradením pravidla $A \Rightarrow \alpha\beta\gamma$ pravidlami $\{\beta \Rightarrow \alpha\beta_i\gamma\}_{i=1..n}$ generujem rovnaký jazyk.

Dôkaz.

Zrejmé.

Lema 8.3.6.2.

Nech G je bezkontextová gramatika bez ε-pravidiel, nech $\{A\Rightarrow A\alpha_i\}_{i=1..n}$ sú všetky A-pravidlá s A na $\{A\Rightarrow\beta_i\}_{i=1..m}$ sú všetky ostatné A-pravidlá. Potom nahradením A-pravidiel pravidlami $\{A\Rightarrow\beta_i, A\Rightarrow\beta_i\beta, \beta\Rightarrow\alpha_i, \beta\Rightarrow\alpha_i\beta\ \}_{i=1..n.m}$ generujem rovnaký jazyk.

Dôkaz.

Zrejmé.

8.3.7. Zásobníkové automaty.

Definícia.

Usporiadanú sedmiticu $A = (K, \Sigma, T, \delta, q_0, Z_0, F)$ nazývame *zásobníkový automat*. K, Σ, δ, q_0 a F T je zásobník a Z_0 zásobníka.

Definícia.

Konfigurácia zásobníkového automatu je usporiadaná trojica (q, w, γ), kde $q\in K$, w je ešte nepr í γ je stav zásobníka.

Poznámka.

Z í ď í

Veta 8.3.7.1.

Rozoznávanie koncovým stavom a prázdnym zásobníkom sú rovnako silné.

Dôkaz.

Triviálne.

Veta 8.3.7.2.

P bezkontextový jazyk L existuje zásobníkový automat A, ktorý ho akceptuje.

Dôkaz.

Nech $G = (N, T_G, P, \sigma)$ je bezkontextová gramatika v greibachovej normálnom tvare, generujúca L. Definujme zásobníkový automat $A = (K, \Sigma, T, \delta, q, Z_0, F)$, kde $K = \{q\}$, $\Sigma = N\cup T_G$, $T = N$, $Z_0 = \sigma$, $F = \emptyset$ a $\delta(q, aw, A\gamma) = (q, w, B_1..B_n\gamma) \Leftrightarrow A \underset{G}{\Rightarrow} aB_1..B_n \in P$. Ak automat A akceptuje vstupné slová prázdnym zásobníkom, tak A ľ G a teda zjavne $L(A)\subseteq L$ A ň $L\subseteq L(A)$ í í í L nedeterministický A „vie“, ktorým pravidlom z G bolo toto í f ť

Veta 8.3.7.3.

P í A je $L(A)$ bezkontextový jazyk.

Dôkaz.

Nech $A = (K, \Sigma, T, \delta, q_0, Z_0, F)$ je zásobníkový automat, akceptujúci vstupné slová prázdnym zásobníkom. Definujme gramatiku $G = (N, T_G, P, \sigma)$, kde $N = \{[q,B,p];\ p,q\in K,\ B\in T\}\cup\{\sigma\}$, $T_G = \Sigma$ a $P = \{\sigma\Rightarrow[q_0, Z_0, q];\ q\in K\}\cup\{[q, B, q_{n+1}]\Rightarrow a[q_1, B_1, q_2]..[q_n, B_n, q_{n+1}];\ (q_1, w, B_1..B_n\gamma) \in \delta(q, aw, B\gamma)\}$. Táto gramatika sleduje A „ “ $L(G) = L(A)$.

8.3.8. Uzáverové vlastnosti.

Veta 8.3.8.1.

L_{CF} nie je uzavretá na prienik a doplnok.

Dôkaz.

$\{a^ib^ic^i;\ a,b,c\in T,\ i\in N) \notin L_{CF}$, ale $\{a^ib^ic^j;\ a,b,c\in T,\ i,j\in N)$ a $\{a^jb^ic^i;\ a,b,c\in T,\ i,j\in N) \in L_{CF} \Rightarrow L_{CF}$ nie je uzvretá na prienik. Navyše $L_1\cap L_2 = \overline{\overline{L_1}\cup\overline{L_2}} \Rightarrow$ ak by L_{CF} ť

Veta 8.3.8.2.

