Příklad

$\mathcal{M} = (\{q_0, p, f\}, \{a, b\}, \{A, B, Z\}, \delta, q_0, Z, \{f\})$, kde

$$
\begin{aligned}
\delta(q_0, a, Z) &= \{(q_0, AZ)\}\\
\delta(q_0, a, A) &= \{(q_0, AA)\}\\
\delta(q_0, a, B) &= \{(q_0, AB)\}\\
\delta(q_0, b, Z) &= \{(q_0, BZ)\}\\
\delta(q_0, b, A) &= \{(q_0, BA)\}\\
\delta(q_0, b, B) &= \{(q_0, BB)\}\\
\delta(q_0, \varepsilon, Z) &= \{(p, Z)\}\\
\delta(q_0, \varepsilon, A) &= \{(p, A)\}\\
\delta(q_0, \varepsilon, B) &= \{(p, B)\}\\
\delta(p, a, A) &= \{(p, \varepsilon)\}\\
\delta(p, b, B) &= \{(p, \varepsilon)\}\\
\delta(p, \varepsilon, Z) &= \{(f, Z)\}
\end{aligned}
$$

Příklad

$\mathcal{M} = (\{q_0\}, \{a, b\}, \{Z, A\}, \delta, q_0, Z, \emptyset)$, kde

$\delta(q_0, a, Z) = \{(q_0, A)\}$
$\delta(q_0, a, A) = \{(q_0, AA)\}$
$\delta(q_0, b, A) = \{(q_0, \varepsilon)\}$

Ekvivalence dvou způsobů akceptování

Věta 3.39. Pro každý jazyk L platí:
$L = L(\mathcal{N})$ pro nějaký PDA $\mathcal{N} \iff L = L_e(\mathcal{M})$ pro nějaký PDA $\mathcal{M}$

Důkaz. koncový stav $\Longrightarrow$ prázdný zásobník

Intuice:

K danému $\mathcal{N}$ zkonstruujeme $\mathcal{M}$ simulující jeho činnost.
Vejde-li $\mathcal{N}$ do koncového stavu, $\mathcal{M}$ se nedeterministicky rozhodne

- pokračovat v simulaci automatu $\mathcal{N}$ **nebo**
- přejít do nově přidaného stavu q_ε, v němž vyprázdní zásobník.

Komplikace:

Řešení: Před zahájením simulace bude u $\mathcal{M}$ na dně zásobníku nový symbol, který nedovolíme odstranit jinde, než ve stavu q_ε.

Konstrukce: Nechť $\mathcal{N} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$.

Klademe $\mathcal{M} = (Q \cup \{q_0', q_\varepsilon\}, \Sigma, \Gamma \cup \{Z'\}, \delta', q_0', Z', \emptyset)$,
kde $Z' \notin \Gamma$, $q_0', q_\varepsilon \notin Q$ a δ' je definována takto:

- $\delta'(q_0', \varepsilon, Z') = \{(q_0, Z_0 Z')\}$

- jestliže $\delta(q, a, Z)$ obsahuje (r, γ), pak $\delta'(q, a, Z)$ obsahuje (r, γ)
- $\delta'(q, \varepsilon, Z)$ obsahuje (q_ε, Z)
  pro všechny $q \in F$ a $Z \in \Gamma \cup \{Z'\}$
- $\delta'(q_\varepsilon, \varepsilon, Z) = \{(q_\varepsilon, \varepsilon)\}$
  pro všechny $Z \in \Gamma \cup \{Z'\}$

Korektnost:

prázdný zásobník $\Longrightarrow$ koncový stav

Intuice:

K danému $\mathcal{M}$ zkonstruujeme $\mathcal{N}$ simulující jeho činnost.

- $\mathcal{N}$ si před simulací přidá na dno zásobníku nový symbol.
- Je-li $\mathcal{N}$ schopen číst tento symbol (tj. zásobník automatu $\mathcal{M}$ je prázdný) tak $\mathcal{N}$ přejde do nově přidaného stavu q_f, který je koncovým stavem.

Konstrukce: Nechť $\mathcal{M} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, \emptyset)$.

