

# Formální jazyky a automaty

Formální jazyky, gramatiky, Chomského hierarchie

Jan Křetínský

Fakulta informatiky, MU Brno

Jaro 2024

# Formální jazyk

**Abeceda** je libovolná konečná množina.

**Slovo** nad abecedou  $\Sigma$  je konečná posloupnost znaků této abecedy.  
(Množinu všech slov nad abecedou  $\Sigma$  značíme  $\Sigma^*$ .)

**Jazyk** nad abecedou  $\Sigma$  je libovolná množina slov nad  $\Sigma$   
(tedy podmnožina  $\Sigma^*$ ).

# Operace nad slovy

Binární operace **zřetězení**, označována  $\cdot$ , která je definována předpisem:  $u \cdot v = uv$

Unární operace  **$i$ -té mocniny** slova, která je definovaná induktivně pro každé  $i \in \mathbb{N}_0$  takto: necht'  $\Sigma$  je libovolná abeceda,  $u$  libovolné slovo nad abecedou  $\Sigma$ . Pak

- ▶  $u^0 = \varepsilon$
- ▶  $u^{i+1} = u \cdot u^i$

Unární operace **zrcadlového obrazu** slova  $w = a_1 \dots a_n$  je definovaná předpisem:  $w^R = a_n \dots a_1$  ( $\varepsilon^R = \varepsilon$ ).

# Operace nad jazyky

Nechť  $L$  je jazyk nad abecedou  $\Sigma$ ,  $K$  je jazyk nad abecedou  $\Delta$ .  
Výsledkem je vždy jazyk nad abecedou  $\Sigma \cup \Delta$ .

- ▶ Standardní množinové operace **sjednocení** ( $\cup$ ), **průnik** ( $\cap$ ) a **rozdíl** ( $\setminus$ ).
- ▶ **Zřetězením** jazyků  $L$  a  $K$  je jazyk  $L.K = \{u.v \mid u \in L, v \in K\}$ .

- ▶  $i$ -tá mocnina jazyka  $L$  definována induktivně pro  $i \in \mathbb{N}_0$ :

$$L^0 =$$

$$L^{i+1} = L.L^i$$

# Operace nad jazyky

Nechť  $L$  je jazyk nad abecedou  $\Sigma$ ,  $K$  je jazyk nad abecedou  $\Delta$ .  
Výsledkem je vždy jazyk nad abecedou  $\Sigma \cup \Delta$ .

- ▶ Standardní množinové operace **sjednocení** ( $\cup$ ), **průnik** ( $\cap$ ) a **rozdíl** ( $\setminus$ ).
- ▶ **Zřetězením** jazyků  $L$  a  $K$  je jazyk  $L.K = \{u.v \mid u \in L, v \in K\}$ .

$$\emptyset.L = L.\emptyset = \emptyset \quad \text{a} \quad \{\varepsilon\}.L = L.\{\varepsilon\} = L.$$

- ▶  $i$ -tá mocnina jazyka  $L$  definována induktivně pro  $i \in \mathbb{N}_0$ :

$$L^0 =$$

$$L^{i+1} = L.L^i$$

# Operace nad jazyky

Nechť  $L$  je jazyk nad abecedou  $\Sigma$ ,  $K$  je jazyk nad abecedou  $\Delta$ .  
Výsledkem je vždy jazyk nad abecedou  $\Sigma \cup \Delta$ .

- ▶ Standardní množinové operace **sjednocení** ( $\cup$ ), **průnik** ( $\cap$ ) a **rozdíl** ( $\setminus$ ).
- ▶ **Zřetěžením** jazyků  $L$  a  $K$  je jazyk  $L.K = \{u.v \mid u \in L, v \in K\}$ .

$$\emptyset.L = L.\emptyset = \emptyset \quad \text{a} \quad \{\varepsilon\}.L = L.\{\varepsilon\} = L.$$

- ▶  $i$ -tá mocnina jazyka  $L$  definována induktivně pro  $i \in \mathbb{N}_0$ :

$$L^0 = \{\varepsilon\}$$

$$L^{i+1} = L.L^i$$

$$\emptyset^0 = \{\varepsilon\}$$

$$\emptyset^i = \emptyset \quad \text{pro libovolné } i \in \mathbb{N}$$

$$\{\varepsilon\}^j = \{\varepsilon\} \quad \text{pro libovolné } j \in \mathbb{N}_0$$

## Operace nad jazyky 2/3

- ▶ Iterace jazyka  $L$  je jazyk  $L^* = \bigcup_{i=0}^{\infty} L^i$ .
  
