Syntax Analysis

Sudakshina Dutta

IIT Goa

8^{th} February, 2022

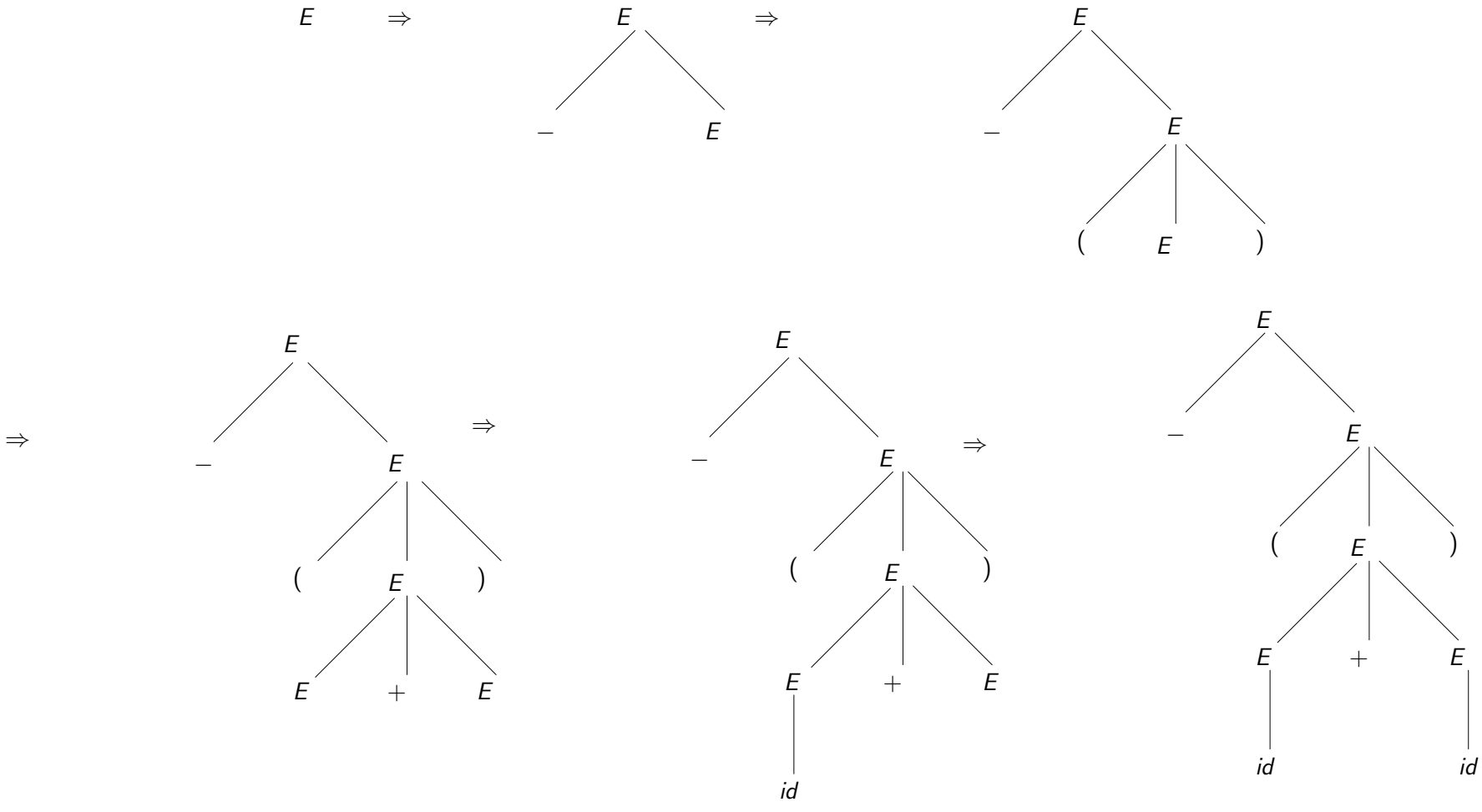
Example

An example context-free grammar

