

Süntaksanalüüs

Kontekstivabad grammatikad
Grammatikate mitmesus

KV grammatika kirjapanek

- Mitte-terminalid esinevad reeglites vasakul!
(antud juhul $N = \{A, B\}$).
- Ülejäänud on terminaalsümbolid
(antud juhul $T = \{a, b\}$).
- Algsümboliks on grammatika esimene
mitte-terminal (antud juhul A).

Esimene
mitte-terminal!

$$A \rightarrow \varepsilon$$

$$B \rightarrow bAa$$

$$A \rightarrow aBb$$

Algsümboliga
peab alustama!

Kõik A reeglid
on mängus!

$$A \rightarrow aBb \rightarrow abAab \rightarrow abab$$

Backus-Naur kuju (BNF)

Näide: olgu $N = \{\text{Exp}\}$ ja $T = \{+, *, (,), \text{num}\}$, siis

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

Exp:
Exp + Exp
Exp * Exp
(Exp)
num

esitab produktsioonireeglite hulka

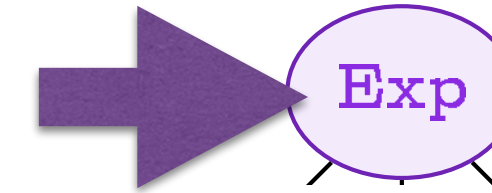
$$P = \{ \text{Exp} \rightarrow \text{Exp} + \text{Exp}, \text{Exp} \rightarrow (\text{Exp}), \\ \text{Exp} \rightarrow \text{Exp} * \text{Exp}, \text{Exp} \rightarrow \text{num} \}.$$

Lause süntaksipuu?

- Selles videos tuletame lause $num + num * num$.
- Tahame teada selle lause struktuuri.
- Väärtustame avaldisi läbides puud, seega on oluline, et saaksime õige puu!

Exp	→	Exp + Exp
		Exp * Exp
		(Exp)
		num

$\text{Exp} \rightarrow$
| $\text{Exp} + \text{Exp}$
| $\text{Exp} * \text{Exp}$
| (Exp)
| num



