

BU*P*

ideas

5 February 2018
Aftab Hussain
University of California, Irvine

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid \text{id}$

$E \leftarrow \text{id} * \text{id}$

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid \text{id}$$

E ←————— id * id

As we do a BUP, notice 2 things:

- 1) How we scan & process the input*
- 2) How we build the parse tree*

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \leftarrow id * id$

$id * id$

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid \text{id}$

$E \leftarrow \text{id} * \text{id}$

$F * \text{id}$

id

$\text{id} * \text{id}$

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \leftarrow id * id$

$F * id$

id

$id * id$

*Note how
the input
has
changed.*

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

E



id * id

F * id

id * id



id

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \leftarrow id * id$

$T * id$

$F * id$

$id * id$

F

id

id

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid \text{id}$$

E

Do we reduce T

by E or do we

shift ahead?

T * id

$$F * \text{id}$$
$$\text{id} * \text{id}$$

F

id

id

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid \text{id}$

$E \leftarrow \text{id} * \text{id}$

$T * \text{id}$
—
 F
 \mid
 id

$F * \text{id}$
 \mid
 id

$\text{id} * \text{id}$

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \leftarrow id * id$

$T * id$

$F * id$

$id * id$

F

id

id

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \leftarrow id * id$

$T * F$

$T * id$

$F * id$

$id * id$

$\begin{array}{c} | \quad | \\ F \quad id \\ | \end{array}$

$\begin{array}{c} | \\ F \\ | \\ id \end{array}$

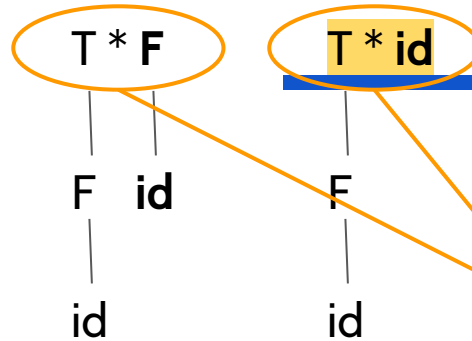
$\begin{array}{c} | \\ id \end{array}$

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

E ←———— id * id



*Note that this time
we have reduced a
part of the scanned
portion, not the
whole*

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E

$$\text{id} * \text{id}$$

T * F

$$T * id$$
$$F * \text{id}$$
$$\text{id} * \text{id}$$

F id

F

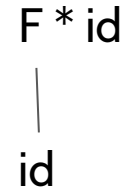
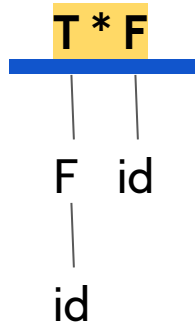
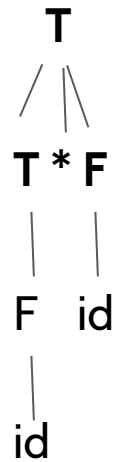
id