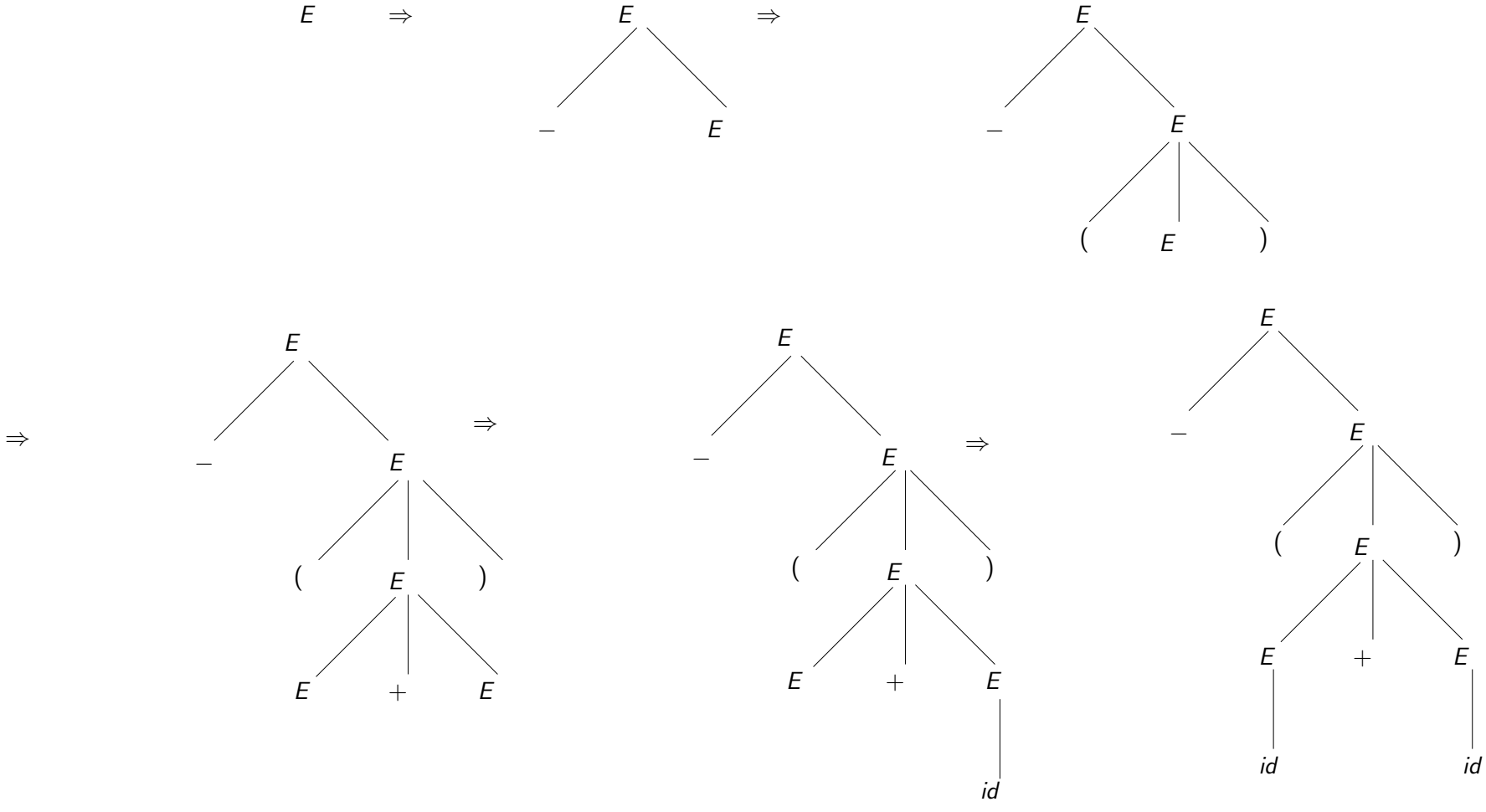
- $E \rightarrow E + E$
- $E \rightarrow E * E$
- $E \rightarrow -E$
- $E \rightarrow (E)$
- $E \rightarrow id$