L_{CF} je uzavretá na prienik s R.

Dôkaz.

Nech $A = (K, \Sigma, T, \delta, q_0, Z_0, F)$ je zásobníkový automat a $A' = (K', \Sigma, \delta', q_0', F')$ Potom zásobníkový automat $A'' = (K \times K', \Sigma, T, \delta'', (q_0, q_0'), Z_0, F \times F')$, kde $\delta''((p,q), aw, \gamma) = \{((p',q'), w, \gamma');$ $\delta(p,a) = p' \wedge (q', w, \gamma') \in \delta'(q, aw, \gamma)\}$ rozpoznáva jazyk $L(A) \cap L(A')$.

8.4. Frázové gramatiky a Turingove stroje.

8.4.1. Turingove stroje.

Definícia.

Usporiadanú šesticu $A = (K, \Sigma, \mathbf{T}, \delta, q_0, F)$ nazývame *Turingov stroj*. K stavov, Σ je vstupná abeceda $\mathbf{T}$ je pásková abeceda (a teda $\Sigma \subseteq \mathbf{T}$), $\delta : K \times \mathbf{T} \rightarrow K \times \mathbf{T} - \{B\} \times \pm 1$ je prechodová funkcia, q_0 F

Definícia.

Konfigurácia Turingovho stroja je usporiadaná trojica (q, w, i), kde $q \in K$, $w \in \mathbf{T}^*$ a $i \in N$ je pozícia hlavy na páske.

Poznámka.

P T g

8.4.2. Modifikované Turingove stroje.

Veta 8.4.2.1.

Viacstopové Turingove stroje sú ekvivalentné štandardným.

Dôkaz.

Na pásku sa nezapisujú symboly, ale vektory symbolov.

Veta 8.4.2.2.

T g

Dôkaz.

P í

Veta 8.4.2.3.

Viacpáskové Turingove stroje sú ekvivalentné štandardným.

Dôkaz.

k pások = $2k$ ô

Definícia.

Stupeň nedeterministickosti Turingovho stroja je $k = \max\{|\delta(q, w, i)|;\ q \in K,\ w \in \mathbf{T}^*,\ i \in N\}$.

Veta 8.4.2.4.

Nedeterministické Turingove stroje sú ekvivalentné deterministickým.

Dôkaz.

Nech k ň T g T. Deterministický T' ť ľ í í

8.4.3. Rekurzívne vypočítateľné jazyky.

Definícia.

Jazyk, rozpoznávaný Turingovým strojom, nazývame *rekurzívne vypočítateľný*. Ak T g í *rekurzívny*. T í í ľ L_{RE}.

8.4.4. Frázové gramatiky.

Definícia.

Gramatika $G = (N, T, P, \sigma)$ je *frázová* x f

Veta 8.4.4.1.

f g g í í ľ

Dôkaz.

Dvojpáskový nedeterministický *TS* g ť g

Veta 8.4.4.2.

P T g *A* existuje gramatika *G* $L(A) = L(G)$.

Dôkaz.

Gramatika postupuje v troch krokoch. Najprv vygener í P g ť T $vqaw \to vbqw$, ktoré simulujú posun hlavy automatu v stave q í í a P ľ P í ť v ť g ä í T ť í p je akceptovací stav. Aby sme tera ť í „ “ ó ť P „ “

Definícia.

Frázová gramatika je v *normálnom tvare*, ak má všetky pravidlá tvaru $\alpha \to \beta$ a $A \to a$, kde $\alpha, \beta \in N^*$, $A \in N$ a $a \in T$.

Lema 8.4.4.1.

P ý rekurzívny jazyk existuje frázová gramatika v normálnom tvare, ktorá ho generuje.

Dôkaz.

Triviálne.

8.4.5. Churchova téza.

Hypotéza.

g ľ T g

8.5. Kontextové gramatiky a lineárne ohraničené automaty.

8.5.1. Kontextové gramatiky.

Definícia.

