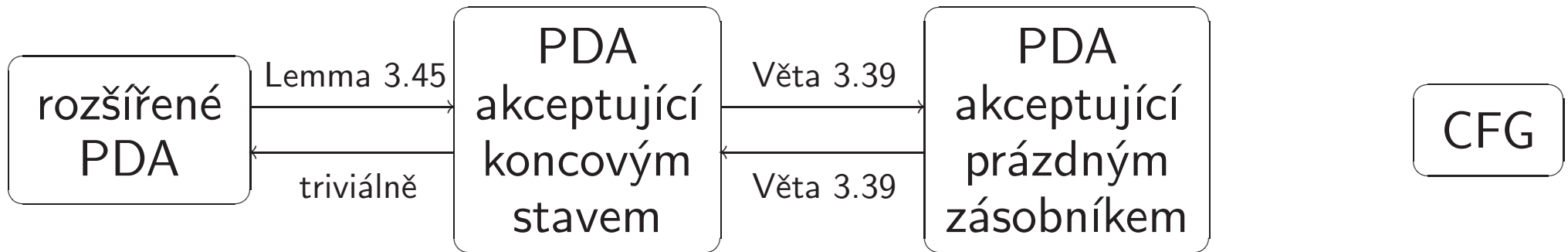


# Zásobníkové automaty a bezkontextové jazyky



# 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 slovo  $w$  sa 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 sestavit CFG  $\mathcal{G}$  takovou, že  $L_e(\mathcal{M}) = L(\mathcal{G})$ .

**Důkaz.** Vynechán.  $\square$

**Věta 3.47.** Ke každé CFG  $\mathcal{G}$  lze sestavit 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  $w$ :  $w \in L(\mathcal{G})$ ?)

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

# Intuice převodu CFG na PDA aneb 0 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 sestavit 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  $\delta$  je definována:

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

$S \rightarrow aAB$   
 $A \rightarrow Aa \mid \varepsilon$   
 $B \rightarrow SaA \mid b$

$\delta(q, \varepsilon, S) = \{(q, aAB)\}$   
 $\delta(q, \varepsilon, A) = \{(q, Aa), (q, \varepsilon)\}$   
 $\delta(q, \varepsilon, B) = \{(q, SaA), (q, b)\}$   
 $\delta(q, a, a) = \delta(q, b, b) = \{(q, \varepsilon)\}$

---

$S \Rightarrow aAB$   
 $\Rightarrow aB$   
 $\Rightarrow aSaA$   
 $\Rightarrow aaABaA$   
 $\Rightarrow aaBaA$   
 $\Rightarrow aabaA$   
 $\Rightarrow aaba$

# Korektnost

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

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

1.  $m = 1$ : zřejmé.
2.  $m > 1$ : 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  $X_i \xRightarrow{m_i} x_i$ ,  $0 \leq m_i < m$   
z definice  $\delta$  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  $\delta$  plyne  $(q, x_i, x_i) \vdash (q, \varepsilon, \varepsilon)$ .

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

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

1.  $n = 1$ : zřejmé.

2.  $n > 1$ : nechť tvrzení platí pro všechna  $n' < n$ .

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

$w$  můžeme napsat jako  $w = x_1 x_2 \dots x_k$  takové, že

• je-li  $X_i \in N$ , pak  $(q, x_i, X_i) \stackrel{n_i}{\vdash} (q, \varepsilon, \varepsilon)$ , kde  $n_i < n$ .

Dle IP  $X_i \Rightarrow^+ x_i$ .

• je-li  $X_i \in \Sigma$ , pak  $X_i \stackrel{0}{\Rightarrow} x_i$ .

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  $\mathcal{G}$ . □

# Intuice pro analýzu zdola nahoru

# Intuice pro analýzu zdola nahoru

# Nedeterministická syntaktická analýza zdola nahoru

**Věta 3.55.** Nechť  $\mathcal{G}$  je libovolná CFG, pak lze zkonstruovat rozšířený PDA  $\mathcal{R}$  takový, že  $L(\mathcal{G}) = L(\mathcal{R})$ .

**Důkaz. Vrchol zásobníku píšeme vpravo.**

Konstruujeme rozšířený PDA  $\mathcal{R}$ , který simuluje pravou derivaci v  $\mathcal{G}$  v obráceném pořadí.

