CSE-233 : Section A
Summer 2020

Derivation/Parse Tree

Reference:
Book2 Chapter 2.1
Reference2, Reference3

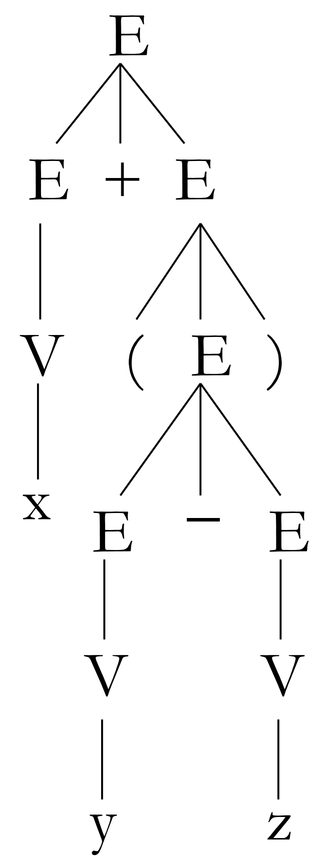
Md. Saidul Hoque Anik
anik@cse.uiu.ac.bd

Derivation/Parse Tree

Derivations can also be represented using derivation/parse trees

$$E \rightarrow E + E \mid E \text{ - } E \mid (E) \mid V$$
$$V \rightarrow x \mid y \mid z$$

$$\begin{aligned}
E &\Rightarrow E + E \\
&\Rightarrow V + E \\
&\Rightarrow x + E \\
&\Rightarrow x + (E) \\
&\Rightarrow x + (E - E) \\
&\Rightarrow x + (V - E) \\
&\Rightarrow x + (y - E) \\
&\Rightarrow x + (y - V) \\
&\Rightarrow x + (y - z)
\end{aligned}$$



Definition

A derivation/parse tree for a CFG G is an ordered tree with labels on the nodes such that

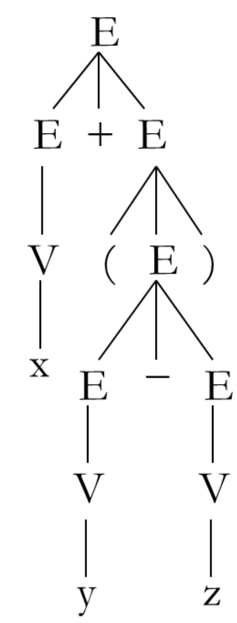
- Every internal node is labeled by a variable
- Every leaf is labeled by a terminal or e
- Leaves labeled by e have no siblings
- If a node is labeled A and has children A_1, ..., A_k from left to right, then the rule

$$A \longrightarrow A_1 \ldots A_k$$

is a production in G.

$$\boxed{\begin{aligned} &E \rightarrow E + E \mid E \text{ - } E \mid (E) \mid V \\ &V \rightarrow x \mid y \mid z \end{aligned}}$$

$$\begin{aligned} E &\Rightarrow E + E \\ &\Rightarrow V + E \\ &\Rightarrow x + E \\ &\Rightarrow x + (E) \\ &\Rightarrow x + (E - E) \\ &\Rightarrow x + (V - E) \\ &\Rightarrow x + (y - E) \\ &\Rightarrow x + (y - V) \\ &\Rightarrow x + (y - z) \end{aligned}$$



Practice

Obtained by applying production rule

$S \rightarrow bSb \mid a \mid b$

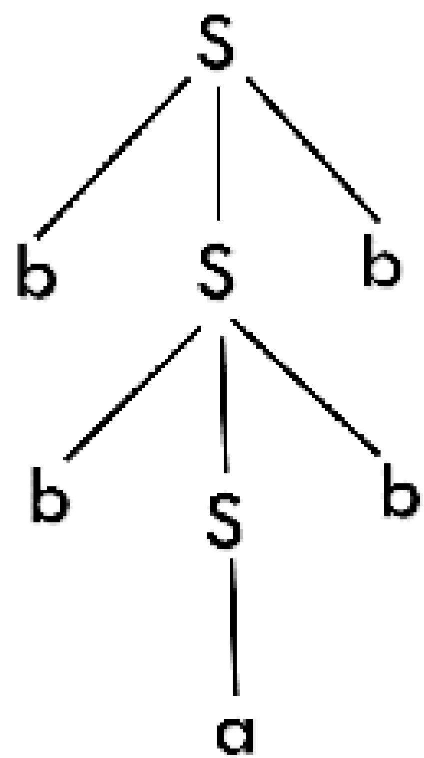
generate **bbabb**

Solution

Obtained by applying production rule

$S \rightarrow bSb \mid a \mid b$

generate **bbabb**



Practice 2

Obtained by applying production rule

$S \rightarrow aB \mid bA$
$A \rightarrow a \mid aS \mid bAA$
$B \rightarrow b \mid bS \mid aBB$