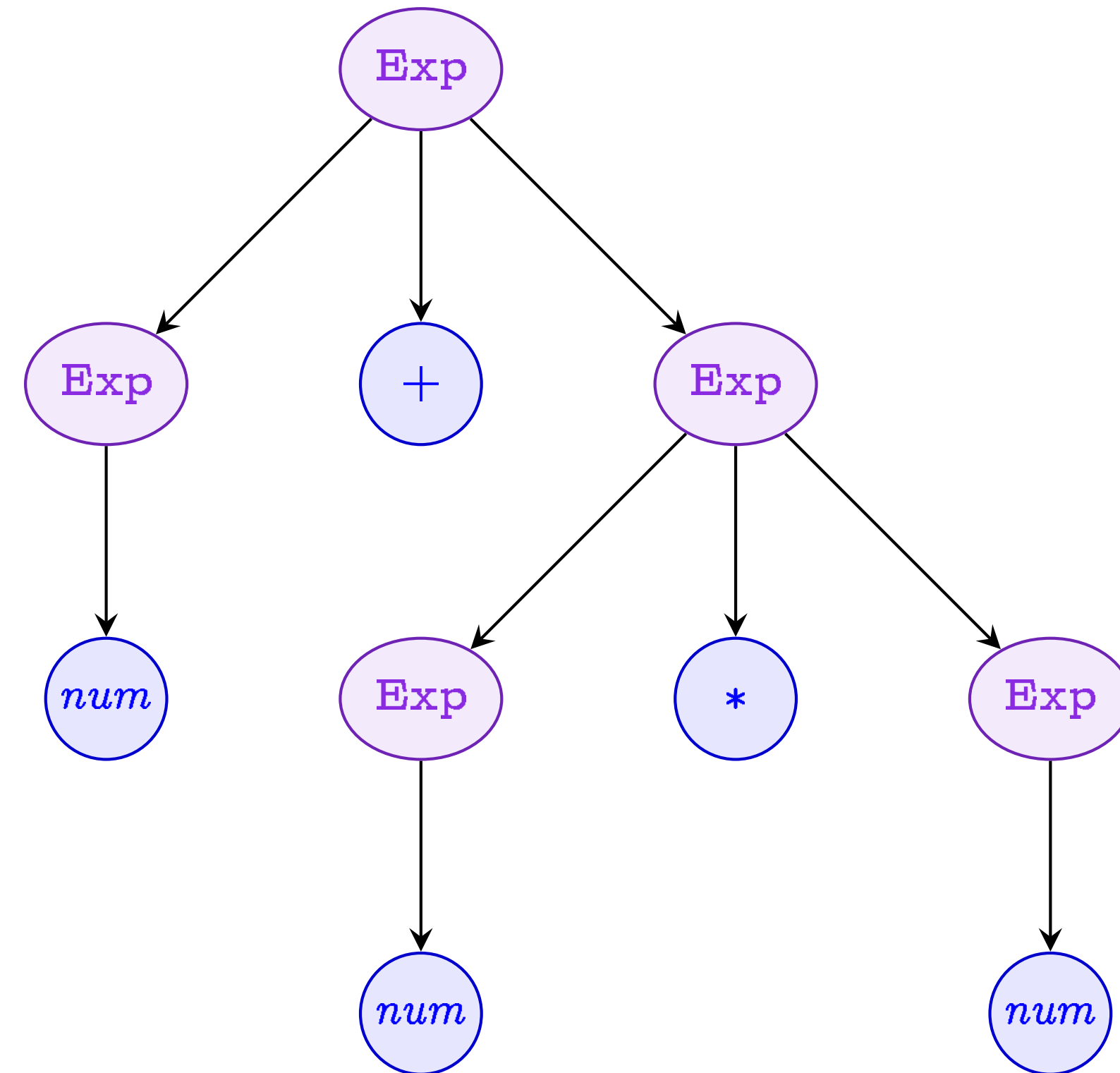
$\text{Exp} \Rightarrow \text{Exp} + \text{Exp}$
 $\Rightarrow \text{num} + \text{Exp}$
 $\Rightarrow \text{num} + \text{Exp} * \text{Exp}$
 $\Rightarrow \text{num} + \text{num} * \text{Exp}$
 $\Rightarrow \text{num} + \text{num} * \text{num}$

Derivatsiooni süntaksipuu

Iga mitte-terminali alluvateks on sümbolid, millega teda asendati!

$\text{Exp} \rightarrow \text{Exp} + \text{Exp}$
 $\text{Exp} \rightarrow \text{Exp} * \text{Exp}$
 $\text{Exp} \rightarrow (\text{Exp})$
 $\text{Exp} \rightarrow \text{num}$

$\text{Exp} \Rightarrow \text{Exp} + \text{Exp}$
 $\Rightarrow \text{num} + \text{Exp}$
 $\Rightarrow \text{num} + \text{Exp} * \text{Exp}$
 $\Rightarrow \text{num} + \text{num} * \text{Exp}$
 $\Rightarrow \text{num} + \text{num} * \text{num}$

