

Süntaksanalüüs

Kontekstivabad grammatikad
Grammatikate mitmesus

Kontekstivabad grammatikad

- Kontekstivabade grammatikate produktsioonireegleid esitatakse tavaliselt **Backus-Naur'i kujul** (BNF).
- Näide: olgu $N = \{\text{Exp}\}$ ja $T = \{+, *, (,), \text{num}\}$, siis

$$\begin{array}{l} \text{Exp} \rightarrow \text{Exp} + \text{Exp} \\ \quad | \text{Exp} * \text{Exp} \\ \quad | (\text{Exp}) \\ \quad | \text{num} \end{array}$$

esitab produktsioonireeglite hulka

$$P = \{ \text{Exp} \rightarrow \text{Exp} + \text{Exp}, \text{Exp} \rightarrow (\text{Exp}), \\ \text{Exp} \rightarrow \text{Exp} * \text{Exp}, \text{Exp} \rightarrow \text{num} \}.$$

Kontekstivabad grammatikad

- Reeglina saab ühte ja sama lauset tuletada paljudel eri viisidel.
- Kanoonilised derivatsioonid:
 - **vasakderivatsioon** – derivatsiooni igal sammul asendatakse vasakpoolseim mitteterminaal;
 - **paremderivatsioon** – derivatsiooni igal sammul asendatakse parempoolseim mitteterminaal.
- Näide:

$$\begin{aligned} \text{Exp} &\Longrightarrow_{lm} \text{Exp} + \text{Exp} \\ &\Longrightarrow_{lm} \text{num} + \text{Exp} \\ &\Longrightarrow_{lm} \text{num} + \text{Exp} * \text{Exp} \\ &\Longrightarrow_{lm} \text{num} + \text{num} * \text{Exp} \\ &\Longrightarrow_{lm} \text{num} + \text{num} * \text{num} \end{aligned}$$
$$\begin{aligned} \text{Exp} &\Longrightarrow_{rm} \text{Exp} + \text{Exp} \\ &\Longrightarrow_{rm} \text{Exp} + \text{Exp} * \text{Exp} \\ &\Longrightarrow_{rm} \text{Exp} + \text{Exp} * \text{num} \\ &\Longrightarrow_{rm} \text{Exp} + \text{num} * \text{num} \\ &\Longrightarrow_{rm} \text{num} + \text{num} * \text{num} \end{aligned}$$