PDA  $\mathcal{R}$  má kroky dvojího typu:

1. může kdykoli číst do zásobníku vstupní symbol,
2. (**redukce**) je-li na vrcholu zásobníku řetěz tvořící pravou stranu nějakého pravidla v  $\mathcal{G}$ , může ho nahradit odpovídajícím levostranným neterminálem (a ze vstupu nic nečte).

Nechť  $\mathcal{G} = (N, \Sigma, P, S)$ .

Položme  $\mathcal{R} = (\{q, r\}, \Sigma, N \cup \Sigma \cup \{\perp\}, \delta, q, \perp, \{r\})$ , kde  $\perp$  je nově přidaný symbol a kde  $\delta$  je definována takto:

1.  $\delta(q, a, \varepsilon) = \{(q, a)\}$  pro všechna  $a \in \Sigma$ ,
2. je-li  $A \rightarrow \alpha$ , pak  $\delta(q, \varepsilon, \alpha)$  obsahuje  $(q, A)$ ,
3.  $\delta(q, \varepsilon, \perp S) = \{(r, \varepsilon)\}$ .



| krok výpočtu  | odpovídající pravidlo z $\mathcal{G}$ |
|---|---------------------------------------|
| $(q, i + i * i, \perp) \xrightarrow{i} (q, +i * i, \perp i)$  | $F \rightarrow i$                     |
| $\xrightarrow{\varepsilon} (q, +i * i, \perp F)$              | $T \rightarrow F$                     |
| $\xrightarrow{\varepsilon} (q, +i * i, \perp T)$              | $E \rightarrow T$                     |
| $\xrightarrow{\varepsilon} (q, +i * i, \perp E)$              |                                       |
| $\xrightarrow{+} (q, i * i, \perp E +)$                       |                                       |
| $\xrightarrow{i} (q, *i, \perp E + i)$                        | $F \rightarrow i$                     |
| $\xrightarrow{\varepsilon} (q, *i, \perp E + F)$              | $T \rightarrow F$                     |
| $\xrightarrow{\varepsilon} (q, *i, \perp E + T)$              |                                       |
| $\xrightarrow{*} (q, i, \perp E + T *)$                       |                                       |
| $\xrightarrow{i} (q, \varepsilon, \perp E + T * i)$           | $F \rightarrow i$                     |
| $\xrightarrow{\varepsilon} (q, \varepsilon, \perp E + T * F)$ | $T \rightarrow T * F$                 |
| $\xrightarrow{\varepsilon} (q, \varepsilon, \perp E + T)$     | $E \rightarrow E + T$                 |
| $\xrightarrow{\varepsilon} (q, \varepsilon, \perp E)$         |                                       |
| $\xrightarrow{\varepsilon} (r, \varepsilon, \varepsilon)$     |                                       |

# Korektnost

$$S \Rightarrow^* \alpha A y \xrightarrow{n} xy \iff (q, xy, \perp) \vdash^* (q, y, \perp \alpha A),$$

kde  $S \Rightarrow^* \alpha A y \xrightarrow{n} xy$  je pravá derivace a  $A$  je nejpravější neterminál.

( $\implies$ ) indukcí k délce odvození

( $\impliedby$ ) indukcí k délce výpočtu

Pro  $A = S$  a  $\alpha, y = \varepsilon$  dostáváme:

$$S \Rightarrow^* x \iff (q, x, \perp) \vdash^* (q, \varepsilon, \perp S) \quad \left[ \vdash (r, \varepsilon) \right]$$

"Výstupem" je pravá derivace v obráceném pořadí. □

# Efektivnost syntaktické analýzy

Nedeterministický PDA  $\implies$  nedeterministický algoritmus  
 $\implies$  exponenciální deterministický algoritmus

## Řešení:

- deterministický algoritmus složitosti  $\mathcal{O}(n^3)$ , kde  $n = |w|$   
(algoritmus Cocke - Younger - Kasami)
- deterministické zásobníkové automaty  
a deterministické bezkontextové jazyky
- lineární algoritmy pro speciální třídy deterministických  
bezkontextových jazyků