Gramatika $G = (N, T, P, \sigma)$ je *kontextová*, ak pre všetky jej pravidlá $x \Rightarrow y$, kde $x, y \in (N \cup T)^*$ platí $|x| \leq |y|$. Jazyk L je *kontextový*, ak je generovaný kontextovou gramatikou. T x L_{CS}.

Definícia.

Kontextová gramatika G je v *normálnom stave*, ak má všetky pravidlá tvaru $\alpha \to \beta$ a $A \to a$, kde $\alpha, \beta \in N^*$, $|\alpha| \leq |\beta|$, $A \in N$ a $a \in T$.

Lema 8.5.1.1.

P x *L* existuje kontextová gramatika *G* v normálnom tvare, ktorá ho generuje.

Dôkaz.

Triviálne.

8.5.2. Lineárne ohraničené automaty.

Definícia.

Turingov stroj, pracujúci v priestore, vyhradenom vstupom (alebo iným spôsobom fixne) *lineárne ohraničený* automat.

Veta 8.5.2.1.

P *A* existuje kontextová gramatika *G* $L(A) = L(G)$.

Dôkaz.

Dô 8 4 4 2 x ť g „ “ ť (osobitných páska) g í ť

8.5.3. Chomského hierarchia gramatík.

Lema 8.5.3.1.

$R \subseteq L_{CF}$.

Dôkaz.

Triviálne.

Lema 8.5.3.2.

$\exists L \in L_{CF} \setminus R$.

Dôkaz.

Definujme $L = \{a^n b^n;\ n \in N\}$. Definujme gramatiku $G = (\{S\}, \{a,b\}, P, S\}$, kde $P = \{S \to aSb|ab\}$ Ľ ť $L(G) = L$ ď G je zjavne bezkontextová, $L \in L_{CF}$. Nech $L \in R$ P ľ - Nerodovej vety (veta 8.2.4.1) je index relácie $\equiv_L$ n U $0^0=\varepsilon, 0^1, .. , 0^n$. Nech $\exists i \neq j \in 0..n$. Po $x=1^i$. Potom $0^i1^i \in L \wedge 0^j1^i \notin L \Rightarrow 0^i \equiv_L 0^j$. Našli sme $n+1$ tried ekvivalencie relácie $\equiv_L$ spor.

Definícia.

$L_{ECS} = \{L \cup \{\varepsilon\};\ L \in L_{CS}\}$ nazývame triedou *rozšírených* kontextových jazykov.

Lema 8.5.3.3.

$L_{CF} \subseteq L_{ECS}$.

Dôkaz.

P ľ 8 3 4 1 x x L bezkontextová gramatika v chomského normálnom tvare, ktorá generuje L - ε g x $\Rightarrow L\text{-}\varepsilon \in L_{CS} \Rightarrow L \in L_{ECS}$.

Lema 8.5.3.4.

$\exists L \in L_{ECS} \setminus L_{CF}$.

Dôkaz.

Definujme $L = \{a^n b^n c^n;\ n \in N\}$. Definujme gramatiku $G = (\{S,B,C\}, \{a,b,c\}, P, S)$, kde $P = \{S \to aSBC|aBC, CB \to BC, aB \to ab, bB \to bb, bC \to bc, cC \to cc\}$ Ľ ť $L(G) = L$ ď G je zjavne kontextová, $L \in L_{ECS}$. Nech $L \in L_{CF}$ P ľ (8 3 6 1) $\exists p,q \in N\ \forall z \in L$ $\exists$jeho rozdelenie na $uvwxy$ ĺň P $l = \max\{p,q\}$ a $a^l b^l c^l$ ď ť vwx í ť q ô ť ô Ale potom $\forall i \in N : uv^iwx^iy$ ô a- c- je spor.

Lema 8.5.3.5.

$L_{ECS} \subseteq L_{REC}$.

Dôkaz.

P x x ĺ g n P x g ĺ n í áno/nie podľ g g

Definícia (Kódovanie gramatík).