Kontekstivabad grammatikad

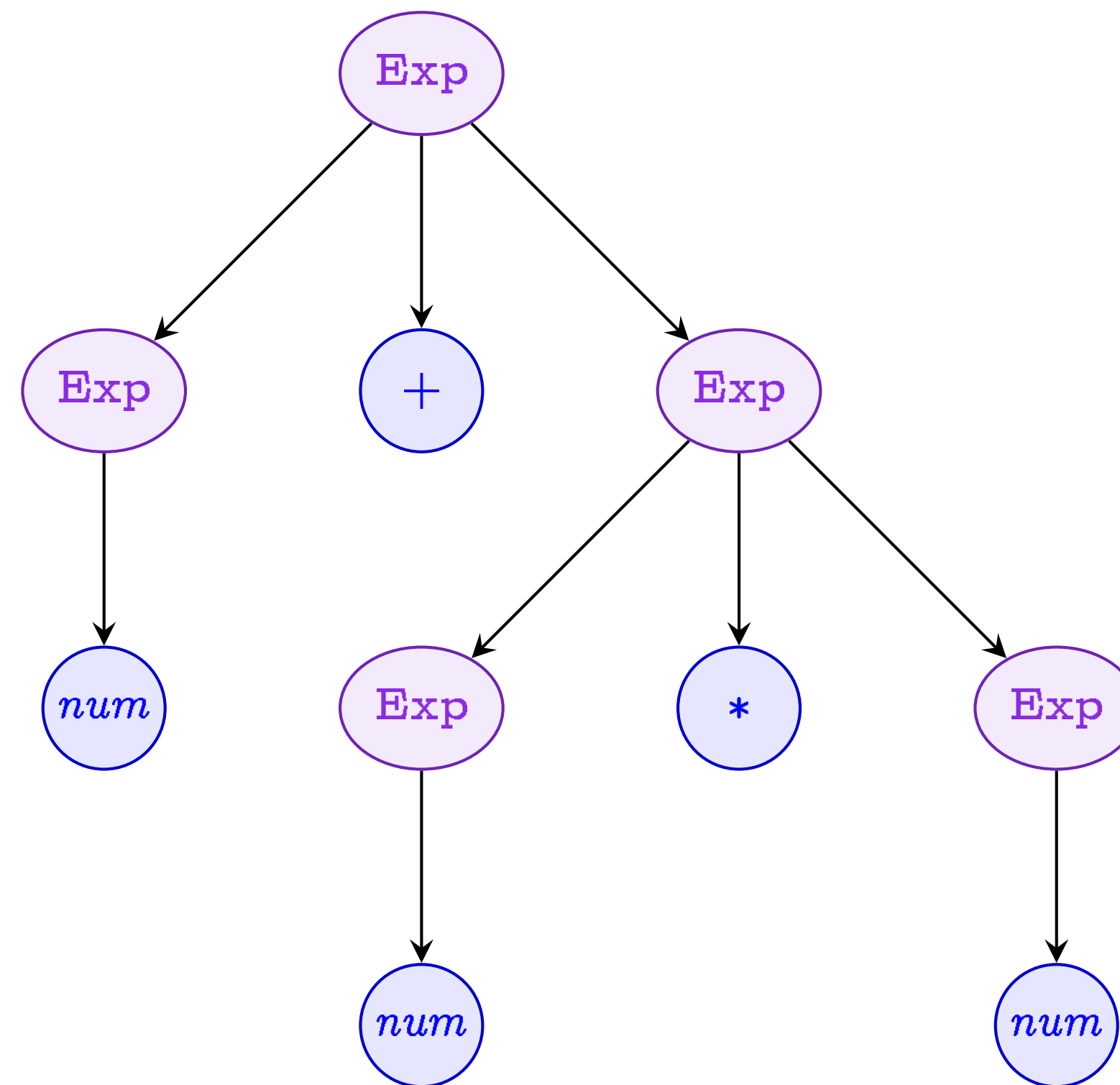
- Iga derivatsioon määrab üheselt **süntaksipuu** (**syntax-tree**, **parse-tree**), mis on järjestatud tippudega puu, kus:
 - puu juur on märgendatud algsümboliga **S**;
 - vahetipud on märgendatud mitteterminaalidega;
 - lehed on märgendatud terminaalsümboliga või tühisümboliga ϵ ;
 - kui vahetipp on märgendatud mitteterminaaliga A ja tema alampuude (vasakult paremale) t_1, \dots, t_n juured on märgendatud vastavalt A_1, \dots, A_n , siis $A \rightarrow A_1 \dots A_n \in P$.
- Tuletatud lause saadakse lugedes puu lehtede märgendid vasakult paremale.
- Süntaksipuu määrab üheselt kasutatud produktsiooni-reeglid, kuid mitte nende rakendamise järjekorda.

Kontekstivabad grammatikad

Näide: eelnevalt toodud vasak- ja parenderivatsioonile vastab mõlemal juhul sama süntaksipuu

$\text{Exp} \Rightarrow_{lm} \text{Exp} + \text{Exp}$
 $\Rightarrow_{lm} \text{num} + \text{Exp}$
 $\Rightarrow_{lm} \text{num} + \text{Exp} * \text{Exp}$
 $\Rightarrow_{lm} \text{num} + \text{num} * \text{Exp}$
 $\Rightarrow_{lm} \text{num} + \text{num} * \text{num}$