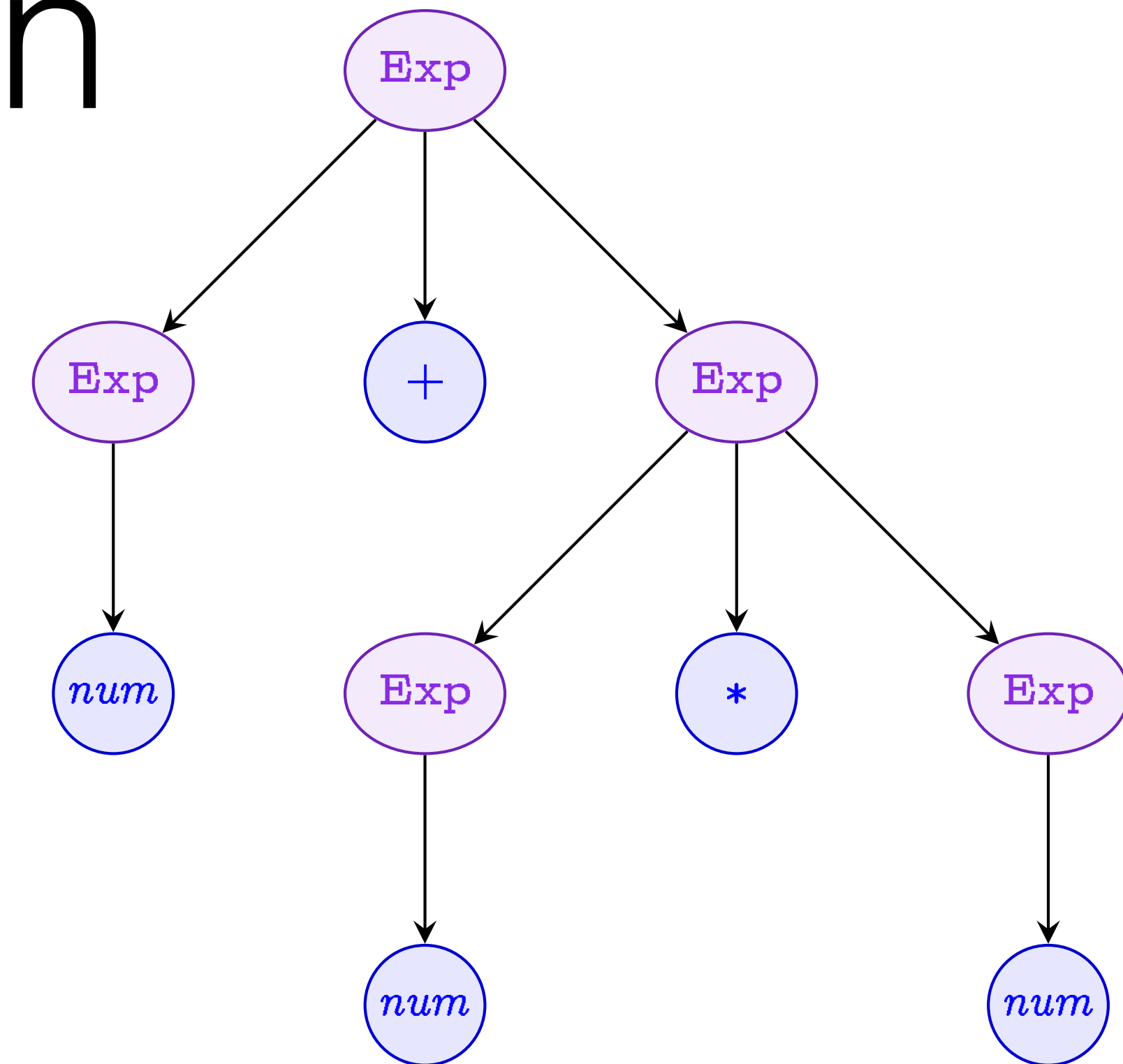


Derivatsiooni süntaksipuu

Puu on valmis. Aga kas see on selle sõna ainus süntaksipuu?
 Kõigepealt, mis on derivatsiooni ja puude seos?