  
  
  
  
  
  
  
  
  
- ▶ Pozitivní iterace jazyka  $L$  je jazyk  $L^+ = \bigcup_{i=1}^{\infty} L^i$ .

# Operace nad jazyky 2/3

- ▶ Iterace jazyka  $L$  je jazyk  $L^* = \bigcup_{i=0}^{\infty} L^i$ .

$$\emptyset^* = \{\varepsilon\}$$

- ▶ Pozitivní iterace jazyka  $L$  je jazyk  $L^+ = \bigcup_{i=1}^{\infty} L^i$ .



# Operace nad jazyky 2/3

- ▶ **Iterace** jazyka  $L$  je jazyk  $L^* = \bigcup_{i=0}^{\infty} L^i$ .

$$\emptyset^* = \{\varepsilon\}$$

- ▶ **Pozitivní iterace** jazyka  $L$  je jazyk  $L^+ = \bigcup_{i=1}^{\infty} L^i$ .

$$\emptyset^+ = \emptyset.$$

## Operace nad jazyky 3/3

- ▶ Doplněk jazyka  $L$  je jazyk  $co-L = \Sigma^* \setminus L$ .
- ▶ Zrcadlový obraz jazyka  $L$  definujeme  $L^R = \{w^R \mid w \in L\}$ .

Nechť  $\mathcal{L}$  je třída jazyků a  $o$  je  $n$ -ární operace na jazycích. Řekneme, že  $\mathcal{L}$  je **uzavřená na  $o$** , pokud pro libovolné jazyky  $L_1, \dots, L_n$  patřící do  $\mathcal{L}$  platí, že také jazyk  $o(L_1, \dots, L_n)$  patří do  $\mathcal{L}$ .

# Konečná reprezentace jazyka

- ▶ automaty, gramatiky, logiky,...
- ▶ Existuje konečná reprezentace pro každý jazyk?
- ▶ Jaké vlastnosti mají jazyky, které jsou konečně reprezentovatelné?

**Gramatika** je popis jazyka pomocí pravidel, podle kterých se vytvářejí všechna slova daného jazyka.

Zadání syntaxe vyšších programovacích jazyků — Backus-Naurova normální forma (BNF)

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$\langle \text{Expr} \rangle \rightarrow$

$\rightarrow a * (b + c)$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$\langle \text{Expr} \rangle \rightarrow$

$\rightarrow a * (b + c)$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$\langle \text{Expr} \rangle \rightarrow$

$\rightarrow a * (b + c)$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\langle \text{Expr} \rangle \rightarrow \langle \text{Term} \rangle$$

$$\rightarrow a * (b + c)$$



# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\langle \text{Expr} \rangle \rightarrow \langle \text{Term} \rangle$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\langle \text{Expr} \rangle \rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\langle \text{Expr} \rangle \rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle\end{aligned}$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle\end{aligned}$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle\end{aligned}$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle\end{aligned}$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$



# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle)\end{aligned}$$

$$\rightarrow a * (b + c)$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \rightarrow a * (\langle \text{Factor} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \rightarrow a * (\langle \text{Factor} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \rightarrow a * (\langle \text{Factor} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$



# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \rightarrow a * (\langle \text{Factor} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \rightarrow a * (\langle \text{Factor} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + \langle \text{Term} \rangle) \rightarrow a * (b + \langle \text{Factor} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad

Abeceda  $\{a, b, c, +, *, (, )\}$  a pravidla

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \\ \langle \text{Expr} \rangle &\rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Factor} \rangle \\ \langle \text{Term} \rangle &\rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ \langle \text{Factor} \rangle &\rightarrow a \mid b \mid c \\ \langle \text{Factor} \rangle &\rightarrow (\langle \text{Expr} \rangle)\end{aligned}$$

Z  $\langle \text{Expr} \rangle$  odvodíme  $a * (b + c)$  :

$$\begin{aligned}\langle \text{Expr} \rangle &\rightarrow \langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle \\ &\rightarrow \langle \text{Factor} \rangle * \langle \text{Factor} \rangle \rightarrow a * \langle \text{Factor} \rangle \\ &\rightarrow a * (\langle \text{Expr} \rangle) \rightarrow a * (\langle \text{Expr} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (\langle \text{Term} \rangle + \langle \text{Term} \rangle) \rightarrow a * (\langle \text{Factor} \rangle + \langle \text{Term} \rangle) \\ &\rightarrow a * (b + \langle \text{Term} \rangle) \rightarrow a * (b + \langle \text{Factor} \rangle) \\ &\rightarrow a * (b + c)\end{aligned}$$