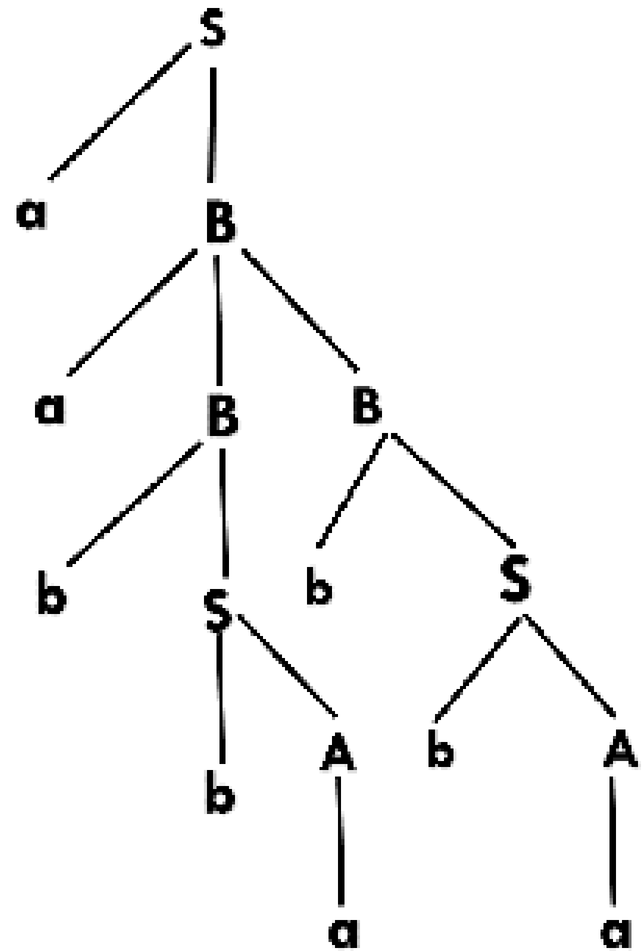
generate **aabbabba**

Solution

Obtained by applying production rule

$S \rightarrow aB \mid bA$
$A \rightarrow a \mid aS \mid bAA$
$B \rightarrow b \mid bS \mid aBB$

generate **aabbabba**



Practice 3

Obtained by applying production rule

$S \rightarrow AB \mid \varepsilon$
$A \rightarrow aB$
$B \rightarrow Sb$

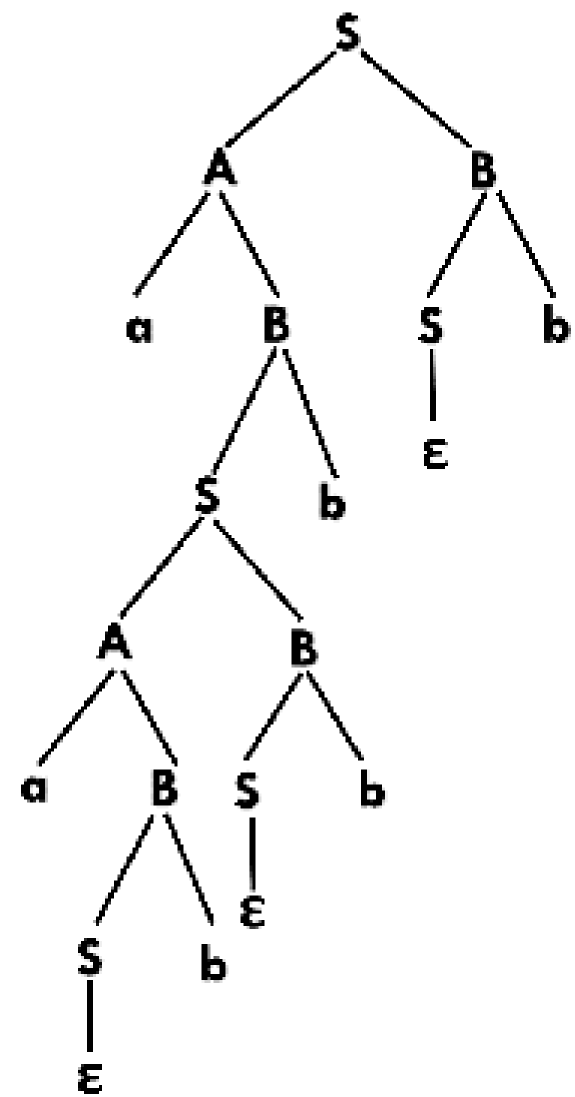
generate **aabbbb**

Solution

Obtained by applying production rule

$S \to AB \mid \varepsilon$
$A \to aB$
$B \to Sb$

generate **aabbbb**



Left Derivation Tree

Obtained by applying production rule to the leftmost variable in each step

$A \rightarrow aBA \mid aAA \mid \varepsilon$
$B \rightarrow AbB \mid ba$

Generate aabaa

Right Derivation Tree

Obtained by applying production rule to the rightmost variable in each step

$A \to aBA \mid aAA \mid \varepsilon$
$B \to AbB \mid ba$

Generate aabaa

Ambiguous Grammar

A grammar is said to be ambiguous if there exists two or more derivation tree for a string w (that means two or more left derivation trees, or two or more right derivation trees)