Süntaksipuu definitsioon

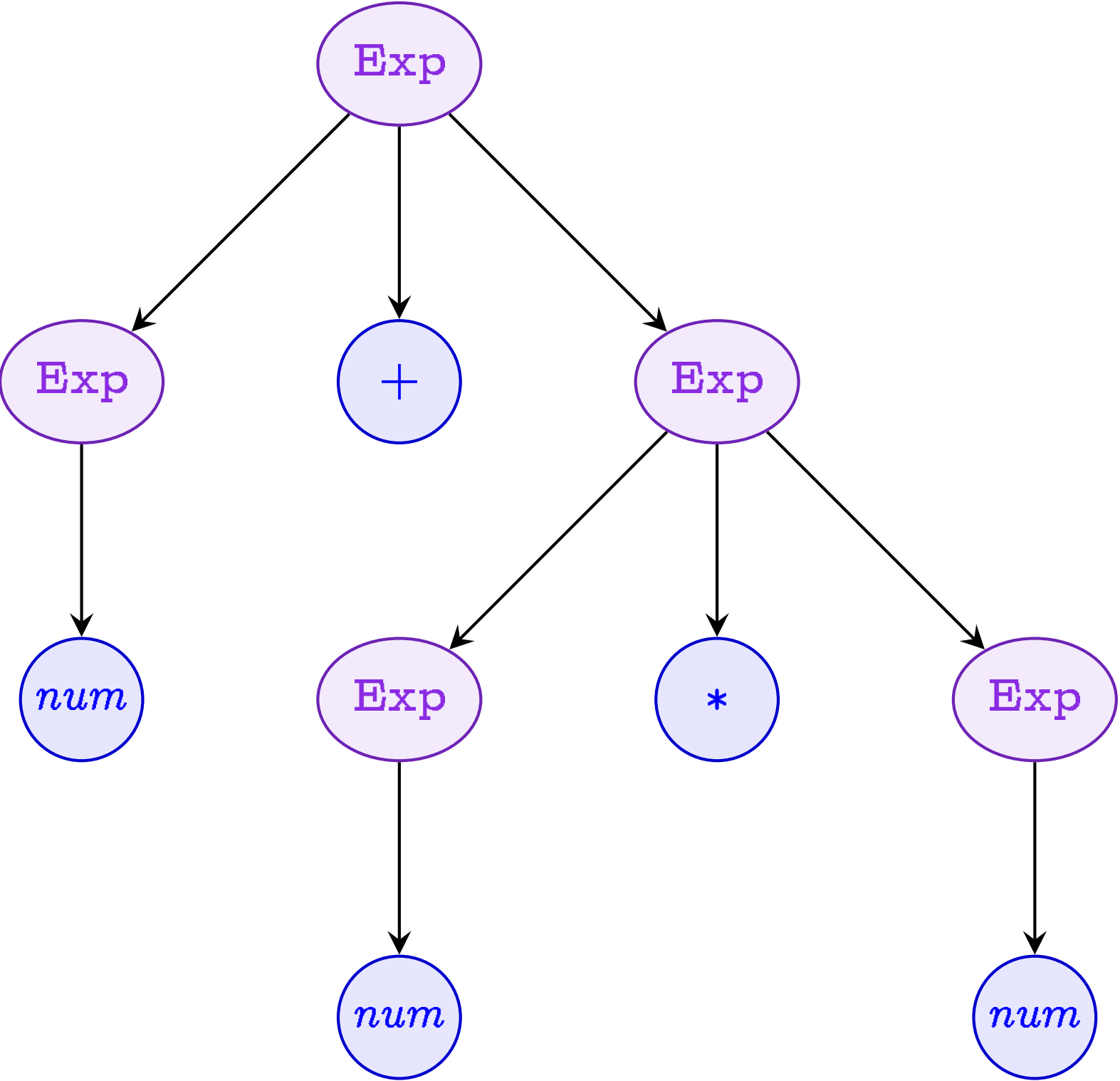
- Iga derivatsioon määrab üheselt **süntaksipuu** (**syntax-tree**, **parse-tree**), mis on järjestatud tippudega puu, kus:
 - puu juur on märgendatud algsümboliga **S**;
 - vahetipud on märgendatud mitteterminaalidega;
 - lehed on märgendatud terminaalsümboliga või tühisümboliga ϵ ;
 - kui vahetipp 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.