# Gramatika: Příklad formálně

▶  $N = \{\langle \text{Expr} \rangle, \langle \text{Term} \rangle, \langle \text{Factor} \rangle\}$

▶  $\Sigma = \{a, b, c, +, *, (, )\}$

▶  $S = \langle \text{Expr} \rangle$

▶ Množina pravidel  $P$  obsahuje

$\langle \text{Expr} \rangle \rightarrow \langle \text{Term} \rangle$

$\langle \text{Expr} \rangle \rightarrow \langle \text{Expr} \rangle + \langle \text{Term} \rangle$

$\langle \text{Term} \rangle \rightarrow \langle \text{Factor} \rangle$

$\langle \text{Term} \rangle \rightarrow \langle \text{Term} \rangle * \langle \text{Factor} \rangle$

$\langle \text{Factor} \rangle \rightarrow a \mid b \mid c$

$\langle \text{Factor} \rangle \rightarrow (\langle \text{Expr} \rangle)$

# Gramatika: Obecně formálně

**Definice.** Gramatika  $\mathcal{G}$  je čtveřice  $(N, \Sigma, P, S)$ , kde

- ▶  $N$  je neprázdná konečná množina **neterminálních symbolů** (stručněji **neterminálů**),
- ▶  $\Sigma$  je konečná množina **terminálních symbolů (terminálů)** taková, že  $N \cap \Sigma = \emptyset$ . Množinu **všech symbolů** gramatiky definujeme jako  $V = N \cup \Sigma$ ,
- ▶  $P \subseteq V^*NV^* \times V^*$  je konečná množina **pravidel**. Pravidlo  $(\alpha, \beta)$  obvykle zapisujeme ve tvaru  $\alpha \rightarrow \beta$  (a čteme jako „ $\alpha$  přepiš na  $\beta$ “),
- ▶  $S \in N$  je speciální **počáteční neterminál** (nazývaný také **kořen gramatiky**).

# Odvození

Gramatika  $\mathcal{G} = (N, \Sigma, P, S)$  určuje

- ▶ relaci  $\Rightarrow_{\mathcal{G}}$  **přímého odvození** na množině  $V^*$

$\gamma \Rightarrow_{\mathcal{G}} \delta$  právě když existuje pravidlo  $\alpha \rightarrow \beta \in P$  a slova  $\eta, \rho \in V^*$  taková, že  $\gamma = \eta\alpha\rho$  a  $\delta = \eta\beta\rho$ .

Používá se i označení **krok odvození**.

► relaci  $\overset{k}{\Rightarrow}_G$  odvození v  $k$  krocích pro  $k \in \mathbb{N}_0$

$\overset{0}{\Rightarrow}_G$  je identická relace

$$\overset{k+1}{\Rightarrow}_G = \overset{k}{\Rightarrow}_G \circ \Rightarrow_G$$

- relaci  $\Rightarrow_G^k$  **odvození v  $k$  krocích** pro  $k \in \mathbb{N}_0$

$\Rightarrow_G^0$  je identická relace

$$\Rightarrow_G^{k+1} = \Rightarrow_G^k \circ \Rightarrow_G$$

- relaci  $\Rightarrow_G^*$  **odvození** a relaci  $\Rightarrow_G^+$  **netriviálního odvození**

$$\Rightarrow_G^* = \bigcup_{i=0}^{\infty} \Rightarrow_G^i \qquad \Rightarrow_G^+ = \bigcup_{i=1}^{\infty} \Rightarrow_G^i$$

Relace  $\Rightarrow_G^*$  je reflexivní a tranzitivní uzávěr relace  $\Rightarrow_G$ .

Relace  $\Rightarrow_G^+$  je tranzitivní uzávěr relace  $\Rightarrow_G$ .



**Větná forma** gramatiky  $\mathcal{G}$  je každý řetěz z množiny  $V^*$ , který lze odvodit z počátečního neterminálu gramatiky.

**Věta** gramatiky  $\mathcal{G}$  je každá větná forma, která obsahuje pouze terminály.

**Jazyk generovaný gramatikou**  $\mathcal{G}$  je množina  $L(\mathcal{G})$  všech vět gramatiky

$$L(\mathcal{G}) = \{w \in \Sigma^* \mid S \Rightarrow_{\mathcal{G}}^* w\}.$$

Gramatiky  $\mathcal{G}_1$  a  $\mathcal{G}_2$  nazveme **jazykově ekvivalentní**, právě když generují tentýž jazyk, tj.  $L(\mathcal{G}_1) = L(\mathcal{G}_2)$ .