with the sentence $id + id * id$

Leftmost derivation



Rightmost derivation

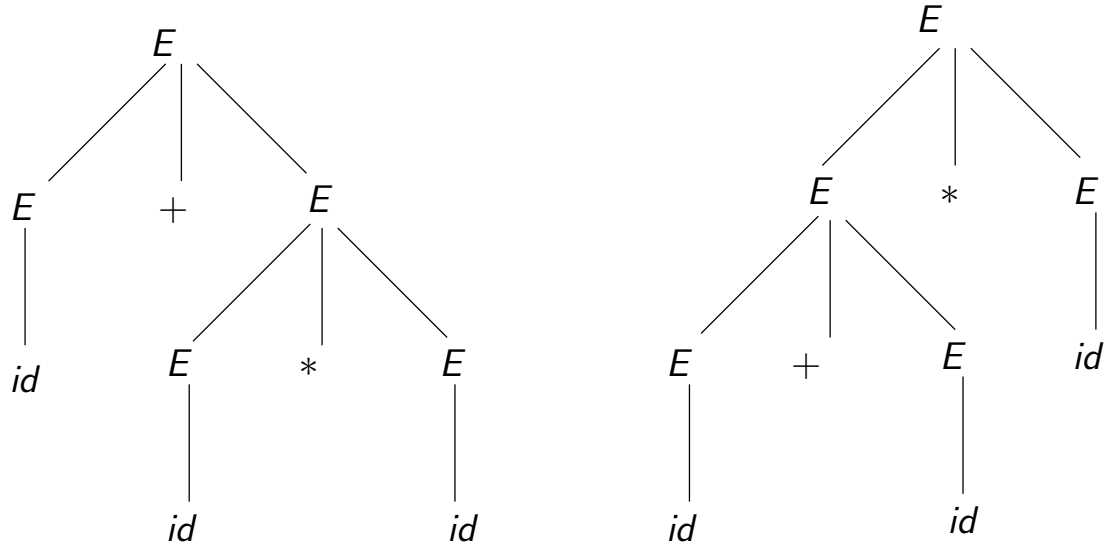


Ambiguity

- From the compiler's perspective, it is important that each sentence in the language defined by the grammar to have a unique rightmost (or leftmost) derivation
- If multiple rightmost (or leftmost) derivations exist for some sentence, then multiple distinct meanings can be possible for a distinct sentence
- Bad for a programming language

Two parse trees

- Consider the string $id + id * id$



- The same tree is generated by both the derivations

Consider the following grammar

- $stmt \rightarrow if\ \ expr\ \ then\ stmt$
- $stmt \rightarrow if\ \ expr\ \ then\ stmt\ \ else\ stmt$
- $stmt \rightarrow other$