Exp $\Rightarrow$ **Exp** + **Exp**
 $\Rightarrow$ *num* + **Exp**
 $\Rightarrow$ *num* + **Exp** * **Exp**
 $\Rightarrow$ *num* + *num* * **Exp**
 $\Rightarrow$ *num* + *num* * *num*

Exp → Exp + Exp
| Exp * Exp
| (Exp)
| num

Exp ⇒ Exp + Exp
⇒ num + Exp
⇒ num + Exp * Exp
⇒ num + num * Exp
⇒ num + num * num

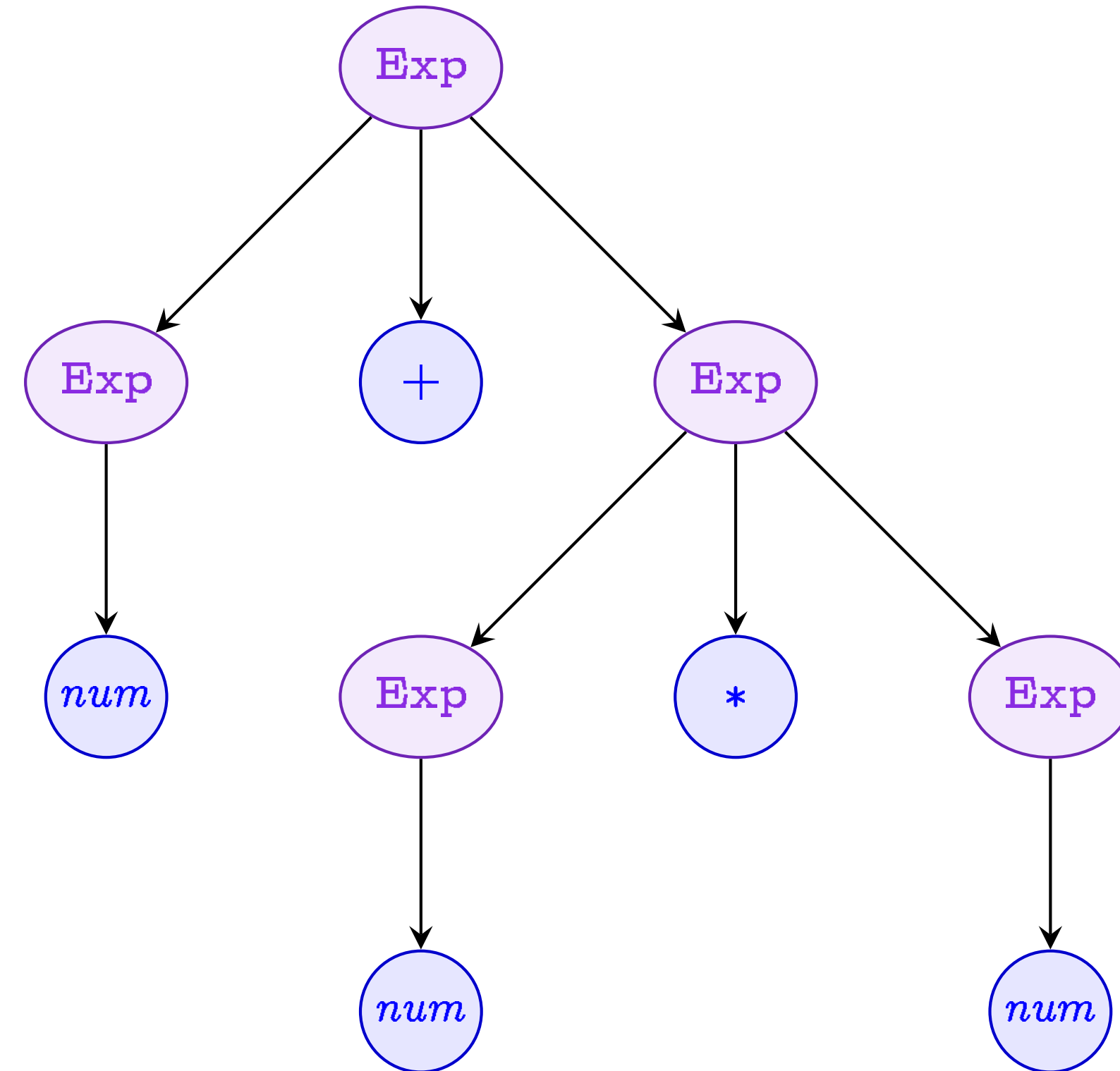


Süntaksipuust derivatsioon

Teeme lihtsalt vastupidi: liigume puust ülalt-alla ja kirjutame asendused järjest.