# Příklad

# Příklad

# Chomského hierarchie gramatik

Klasifikace gramatik podle tvaru prepisovacích pravidel:

typ 0 (frázové gramatiky)

pravidla v obecném tvaru

typ 1 (kontextové gramatiky)

pro každé pravidlo  $\alpha \rightarrow \beta$  platí  $|\alpha| \leq |\beta|$

typ 2 (bezkontextové gramatiky (bez  $\varepsilon$ -pravidel))

každé pravidlo je tvaru  $A \rightarrow \alpha$ , kde  $|\alpha| \geq 1$

typ 3 (regulární gramatiky)

každé pravidlo je tvaru  $A \rightarrow aB$  nebo  $A \rightarrow a$

# Chomského hierarchie gramatik

Klasifikace gramatik podle tvaru přepisovacích pravidel:

typ 0 (frázové gramatiky)

pravidla v obecném tvaru

typ 1 (kontextové gramatiky)

pro každé pravidlo  $\alpha \rightarrow \beta$  platí  $|\alpha| \leq |\beta|$

s eventuální výjimkou pravidla  $S \rightarrow \varepsilon$ , pokud se  $S$  nevyskytuje na pravé straně žádného pravidla

typ 2 (bezkontextové gramatiky (bez  $\varepsilon$ -pravidel))

každé pravidlo je tvaru  $A \rightarrow \alpha$ , kde  $|\alpha| \geq 1$

s eventuální výjimkou pravidla  $S \rightarrow \varepsilon$ , pokud se  $S$  nevyskytuje na pravé straně žádného pravidla

[zde skripta povolují  $|\alpha| \geq 0$ , tj. obecně všechna pravidla tvaru  $A \rightarrow \varepsilon$ , kterých se však lze zbavit, jak uvidíme později]

typ 3 (regulární gramatiky)

každé pravidlo je tvaru  $A \rightarrow aB$  nebo  $A \rightarrow a$

s eventuální výjimkou pravidla  $S \rightarrow \varepsilon$ , pokud se  $S$  nevyskytuje na pravé straně žádného pravidla

# Chomského hierarchie jazyků

Hierarchie gramatik určuje hierarchii jazyků.

**Jazyk  $L$  je typu 0 (rekursivně spočetný)**, pokud existuje gramatika  $\mathcal{G}$  typu 0 taková, že  $L(\mathcal{G}) = L$ .

Analogicky: **kontextový, bezkontextový, regulární**

$\mathcal{L}_0$  třída všech rekursivně spočetných jazyků

$\mathcal{L}_1$  třída všech kontextových jazyků

$\mathcal{L}_2$  třída všech bezkontextových jazyků

$\mathcal{L}_3$  třída všech regulárních jazyků

$$\mathcal{L}_0 \supsetneq \mathcal{L}_1 \supsetneq \mathcal{L}_2 \supsetneq \mathcal{L}_3$$

*(Dokážeme později.)*

# (Ne)reprezentovatelnost

Věta.

Nad abecedou  $\{a\}$  existuje jazyk, který není typu 0.

# (Ne)reprezentovatelnost

## Věta.

Nad abecedou  $\{a\}$  existuje jazyk, který není typu 0.

## Důkaz.

Množina všech slov nad abecedou  $\{a\}$  je spočetně nekonečná. Množina všech jazyků nad touto abecedou má proto mohutnost  $2^{\aleph_0}$  (je tedy **nespočetná**).

Gramatik typu 0 nad abecedou  $\{a\}$  je pouze **spočetně mnoho**:

- ▶ buď  $M$  libovolná, ale pevně zvolená spočetná množina
- ▶ b.ú.n.o. každá gramatika má neterminály z  $M$
- ▶ každá gramatika je slovo nad abecedou

$$M \cup \{a, \rightarrow, \varepsilon, \underline{\quad}, \underline{\quad}, \underline{\quad}, \underline{\quad}, \underline{\quad}\}$$

- ▶ všech slov délky  $i$  nad touto abecedou je  $\aleph_0^i = \aleph_0$  pro lib.  $i \in \mathbb{N}$
- ▶ **všech** slov nad touto abecedou je tedy spočetně mnoho (*sjednocení spočetně mnoha spočetných množin je spočetné*)