id

id

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

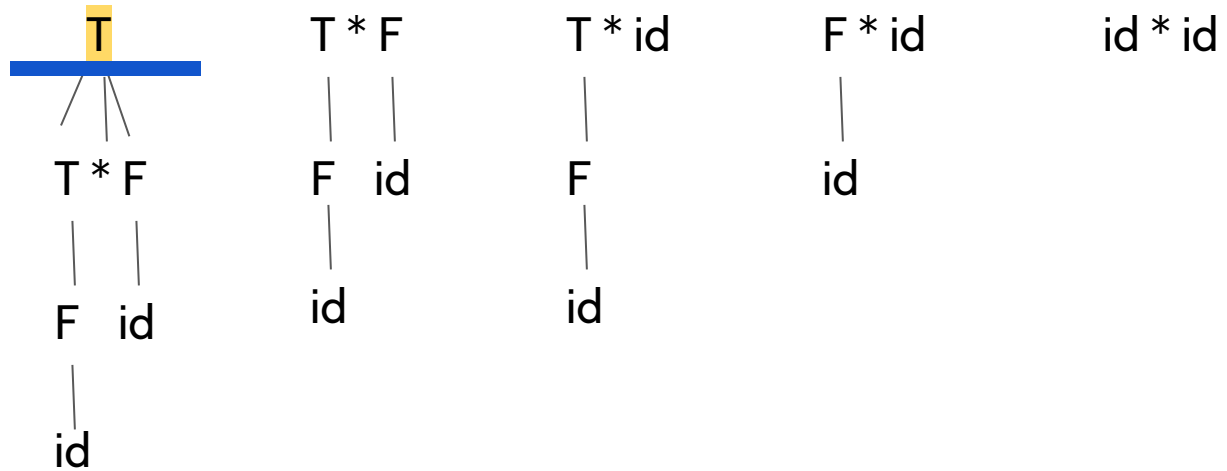
E ← id * id



id * id

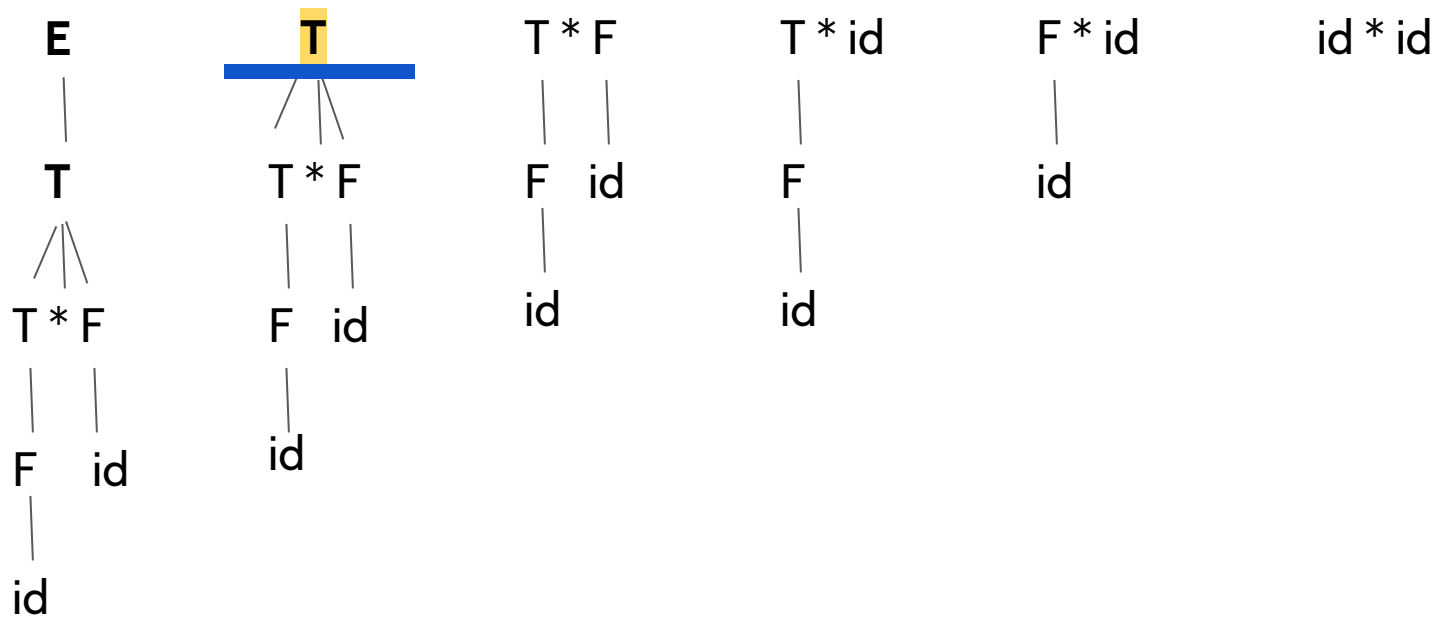
$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E ← id * id