Klademe $\mathcal{N} = (Q \cup \{q_0', q_f\}, \Sigma, \Gamma \cup \{Z'\}, \delta', q_0', Z', \{q_f\})$,
kde $Z' \notin \Gamma$, $q_0', q_f \notin Q$ a δ' je definována takto:

- $\delta'(q_0', \varepsilon, Z') = \{(q_0, Z_0 Z')\}$
- jestliže $\delta(q, a, Z)$ obsahuje (r, γ), pak $\delta'(q, a, Z)$ obsahuje (r, γ)
- $\delta'(q, \varepsilon, Z') = \{(q_f, \varepsilon)\}$
  pro všechny $q \in Q$ $\square$

Grafická reprezentace konfigurací

Rozšířený zásobníkový automat

Definice 3.44. Rozšířený zásobníkový automat je sedmice $\mathcal{R} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, kde

- všechny symboly až na δ mají tentýž význam jako v definici PDA,
- δ je zobrazením
  z konečné podmnožiny množiny $Q \times (\Sigma \cup \{\varepsilon\}) \times \Gamma^*$
  do konečných podmnožin množiny $Q \times \Gamma^*$.

Pojmy konfigurace a akceptovaný jazyk (koncovým stavem, prázdným zásobníkem) zůstávají beze změny. **Krok výpočtu** $\underset{\mathcal{R}}{\vdash}$ definujeme takto:

$$(p, aw, \gamma_1\alpha) \underset{\mathcal{R}}{\vdash} (q, w, \gamma_2\alpha) \overset{def}{\Longleftrightarrow} \exists(q, \gamma_2) \in \delta(p, a, \gamma_1) \text{ pro } a \in \Sigma \cup \{\varepsilon\}$$

Příklad

$\mathcal{R} = (\{q_0, p, f\}, \{a, b, c, d\}, \{A, B, Z\}, \delta, q_0, Z, \{f\})$, kde

$$
\begin{aligned}
\delta(q_0, a, \varepsilon) &= \{(q_0, A)\} \\
\delta(q_0, b, \varepsilon) &= \{(q_0, B)\} \\
\delta(q_0, \varepsilon, \varepsilon) &= \{(p, \varepsilon)\} \\
\delta(p, a, A) &= \{(p, \varepsilon)\} \\
\delta(p, b, B) &= \{(p, \varepsilon)\} \\
\delta(p, c, AA) &= \{(p, \varepsilon)\} \\
\delta(p, c, BBB) &= \{(p, \varepsilon)\} \\
\delta(p, d, Z) &= \{(f, \varepsilon)\}
\end{aligned}
$$

Ekvivalence rozšířených PDA a PDA

Lemma 3.45. Ke každému rozšířenému PDA existuje ekvivalentní (*obyčejný*) PDA.

Intuice:

Důkaz. Nechť $\mathcal{R} = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$ je rozšířený PDA a $m = \max\{|\alpha| \mid \delta(q, a, \alpha)$ je definováno pro nějaké $q \in Q, a \in \Sigma \cup \{\varepsilon\}\}$.

Definujeme $\mathcal{P} = (Q_1, \Sigma, \Gamma_1, \delta_1, q_1, Z_1, F_1)$, kde

- $Q_1 = \{[q, \alpha] \mid q \in Q, \alpha \in \Gamma_1^*,\ 0 \leq |\alpha| \leq m\}$,
- $\Gamma_1 = \Gamma \cup \{Z_1\}$, kde Z_1 je nový symbol,
- $q_1 = [q_0, Z_0 Z_1^{m-1}]$,
- $F_1 = \{[q, \alpha] \mid q \in F, \alpha \in \Gamma_1^*,\ 0 \leq |\alpha| \leq m\}$.

- δ_1 je definována takto:
  - jestliže $\delta(q, a, X_1 \dots X_k)$ obsahuje $(r, Y_1 \dots Y_l)$, pak

    $l \geq k$: $\delta_1([q, X_1 \dots X_k\alpha], a, Z)$ obsahuje $([r, \beta], \gamma Z)$,
    kde $\beta\gamma = Y_1 \dots Y_l\alpha$ a $|\beta| = m$,
    pro všechny $Z \in \Gamma_1$ a $\alpha \in \Gamma_1^*$ takové, že $|\alpha| = m - k$