$\text{Exp} \rightarrow \text{Exp} + \text{Exp}$
 $\text{Exp} \rightarrow \text{Exp} * \text{Exp}$
 $\text{Exp} \rightarrow (\text{Exp})$
 $\text{Exp} \rightarrow \text{num}$

$\text{Exp} \Rightarrow \text{Exp} + \text{Exp}$
 $\Rightarrow \text{Exp} + \text{Exp} * \text{Exp}$
 $\Rightarrow \text{Exp} + \text{Exp} * \text{num}$
 $\Rightarrow \text{Exp} + \text{num} * \text{num}$
 $\Rightarrow \text{num} + \text{num} * \text{num}$



Süntaksipuust derivatsioon

Paaristöö tulemus: meil on erinevad derivatsioonid!?

Kanoonilised derivatsioonid!

- Fikseerime derivatsiooni järjekorra!
- **Vasakderivatsioon**: derivatsiooni igal sammul asendatakse vasakpoolseim mitte-terminal (*left-most derivation*).
- **Parenderivatsioon**: derivatsiooni igal sammul asendatakse parempoolseim mitte-terminal (*right-most derivation*).
- Kui süntaksipuu teisendada näiteks vasak-derivatsiooniks, siis on üks kindel tulemus!

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

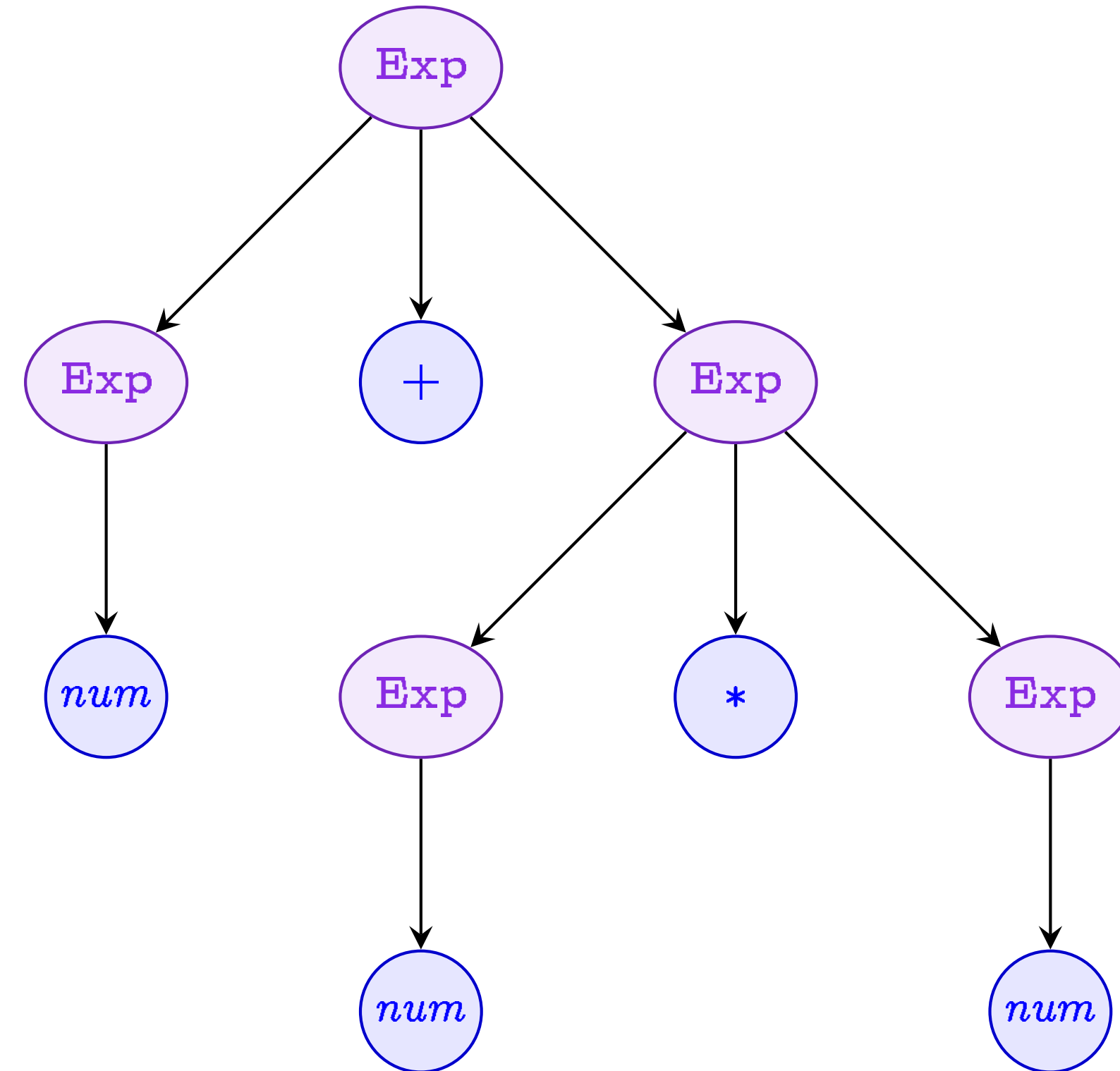
$$\begin{aligned}\text{Exp} &\Rightarrow_{rm} \text{Exp} + \text{Exp} \\ &\Rightarrow_{rm} \text{Exp} + \text{Exp} * \text{Exp} \\ &\Rightarrow_{rm} \text{Exp} + \text{Exp} * \text{num} \\ &\Rightarrow_{rm} \text{Exp} + \text{num} * \text{num} \\ &\Rightarrow_{rm} \text{num} + \text{num} * \text{num}\end{aligned}$$