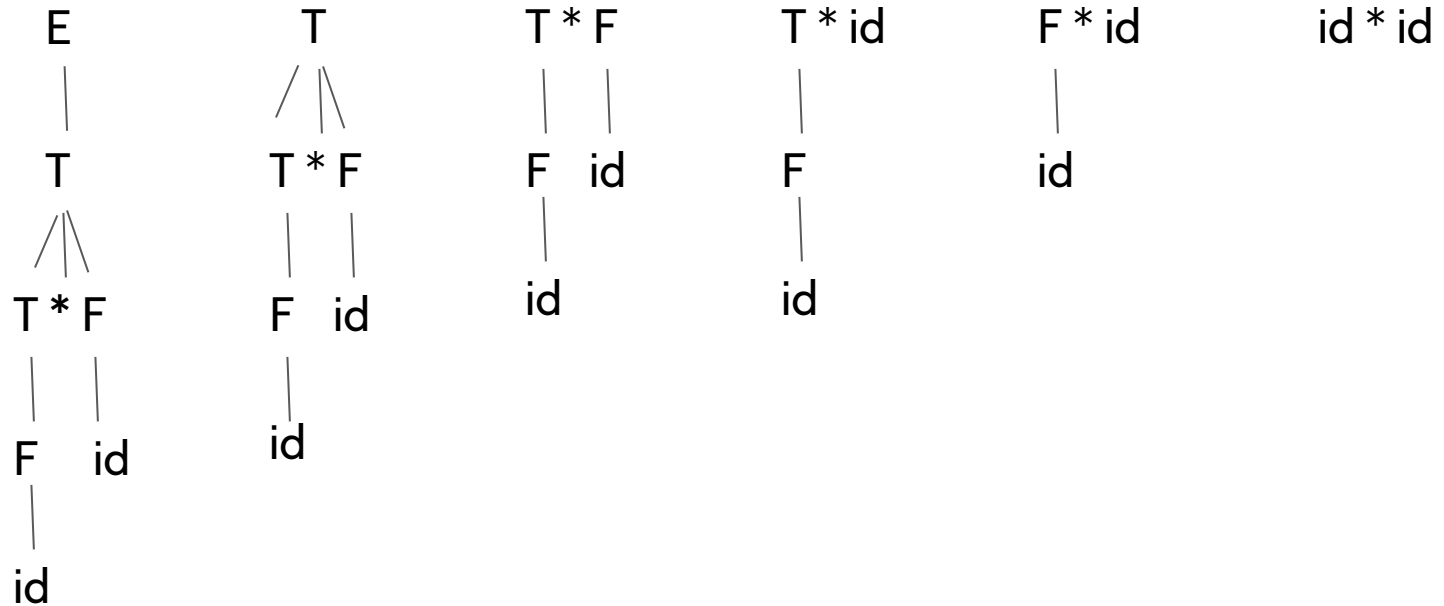
$$F \rightarrow (E) \mid id$$

$E \longleftarrow \text{id} * \text{id}$



$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E ← id * id



BUP in a nutshell



Grammar

The diagram consists of two rectangular boxes. The top box is light blue and contains the word 'Grammar'. The bottom box is light gray and contains the word 'Input'. The boxes are positioned one above the other, with the gray box being wider than the blue box.

