Matematička teorija računarstva

Predavanja 12

Matko Botinčan

PMF – Matematički odjel

17.01.2007.

Potisni automati

- Potisni automat (**PDA**)
  - konačni automat koji može brojati bez ograničenja
  - posjeduje neograničenu količinu memorije u obliku stoga
  - ograničenje: vidljiv je samo simbol na vrhu stoga
- **PDA** = **NKA** + stog
- Teorem ekvivalencije: **PDA** = **KS**
- Deterministički **PDA** $\subsetneq$ nedeterministički **PDA**

Važnost **PDA** — parseri za **KS** jezike:

- Sintaksa programskog jezika tipično se zadaje u formi **KS** gramatike
- Naivni algoritam (top-down + backtracking) – eksponencijalan :-(
- Sofisticiraniji algoritmi (Cocke-Younger-Kasami i Earley):
  - prethodna konverzija gramatike u Chomskyjevu normalnu formu (produkcije imaju oblik $A \to BC$ ili $A \to a$)
  - dinamičko programiranje: za dani string $a_1 \dots a_n$ konstruira se $n \times n$ tabela T t.d. T_{ij} predstavlja skup varijabli iz kojih se može generirati $a_i \dots a_j$
  - složenost: $\mathcal{O}(n^3)$:-(

- Za opis programskih jezika u praksi su dostatne potklase klase **KS** gramatika
- top-down parsiranje – prediktivne rekurzivno silazne gramatike **LL**(k)
  - složenost: $\mathcal{O}(n^k)$:-|
- bottom-up parsiranje – **LR** i **LALR** klasa gramatika
  - "linearna" složenost :-)
  - $\bigcup_{k \geq 0}$ **LL**$(k) \subsetneq$ **LR**
  - shift-reduce algoritam – koristi potisne automate
  - **LR** $\approx$ deterministički **PDA**

PDA tranzicija:

- čitanje ulaznog simbola s trake
- prelazak u novo stanje
- zamjena simbola na vrhu stoga bilo kojim nizom znakova ($\rightarrow$ push ili pop)

Definicija 1

*Potisni automat (**PDA**) $\mathcal{A} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ sastoji se od slijedećih komponenti:*

- Q — *konačan skup stanja*
- Σ — *alfabet ulaznog niza znakova*
- Γ — *alfabet stoga*
- $\delta \colon Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma \rightarrow \mathscr{P}(Q \times \Gamma^*)$ — *tranzicijska funkcija*
- $q_0 \in Q$ — *početno stanje*
- $Z_0 \in \Gamma$ — *početni simbol na stogu*
- $F \subseteq Q$ — *skup završnih stanja*

Tranzicijska funkcija δ uzima kao argument (q, a, A), gdje je:

- $q \in Q$
- $a \in \Sigma \cup \{\varepsilon\}$
- $X \in \Gamma$

Vrijednost od $\delta(q, a, X)$ je konačan skup parova (p, γ), gdje je p novo stanje, a $\gamma \in \Gamma^*$ niz simbola koji zamjenjuje X:

- $\gamma = \varepsilon \rightarrow$ pop()
- $\gamma = X \rightarrow$ stog ostaje nepromjenjen
- $\gamma = YZ \rightarrow$ zamjena X sa Z, te push(Y)

Primjeri **PDA**

Primjer:

Odredite **PDA** koji prepoznaje jezik $L = \{0^n 1^n \mid n \geq 0\}$.

Primjer:

Odredite **PDA** koji prepoznaje jezik $L = \{ww^T \mid w \in \{0,1\}^*\}$.

Primjer:

Neka je $\mathcal{G}$ gramatika koja opisuje korektne sekvence **if**-**else** naredbi definirana na slijedeći način:

$$S \rightarrow \varepsilon \mid SS \mid \mathbf{i}S \mid \mathbf{i}S\mathbf{e}$$

Odredite **PDA** koji prepoznaje komplement jezika $L(\mathcal{G})$.

Konfiguracija **PDA** $\mathcal{A} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je (q, w, γ), pri čemu je:

- $q \in Q$ — stanje
- $w \in \Sigma^*$ — preostala riječ na traci
- $\gamma \in \Gamma^*$ — riječ na stogu

Relacija $\vdash$ na skupu konfiguracija definirana je na slijedeći način $(\forall w \in \Sigma^*, \forall \beta \in \Gamma^*)$:

$$(p, \alpha) \in \delta(q, a, X) \quad \Leftrightarrow \quad (q, aw, X\beta) \vdash (p, w, \alpha\beta)$$

$\vdash^* \equiv$ refleksivno i tranzitivno zatvorenje od $\vdash$

Lema 1

$\forall w \in \Sigma^*, \forall \gamma \in \Gamma^*$

$$(q, x, \alpha) \vdash^* (p, y, \beta) \quad \Rightarrow \quad (q, xw, \alpha\gamma) \vdash^* (p, yw, \beta\gamma)$$

Napomena 1

- *Što predstavljaju slučajevi kada je $\gamma = \varepsilon$, odnosno $w = \varepsilon$?*
- *Obrat leme ne vrijedi.*

Lema 2

$\forall w \in \Sigma^*$

$$(q, xw, \alpha) \vdash^* (p, yw, \beta) \quad \Rightarrow \quad (q, x, \alpha) \vdash^* (p, y, \beta)$$

Neka je $\mathcal{A} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ **PDA**.

- Jezik prihvaćen praznim stogom:
  $L_E(\mathcal{A}) = \{w \in \Sigma^* \mid (q_0, w, Z_0) \vdash^* (q, \varepsilon, \varepsilon)$ za neko $q \in Q\}$
- Jezik prihvaćen završnim stanjem:
  $L_F(\mathcal{A}) = \{w \in \Sigma^* \mid (q_0, w, Z_0) \vdash^* (q_F, \varepsilon, \gamma)$ za neko $\gamma \in \Gamma^*$ i $q_F \in F\}$

Teorem 1

- *Za svaki* **PDA** $\mathcal{A}$ *postoji* **PDA** $\mathcal{A}'$ *t.d.* $L_E(\mathcal{A}) = L_F(\mathcal{A}')$.
- *Za svaki* **PDA** $\mathcal{A}$ *postoji* **PDA** $\mathcal{A}'$ *t.d.* $L_F(\mathcal{A}) = L_E(\mathcal{A}')$.