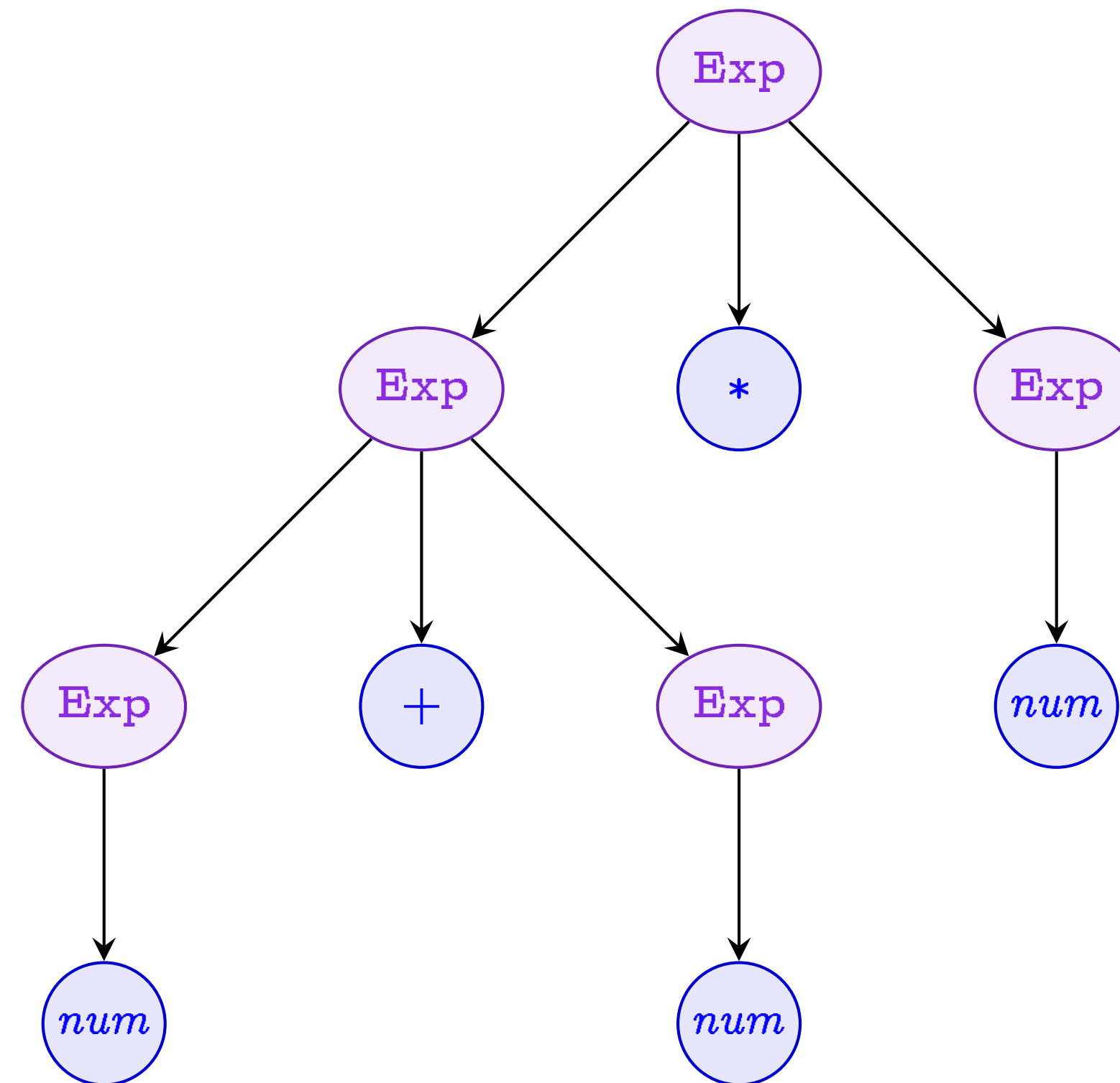
$\text{Exp} \Rightarrow_{rm} \text{Exp} + \text{Exp}$
 $\Rightarrow_{rm} \text{Exp} + \text{Exp} * \text{Exp}$
 $\Rightarrow_{rm} \text{Exp} + \text{Exp} * \text{num}$
 $\Rightarrow_{rm} \text{Exp} + \text{num} * \text{num}$
 $\Rightarrow_{rm} \text{num} + \text{num} * \text{num}$



Kontekstivabad grammatikad

NB! Ühel lausel võib olla mitu erinevat süntaksipuud!

Exp \implies_{lm} **Exp** * **Exp**
 \implies_{lm} **Exp** + **Exp** * **Exp**
 \implies_{lm} *num* + **Exp** * **Exp**
 \implies_{lm} *num* + *num* * **Exp**
 \implies_{lm} *num* + *num* * *num*



Kontekstivabad grammatikad

- KV-grammatika on **mitmene** (**ambiguous**), kui ühe lause jaoks leidub mitu süntaksipuud.
- Iga süntaksipuu jaoks leidub täpselt üks vasak- ja üks parenderivatsioon; seega:
 - ühesel lausel on täpselt üks vasak- ja üks parenderivatsioon;
 - mitmesel lausel on vähemalt kaks vasak- ja parenderivatsiooni.
- Lause erinevad süntaksipuud vastavad reeglina lause semantiliselt erinevatele tõlgendusvõimalustele.
- Mitmese grammatika saab teatud juhtudel (aga mitte alati) teisendada ekvivalentseks üheseks grammatikaks.

Kodutöö (2. nädalat)

- Avaldise grammatika moodustamine
AST klassidega tutvumine
- Käsitsi parseri implementeerimine
Vasakrekursiooni elimineerimine