Input

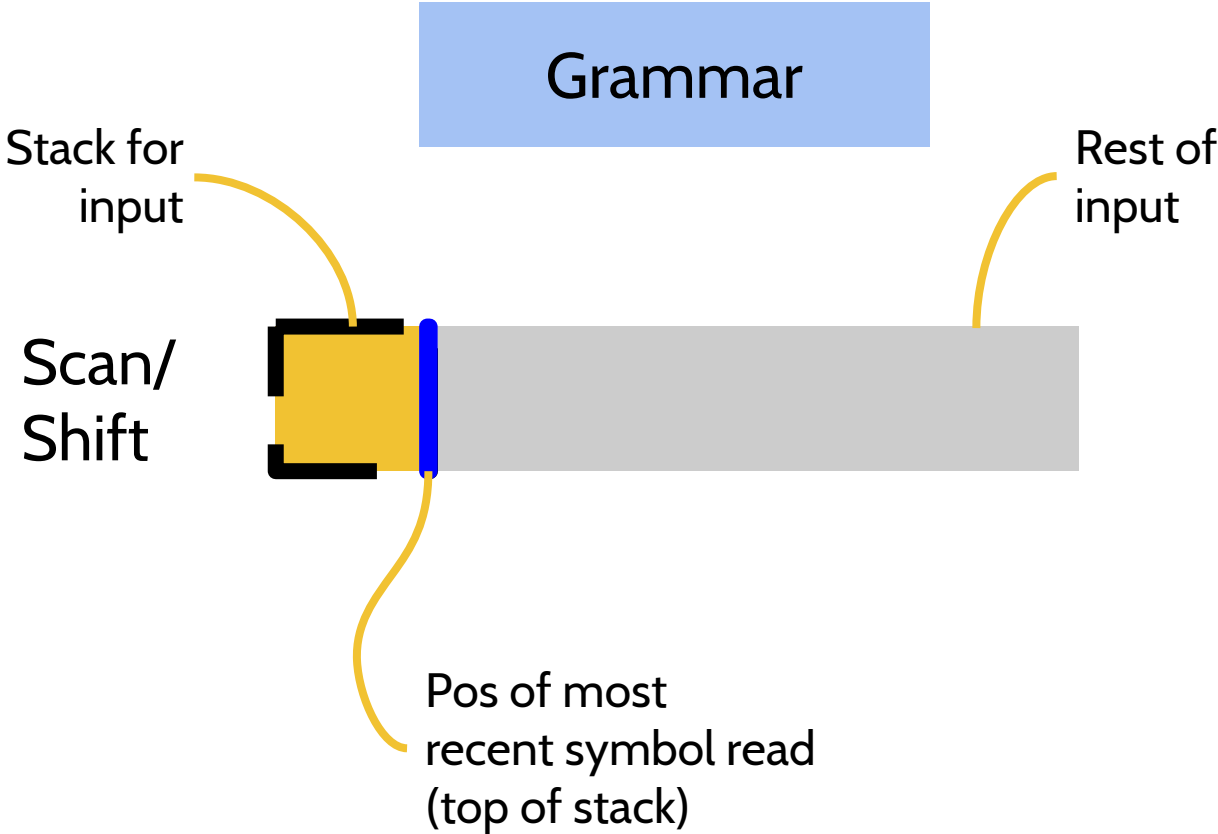
BUP in a nutshell

Grammar

Scan/
Shift



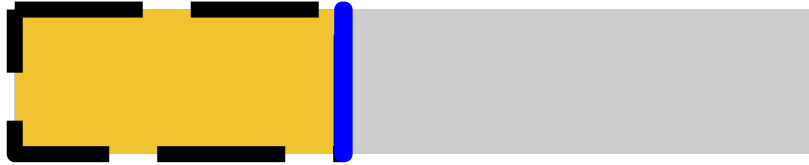
BUP in a nutshell