Nech T je abeceda, $|T| = k$. nech $G = (N, T, P, S)$ je gramatika, kde $|N| = l$. G kódujeme binárne nasledovným spôsobom : $10^1..10^k$ sú kódy terminálov, $10^{k+1}..10^{k+6}$ sú kódy postune {**,** , $\to$, **{ , } , (,)**) a vyššie mocniny dvojky kódujú neterminály.

Kontextový jazyk L definujeme ako $C1E$, kde C je kód kontextovej gramatiky, generujúcej L a $E = 0$ ak $\varepsilon \in L$ a 00, ak $\varepsilon \notin L$.

Lema 8.5.3.6.

O T g ť g x g touto terminálnou abecedou.

Dôkaz.

ó f í

Definícia.

Kanonická ť g í T ť gramatík nad terminálnou abecedou T ó ť binárnou abecedou.

Lema 8.5.3.7.

Ex g g ť x g í abecedou T.

Dôkaz.

Algoritmus generuje kanonickú postup ť kóduje kontextovú gramatiku, a ak áno, zaradí ju do postupnosti.

Definícia.

Nech $\{G_i\}_{i\in N}$ ť g í T. Jazyk $L_{DIAG} = \{w_i;\ w_i \notin G_i\}$ nazývame *diagonálny* jazyk nad abecedou T.

Lema 8.5.3.8.

$\exists L \in L_{REC} \setminus L_{ECS}$.

Dôkaz.

L_{DIAG} nad kontextovými gramatikami zjavne nie je kontextový. Ale pre dané $w \in T$ í ť index v kanonickej postupnosti (nech je to i) g ť ó i-tej gramatiky G_i postupnosť kontextových gramatík nad terminálnou abecedou T ť í ť w do $L(G_i)$ a tým aj do $L \Rightarrow$ L je rekurzívny.

Lema 8.5.3.9.

$L_{REC} \subseteq L_{RE}$.

Dôkaz.

Triviálne.

Lema 8.5.3.10.

$\exists L \in L_{RE} \setminus L_{REC}$.

Dôkaz.

V í T g ľ

Dôsledok 8.5.3.1 (Chomského hierarchia).

$R \subset L_{CF} \subset L_{ECS} \subset L_{REC} \subset L_{RE}$.

Dôkaz.

Tvrdenie je priamym dôsledkom predchádzajúcich liem.

8.5.4. Uzáverové vlastnosti.

Veta 8.5.4.1.

L_{CS}, L_{REC}, L_{RE} ť (L_{CS} iba na kladný uzáver) a zrkadlový obraz.

Dôkaz.

- Zjednotenie : Nech $G_1 = (N_1, T_1 P_1, S_1)$ a $G_2 = (N_2, T_2, P_2, S_2)$ sú kontextové (rekurzívne, frázové) gramatiky a nech $N_1 \cap N_2 = T_1 \cap T_2 = \emptyset$. Definujme gramatiku $G = (N, T, P, S)$, kde $S \notin N_1 \cup N_2 \cup T_1 \cup T_2$, $N = N_1 \cup N_2 \cup S$, $T = T_1 \cup T_2$ a $P = P_1 \cup P_2 \cup \{S \rightarrow S_1 | S_2\}$. G patrí do rovnakej triedy chomského hierarchie ô g ľ ť g $L(G_1) \cup L(G_2)$.

- Z ť : $G_1 = (N_1,T_1P_1,S_1)$ a $G_2 = (N_2,T_2,P_2,S_2)$ sú kontextové (rekurzívne, frázové) gramatiky a nech $N_1 \cap N_2 = T_1 \cap T_2 = \emptyset$. Definujme gramatiku $G = (N,T,P,S)$, kde $S \notin N_1 \cup N_2 \cup T_1 \cup T_2$, $N = N_1 \cup N_2 \cup S$, $T = T_1 \cup T_2$ a $P = P_1 \cup P_2 \cup \{S \to S_1S_2\}$. G patrí do rovnakej triedy chomského hierarchie ô g ľ ť g ť $L(G_1)$ a $L(G_2)$.
- Uzáver : Nech $G = (N,T,P,S)$ je kontextová (rekurzívna, frázová) gramatika. Pridaním pravidla $S \to SS$ vytvoríme gramatiku, prislúchajúcu do rovnakej triedy chomského hierarchie ako pôvodná gramatika, ak genrujúcu uzáver $L(G)$.
- Zrkadlový obraz : Nech $G = (N,T,P,S)$ je kontextová (rekurzívna, frázová) gramatika. Potom gramatika $G' = (N,T,P',S)$, kde $P' = \{\alpha^R \to \beta^R;\ \alpha \to \beta \in P\}$, patrí do rovnakej triedy chomského hierarchie ako pôvodná gramatika a generuje zrkadlový obraz $L(G)$.