Valige üks...

- Nüüd teeme ühe valiku (näiteks vasakderivatsioonid) ja elu lõpuni teeme ainult vasakderivatsioonid.
- See on väga oluline, sest see järjekord ei ole meie jaoks oluline! :)

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

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

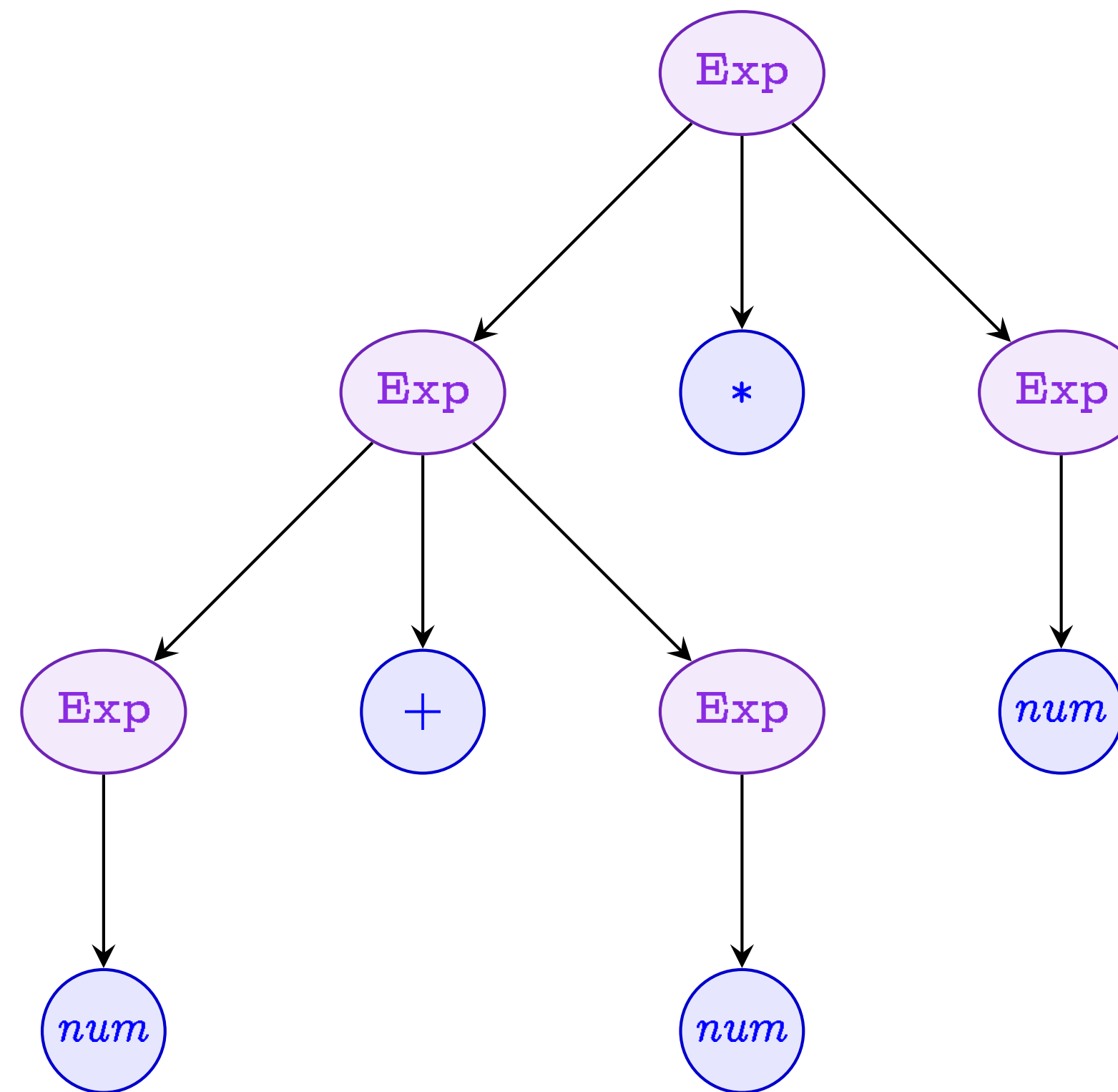


See ei ole oluline!

Need mõlemad derivatsioonid erinevad ainult järjekorra poolest, aga iga vastava mitte-terminali asendamiseks on kasutatud sama reegelit!

$\text{Exp} \rightarrow \text{Exp} + \text{Exp}$
 $\text{Exp} \rightarrow \text{Exp} * \text{Exp}$
 $\text{Exp} \rightarrow (\text{Exp})$
 $\text{Exp} \rightarrow \text{num}$

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



See on aga oluline!

Siin on jälle vasak-derivatsioon, aga juurtipus kasutati teist reeglit (korrutamist).
 Tulemuseks on teine puu ja kui avaldise tähendust jälgida, siis vale puu...

Mitmesus!

- Grammatika on **mitmene** (*ambiguous*), kui ühe lause jaoks leidub mitu süntaksipuud.
- Iga süntaksipuu jaoks leidub täpselt üks vasakderivatsioon, seega (parenderivatsioonide kohta samamoodi)
 - Ühesel lausel on täpselt üks vasakderivatsioon.
 - Mitmesel lausel on vähemalt kaks vasakderivatsiooni.
- Mitmesus on halb, sest avaldise väärtustamine käib süntaksipuu põhjal ja vale puu annab vale tulemuse.
- Mõnedest grammatikatest on võimalik mitmesust eemaldada!

Kokkuvõte

- Kanoonilised derivatsioonid fikseerivad lausevormis ümberkirjutatava mitte-terminali: vasakderivatsioonis alati vasakpoolseim.
- Igale (suvalisele) derivatsioonile vastab täpselt üks süntaksipuu. Süntaksipuule vastab täpselt üks vasakderivatsioon.
- Mitmesel sõnal on rohkem kui üks süntaksipuu.
Mitmesel sõnal on rohkem kui üks vasakderivatsioon.

(Sellel slaidil võid sõna „vasak“ asendada sõnaga „parem“.)