BUP in a nutshell

Grammar

Scan/
Shift



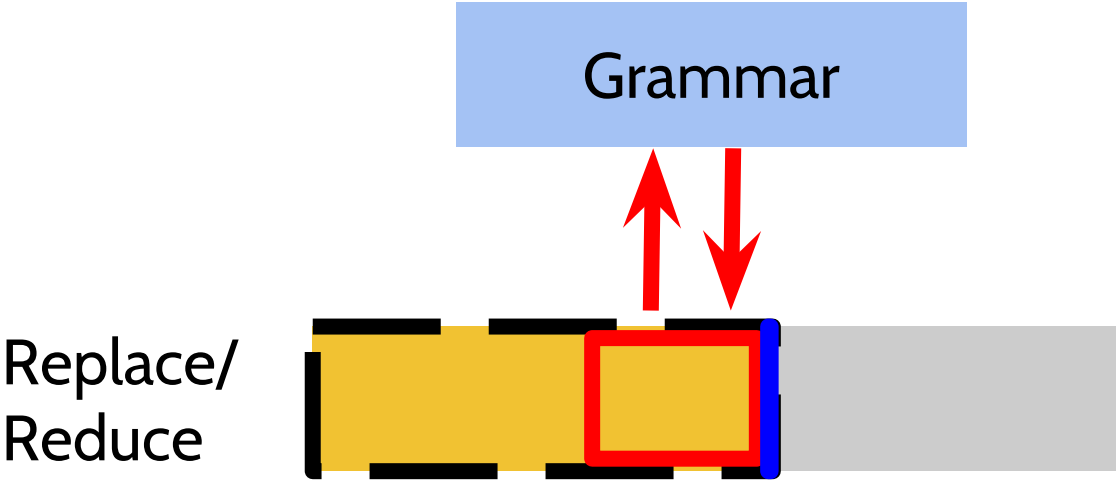
BUP in a nutshell

Grammar

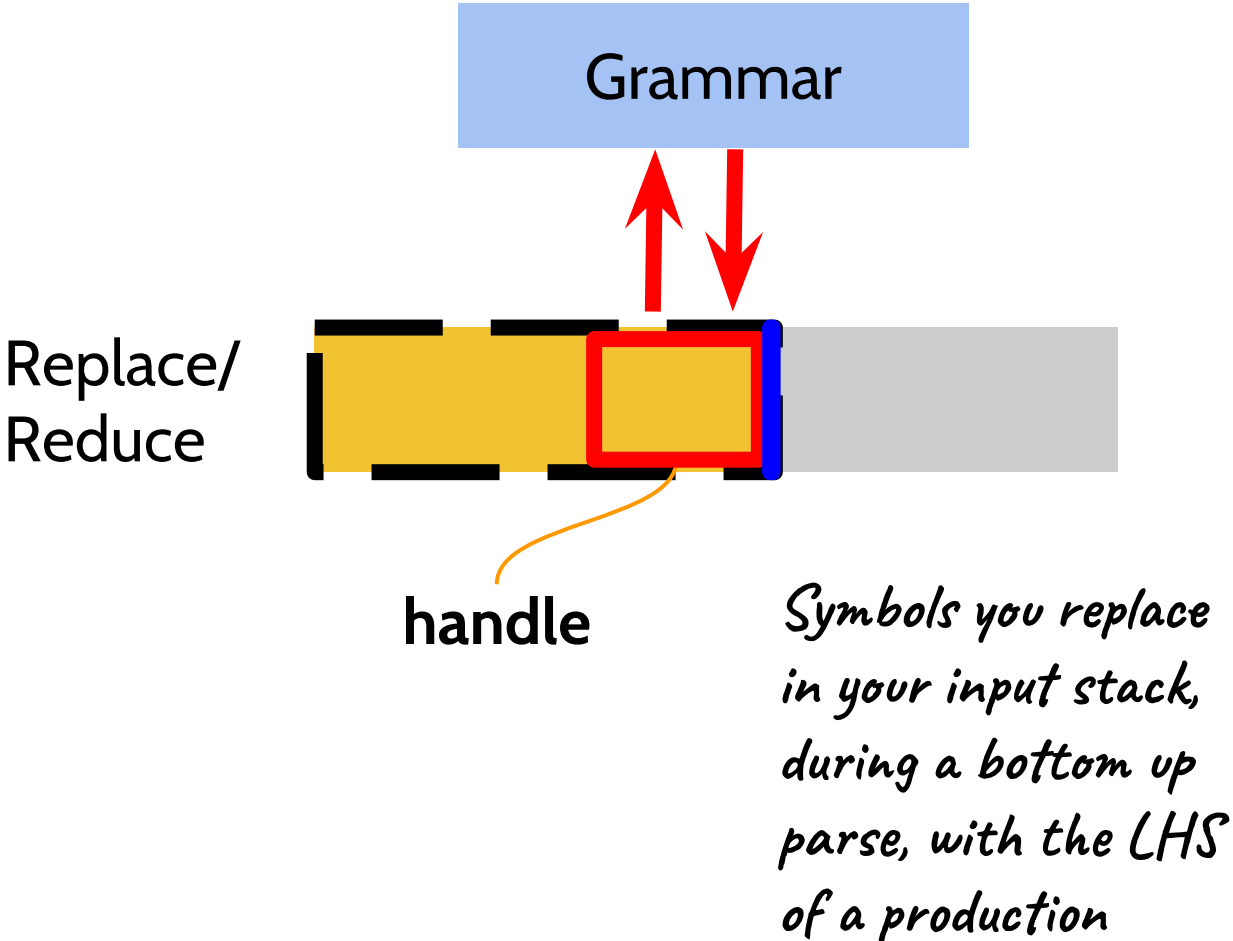