    $l < k$: $\delta_1([q, X_1 \dots X_k\alpha], a, Z)$ obsahuje $([r, Y_1 \dots Y_l\alpha Z], \varepsilon)$
    pro všechny $Z \in \Gamma_1$ a $\alpha \in \Gamma_1^*$ takové, že $|\alpha| = m - k$
  - $\delta_1([q, \alpha], \varepsilon, Z) = \{([q, \alpha Z], \varepsilon)\}$
    pro všechny $q \in Q$, $Z \in \Gamma_1$ a $\alpha \in \Gamma_1^*$ takové, že $|\alpha| < m$

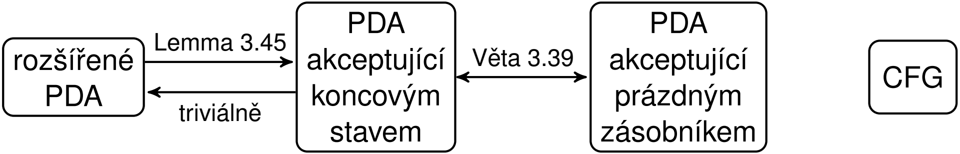
Korektnost: Ověříme, že platí

$$(q, aw, X_1 \dots X_k X_{k+1} \dots X_n) \vdash_{\mathcal{R}} (r, w, Y_1 \dots Y_l X_{k+1} \dots X_n)$$
$$\iff \quad ([q, \alpha], aw, \beta) \vdash^{+}_{\mathcal{P}} ([r, \alpha'], w, \beta'),$$

kde

1. $\alpha\beta = X_1 \dots X_n Z_1^m$,
2. $\alpha'\beta' = Y_1 \dots Y_l X_{k+1} \dots X_n Z_1^m$,
3. $|\alpha| = |\alpha'| = m$ a
4. mezi dvěma výše uvedenými konfiguracemi PDA $\mathcal{P}$ neexistuje taková konfigurace, kde druhá komponenta stavu (tj. buffer) by měla délku m. □

Zásobníkové automaty a bezkontextové jazyky



Zásobníkové automaty a bezkontextové jazyky

Motivace 1: Jaká je třída jazyků rozpoznávaných zásobníkovými automaty?

Motivace 2: Je dána bezkontextová gramatika $\mathcal{G}$ a slovo w.
Jak zjistit, zda se slovo w dá vygenerovat v gramatice $\mathcal{G}$?

Problém syntaktické analýzy pro bezkontextové gramatiky:
pro danou bezkontextovou gramatiku $\mathcal{G}$ a slovo w rozhodnout, zda $w \in L(\mathcal{G})$.

Ekvivalence bezkontextových gramatik a zásobníkových automatů

Věta 3.51. Ke každému PDA $\mathcal{M}$ lze sestrojit CFG $\mathcal{G}$ takovou, že $L_e(\mathcal{M}) = L(\mathcal{G})$.

Důkaz. Vynechán. $\square$

Věta 3.47. Ke každé CFG $\mathcal{G}$ lze sestrojit PDA $\mathcal{M}$ takový, že $L(\mathcal{G}) = L_e(\mathcal{M})$.

Důkaz. Uvedeme za chvíli. $\square$

Důsledek 3.52. Třída jazyků rozpoznávaných zásobníkovými automaty je právě třída bezkontextových jazyků.

Intuice převodu PDA na CFG

1. $|Q| = 1$

Intuice převodu PDA na CFG

2. $|Q| \geq 2$

Intuice převodu CFG na PDA

Konstrukce PDA řeší problém syntaktické analýzy.
(Platí pro dané $\mathcal{G}$ a w: $w \in L(\mathcal{G})$?)

$$w \in L(\mathcal{G}) \iff \text{v } \mathcal{G} \text{ existuje derivační strom s výsledkem } w$$

Intuice převodu CFG na PDA aneb O nedeterministické syntaktické analýze

PDA se bude snažit budovat derivační strom pro w.

shora dolů	zdola nahoru

Intuice pro analýzu shora dolů

Budování derivačního stromu simuluje levé derivace, tj. vždy rozvíjíme nejlevější neterminál.

Nedeterministická syntaktická analýza shora dolů

Věta 3.47. Ke každé CFG $\mathcal{G}$ lze sestrojit PDA $\mathcal{M}$ takový, že $L(\mathcal{G}) = L_e(\mathcal{M})$.