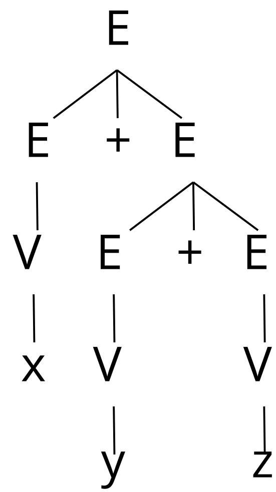
Terminals: a,b,+,*

$S \rightarrow S*S \mid S+S \mid a \mid b$

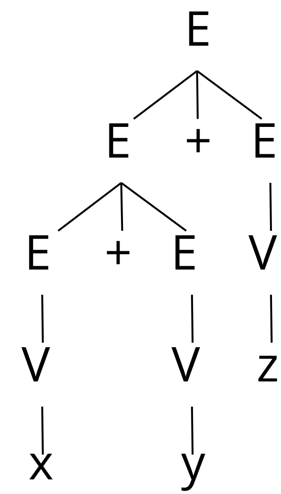
Generate a + a * b

Why ambiguity is important

The parse tree represents the intended meaning



x + y + z



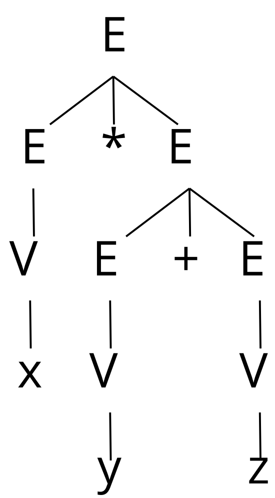
“first add y and z,
and then add this to x”

“first add x and y,
and then add z to this”

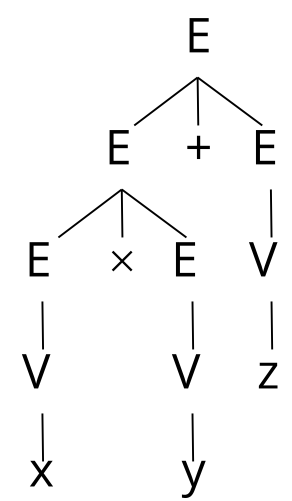
Problem in ambiguity

For adding a multiplication rule directly, we may get undesirable results

$E \rightarrow E + E \mid E - E \mid E \times E \mid (E) \mid V$
$V \rightarrow x \mid y \mid z$



$x \times y + z$



“first y + z, then x ×”

“first x × y, then + z”

Disambiguation

Sometimes we can rewrite the grammar to remove the ambiguity

$$E \rightarrow E + E \mid E - E \mid E \times E \mid (E) \mid V$$
$$V \rightarrow x \mid y \mid z$$

Rewritten grammar so $\times$ cannot be broken by +:

$$E \rightarrow T \mid E + T \mid E - T$$
$$T \rightarrow F \mid T \times F$$
$$F \rightarrow (E) \mid V$$
$$V \rightarrow x \mid y \mid z$$

T stands for term: x * (y + z)
F stands for factor: x, (y + z)

A term always splits into factors

A factor is either a variable or a parenthesized expression

Remarks

- It’s not always possible to disambiguate the grammar as the rules are context dependent
- Rules used in Compilers must be unambiguous

Example

$S \rightarrow AB \mid aaB$
$A \rightarrow a \mid Aa$
$B \rightarrow b$

Derive the string **aab** from the grammar. Convert the grammar to unambiguous

Example

$S \rightarrow AB \mid aaB$
$A \rightarrow a \mid Aa$
$B \rightarrow b$

Ask yourself-

1. What is causing the ambiguity?
2. Where are we allowed to make multiple choices?
3. Can we remove the multiple choice option without the removing the power of generating the same strings?

Example

$S \rightarrow AB \mid aaB$
$A \rightarrow a \mid Aa$
$B \rightarrow b$

Ask yourself-

1. What is causing the ambiguity?
2. Where are we allowed to make multiple choices?
3. Can we remove the multiple choice option without the removing the power of generating the same strings?

Solution

$$S \rightarrow AB$$

$$A \rightarrow a \mid Aa$$

$$B \rightarrow b$$