Veta 8.5.4.2.

L_{CS} a L_{REC} sú uzavreté na komplement.

Dôkaz.

Szelepcsenyho veta dokazuje tvrdenie pre L_{CS}. Pre L_{REC} je tvrdenie triviálne.

Veta 8.5.4.3.

L_{CS} sú uzavreté na prienik.

Dôkaz.

Triviálne.

8.6. Aplikácie Turingových strojov.

8.6.1. Turingove stroje, počítajúce celočíselné funkcie.

- $f(k_1, .., k_n)$ reprezentujeme vstupom $0^{k_1}1\Lambda 10^{k_i}1\Lambda 10^{k_n}$ f 0^m, ak $f(k_1,..,k_n) = m$.

8.6.2. Turingove stroje ako enumerátory.

Definícia.

Enumerátor je Turingov stroj bez vstupu so špeciálnou výstupnou páskou, na ktorej sa ť V ú slová oddelené špeciálnym symbolom ľ (#). Všetky slová, ktoré takýto enumerátor A vyprodukuje, tvoria $L(A)$.

Veta 8.6.2.1.

Jazyk L í í ľ ď x A, ktorý ho generuje.

Dôkaz.

Nech L je akceptovaný TS A'. Ak ekvivalentný DTS (konštrukcia v dôkaze vety 8.4.2.4) upravíme na enumerátor A, produkujúci všetky slová akceptované A', potom $L(A) = L(A')$. Obrátene TS enumerátora a porovnáva slová so vstupom.

8.7. Rozhodnuteľnosť.

8.7.1. Niektoré štandardné problémy.

Definícia.

Problém *Find* – pre danú rekurzívnu gramatiku a vstup w nad jej terminálnou abecedou ť ľ w v G a NIE, ak $w \notin L(G)$.

Problém *Inclusion* – pre danú rekurzívnu gramatiku a vstup w nad jej terminálnou ť Á O/ IE ľ $w \in / \notin L(G)$.

Veta 8.7.1.1.

Find ľ $\Leftrightarrow$ *Inclusion* ľ

Dôkaz.

Triviálne.

Veta 8.7.1.2 (Post).

L í ď L í í ľ

Dôkaz.
Triviálne.

Definícia.
Problém *Accept* – A akceptuje vstup w.
Problém *Stop* – A zastaví na vstupe w.

Lema 8.7.1.1.
Ak $L \subseteq \{0,1\}^*$, tak $\exists$TS A s páskovou abecedou $\mathbf{T} = \{0,1,B\}$, akceptujúci L.

Dôkaz.
ó ť k-ticami abecedy $\mathbf{T} = \{0,1,B\}$ (k je pre daný jazyk pevné a teda k- ä ť)

Definícia (Kódovanie Turingových strojov).
Nech $A = (\{q_1, .., q_n\}, \{0,1\}, \{0,1,B\}, \delta, q_1, B, \{q_2\})$ je Turingov stroj, ktorého δ-funkcia obsahuje k pravidiel. O $0 = x_1$, $1 = x_2$ a $B = x_3$. Pravidlo $\delta(q_i, x_j) = (q_m, x_p, \pm 1)$ ó ť $0^i10^j10^m10^p10^e$, kde $e = 1$ ľ 1 doprava. Automat A potom kódujeme $111C_111 .. 11C_k111$, kde C_i je kód i-teho pravidla. Kód automatu A o $\langle A \rangle$.

Ak w je vstup automatu A, tak dvojicu A, w ó ť í $\langle A \rangle w$ í $\langle A, w \rangle$.