Důkaz. K dané gramatice $\mathcal{G}$ konstruujeme PDA $\mathcal{M}$, který simuluje levé derivace v $\mathcal{G}$.

- V levé derivaci je v jednom kroku odvození nahrazen (nejlevější) neterminál A pravou stranou $X_1 \dots X_n$ nějakého A-pravidla.
- V $\mathcal{M}$ této situaci odpovídá náhrada A na vrcholu zásobníku řetězem $X_1 \dots X_n$.

$\mathcal{M} = (\{q\}, \Sigma, N \cup \Sigma, \delta, q, S, \emptyset)$, kde δ je definována:

- $\delta(q, \varepsilon, A)$ obsahuje (q, α) právě když $A \to \alpha \in P$
- $\delta(q, a, a) = \{(q, \varepsilon)\}$ pro všechna $a \in \Sigma$

$$
\begin{aligned}
S &\rightarrow aAB \\
A &\rightarrow Aa \mid \varepsilon \\
B &\rightarrow SaA \mid b
\end{aligned}
\qquad
\begin{aligned}
\delta(q, \varepsilon, S) &= \{(q, aAB)\} \\
\delta(q, \varepsilon, A) &= \{(q, Aa), (q, \varepsilon)\} \\
\delta(q, \varepsilon, B) &= \{(q, SaA), (q, b)\} \\
\delta(q, a, a) &= \delta(q, b, b) = \{(q, \varepsilon)\}
\end{aligned}
$$

$$
\begin{aligned}
S &\Rightarrow aAB \\
&\Rightarrow aB \\
&\Rightarrow aSaA \\
&\Rightarrow aaABaA \\
&\Rightarrow aaBaA \\
&\Rightarrow aabaA \\
&\Rightarrow aaba
\end{aligned}
$$

Korektnost

$$A \Rightarrow^* w \iff (q, w, A) \vdash^* (q, \varepsilon, \varepsilon)$$

$(\Longrightarrow)$ Indukcí vzhledem k délce odvození m.

- $m = 1$: zřejmé.
- $m > 1$: nechť tvrzení platí pro všechna $m' < m$.

  $A \Rightarrow X_1 X_2 \dots X_k \Rightarrow^* x_1 x_2 \dots x_k = w$, kde $X_i \overset{m_i}{\Rightarrow} x_i$, $0 \le m_i < m$
  z definice δ plyne $(q, w, A) \vdash (q, w, X_1 X_2 \dots X_k)$.

  Je-li $X_i \in N$, pak dle indukčního předpokladu máme
  $(q, x_i, X_i) \vdash^* (q, \varepsilon, \varepsilon)$.

  Je-li $X_i \in \Sigma$, pak $X_i = x_i$ a z definice δ plyne $(q, x_i, x_i) \vdash (q, \varepsilon, \varepsilon)$.

  Kompozicí dostáváme $(q, w, A) \vdash^+ (q, \varepsilon, \varepsilon)$.

$(\Longleftarrow)$ Předpokládejme $(q, w, A) \vdash^{n} (q, \varepsilon, \varepsilon)$ a ukažme $A \Rightarrow^{+} w$.
Indukcí vzhledem k délce výpočtu n.

- $n = 1$: zřejmé.
- $n > 1$: nechť tvrzení platí pro všechna $n' < n$.

  $(q, w, A) \vdash (q, w, X_1 X_2 \dots X_k)$, tj. $A \to X_1 X_2 \dots X_k \in P$

  w můžeme napsat jako $w = x_1 x_2 \dots x_k$ takové, že
  - je-li $X_i \in N$, pak $(q, x_i, X_i) \vdash^{n_i} (q, \varepsilon, \varepsilon)$, kde $n_i < n$.
    Dle IP $X_i \Rightarrow^{+} x_i$.
  - je-li $X_i \in \Sigma$, pak $X_i \overset{0}{\Rightarrow} x_i$.

Vhodnou kompozicí obdržíme
$A \Rightarrow X_1 X_2 \dots X_k \Rightarrow^{*} x_1 X_2 \dots X_k \Rightarrow^{*} x_1 \dots x_k = w$
což je levá derivace slova w v gramatice $\mathcal{G}$. $\square$