and the following string

- $if\ \ E_1\ \ then\ if\ \ E_2\ \ then\ S_1\ \ else\ S_2$

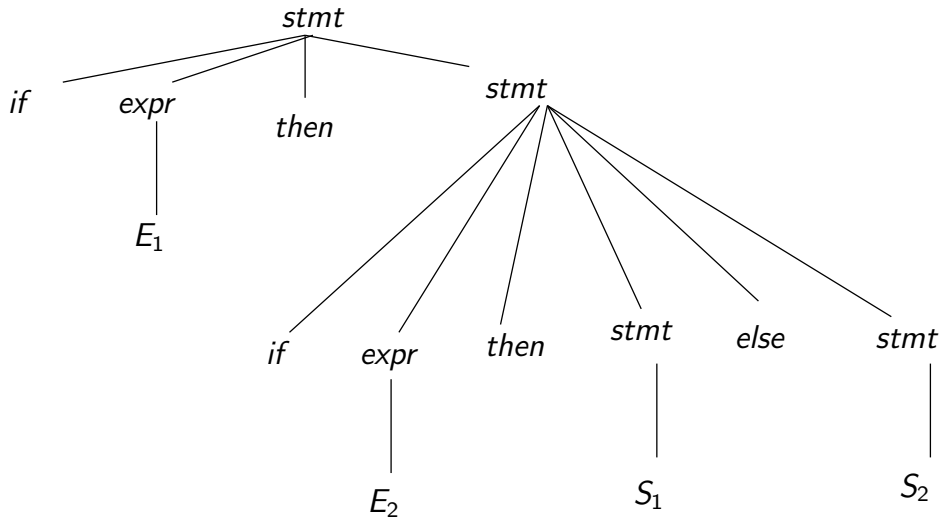
Ambiguity

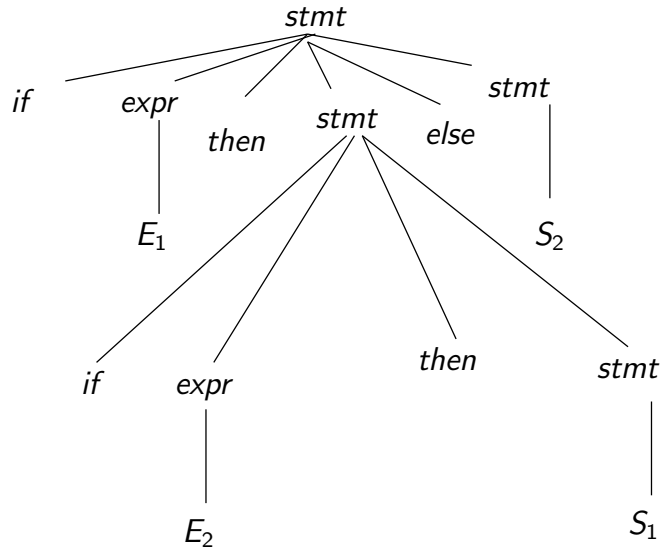
Consider the follpwing grammar

- $stmt \rightarrow if\ expr\ then\ stmt$
- $stmt \rightarrow if\ expr\ then\ stmt\ else\ stmt$
- $stmt \rightarrow other$

and the following string

- $if\ E_1\ then\ (if\ E_2\ then\ S_1\ else\ S_2)$
- $if\ E_1\ then\ (if\ E_2\ then\ S_1)\ else\ S_2$





The ambiguity can be resolved if some convention from the programming language can be specified in the grammar

Here, The grammar is said to be having an ambiguity since the SAME STRING can have 2 parse/syntax trees under the given grammar.