Lema 8.7.1.2.
O danom slove w ť ó TS

Dôkaz.
Rovnako ako pri kódovaní gramatík.

Lema 8.7.1.3.
Existuje algoritmus, generujúci kódy TS v kanonickom poradí.

Dôkaz.
Rovnako ako pri kódovaní gramatík.

Definícia.
Nech A_i je i-ty TS v kanonickej postupnosti. $L_{DIAG} = \{w_i;\ w_i \notin L\}A_i|\}$.

Veta 8.7.1.3.
$L_{DIAG} \notin L_{REC}$.

Dôkaz.
Triviálne.

Dôsledok 8.7.1.1.
Doplnok k $L_{DIAG} \notin L_{RE}$.

Dôkaz.
Triviálne.

Definícia.
$L_{UNIV} = \{\langle A, w \rangle;\ w \in L(A)\}$.

Veta 8.7.1.4.
$L_{UNIV} \notin L_{REC}$.

Dôkaz.
Triviálne.

Dôsledok 8.7.1.2.
Problémy *Accept* a *Stop* ľ

Dôkaz.
Trviálne.

8.7.2. Vlastnosti rekurzívnych jazykov.

Definícia.

φ í *vlastnosť* rekurzívnych H í í L ť φ, ak $L\in\varphi$ V ť φ sa nazýva *triviálna*, ak $\varphi\in\{\emptyset, L_{REC}\}$ J ť φ je jazyk $L_\varphi = \{\langle A\rangle;\ L(A)\in\varphi\}$.

Veta 8.7.2.1 (Rice).

Ž ť í ľ

Dôkaz.

Nech φ ľ ť í A_φ í $L(A_\varphi) = L_\varphi$. Nech $\langle A,w\rangle\in L_{UNIV}$ U A', pracujúci nasledovným spôsobom : najprv (ignorujúc svoj vstup x) simuluje A na vstupe w. Ak $w\notin L(a)$, $x\notin L(A')$. Ak $w\in L(A)$ (ď $\langle A,w\rangle\in L_{UNIV}$), $x\in L(A')\Leftrightarrow x\in L(A_\varphi)$. Potom $\langle A,w\rangle\in L_{UNIV} \Leftrightarrow L(A')\in\varphi$, í ľ φ ľ ť *Accept*

Dôsledok 8.7.2.1.

Pre daný TS A ľ : $L(A) = \emptyset$, $L(A) = L_{REC}$, $L(A)\in L_{REC}$, $L(A)\in L_{CF}$.

Dôkaz.

Tvrdenie priamo vyplýva z predchádzajúcej vety.

8.7.3. Postov korešpondenčný problém.

Definícia (Postov korešpondenčný problém).

Nech Σ je abeceda a $A,B\subseteq\Sigma^*$. Postov problém priradenia pre jazyky A a B je problém existencie postupnosti $\{u_i\in A\}_{i\in 1..n}$ a $\{v_i\in B\}_{i\in 1..n}$: $u_1u_2..u_n = v_1v_2..v_n$ T ť a nazýva *riešenie* problému pre dané jazyky.

Veta 8.7.3.1.

P ľ

Dôkaz.

Nedokazujeme.

8.7.4. Nerozhodnuteľné problémy pre kontextové a bezkontextové jazyky.

Veta 8.7.4.1.

Problém prázdnosti kontextového jazyka je ľ

Dôkaz.

Ak pre dané jazyky A a B í A, akceptujúci vstup x ď x í P A a B ľ $L(A)$ by vyplývala rozho ľ ť P

Veta 8.7.4.2.

P x ľ

Dôkaz.

Nech $A = \{u_1, .. , u_n\}$, $B = \{v_1, .. , v_n\}$ a nech $a_1, .. , a_n$ sú slová nad inou abecedou U g $G_A = \{S\to u_iSa_i|u_ia_i;\ i\in 1..n\}$ a $G_B = \{S\to v_iSa_i|v_ia_i;\ i\in 1..n\}$. $L(G_A)\cap L(G_B)\neq\emptyset \Leftrightarrow$ P má pre jazyky A a B Z ľ x ľ ť P ného problému.