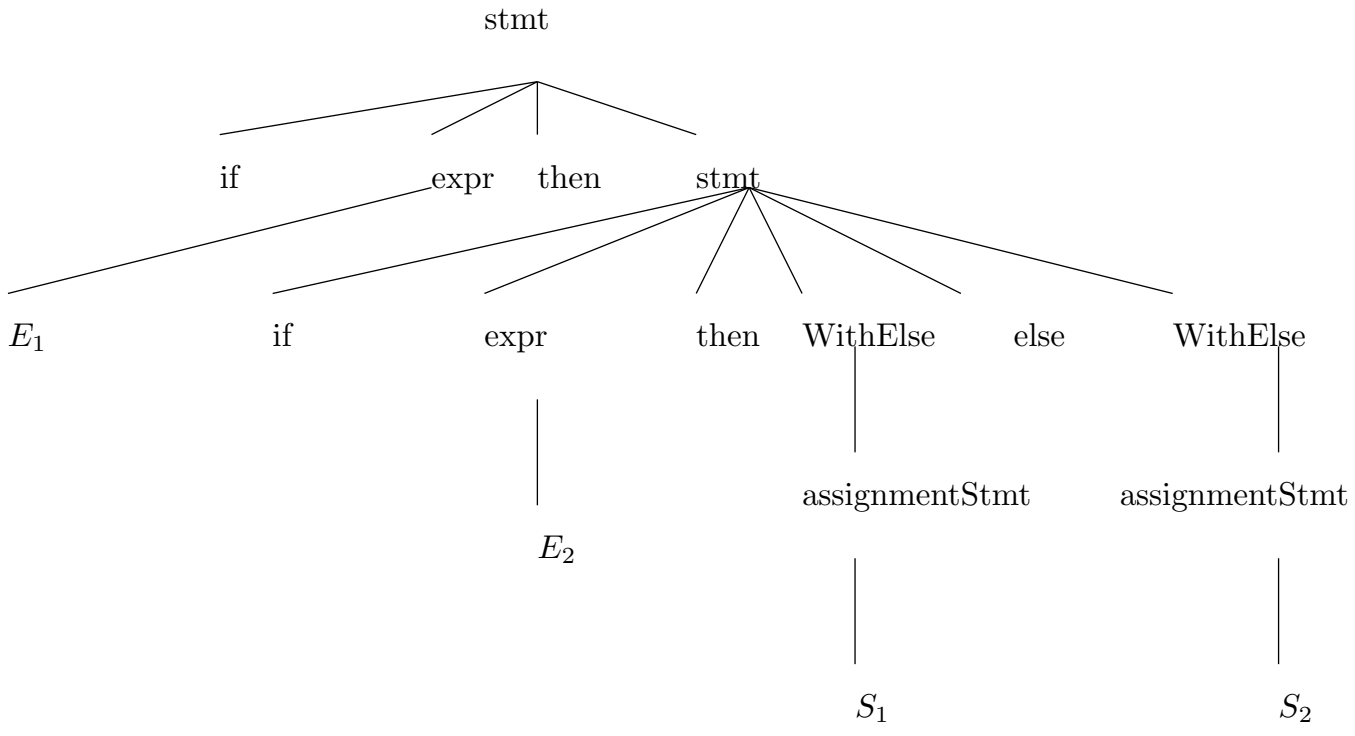
Resolving ambiguity

- $stmt \rightarrow if\ \ expr\ \ then\ stmt$
- $stmt \rightarrow if\ \ expr\ \ then\ \ \ WithElse\ else\ stmt$
- $stmt \rightarrow assignmentStmt$
- $WithElse \rightarrow if\ \ expr\ \ then\ WithElse\ else\ WithElse$
- $WithElse \rightarrow assignmentStmt$

We need to parse

- $if\ \ E_1\ \ then\ if\ \ E_2\ \ then\ S_1\ \ else\ S_2$

Resolving ambiguity



To resolve ambiguity..

- Generally meaning and grammatical structure interact

For example,

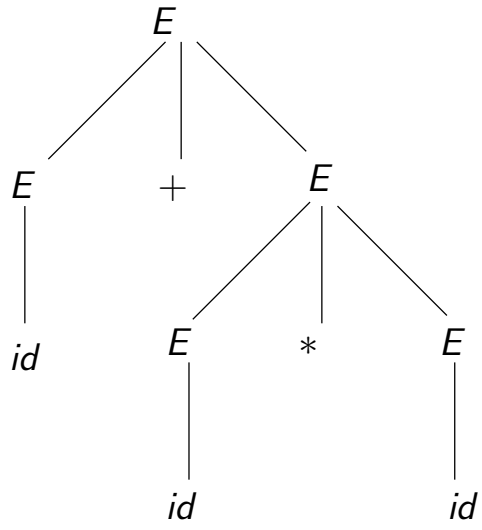
- $E \rightarrow E + E$
- $E \rightarrow E * E$
- $E \rightarrow id$

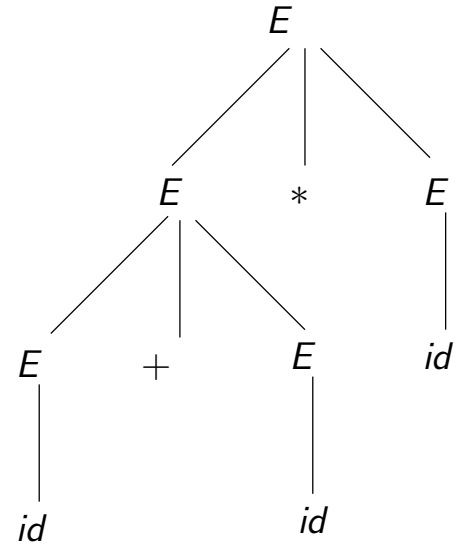
is changed to

- $E \rightarrow E + T \mid T$
- $T \rightarrow T * F \mid F$
- $F \rightarrow id$

The change considers the precedence of operators

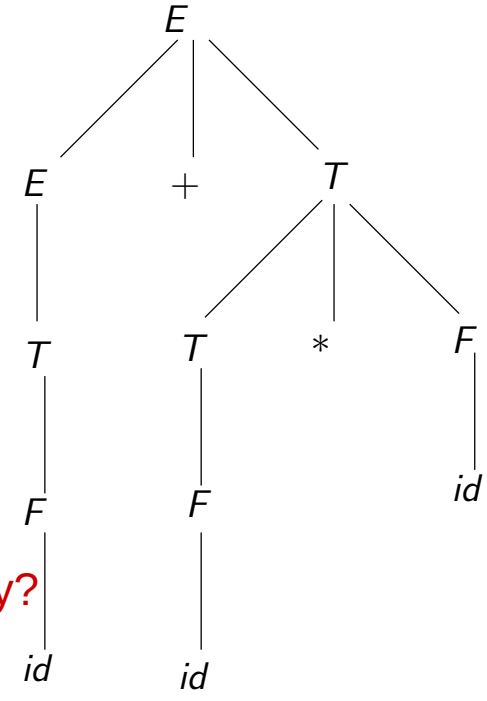
Don't get confused. Here T & F refers to variable/Expression Names only.. They dont stand for True & False resp.





$E \to E + E$
$E \to E * E$
$E \to id$

$E \to E + T | T$
$T \to T * F | F$
$F \to id$



How to get this type of idea that resolves ambiguity?