Scan/
Shift



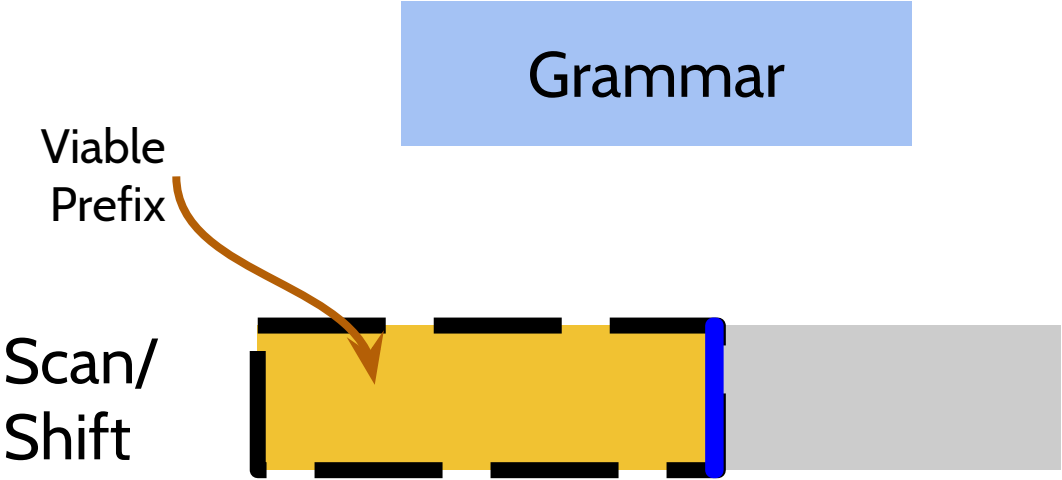
BUP in a nutshell



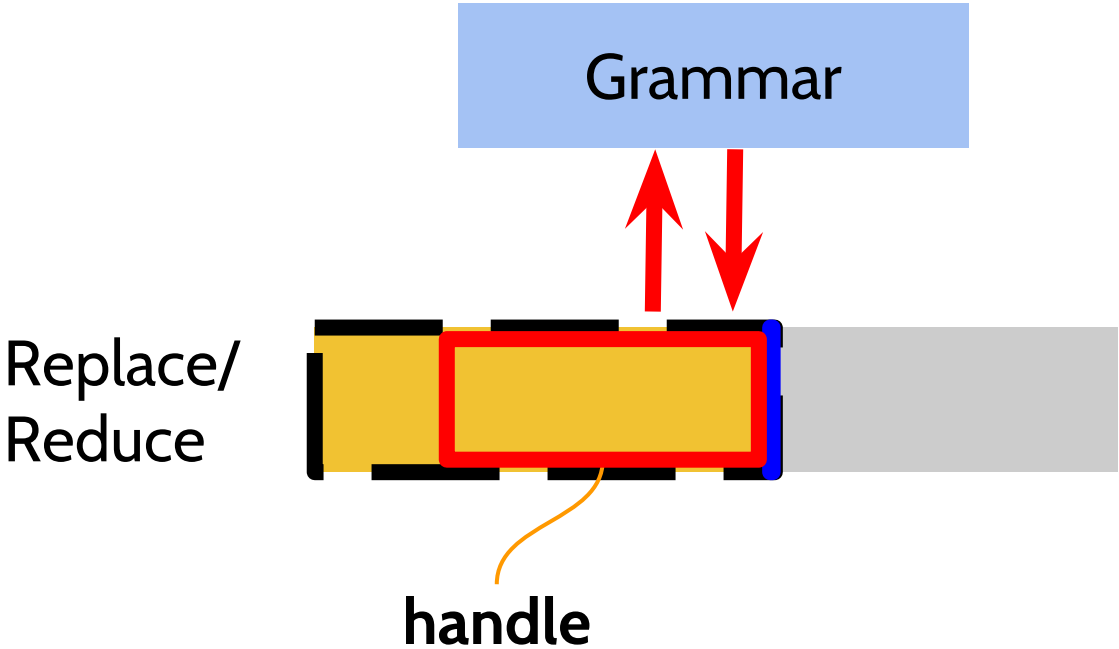
BUP in a nutshell



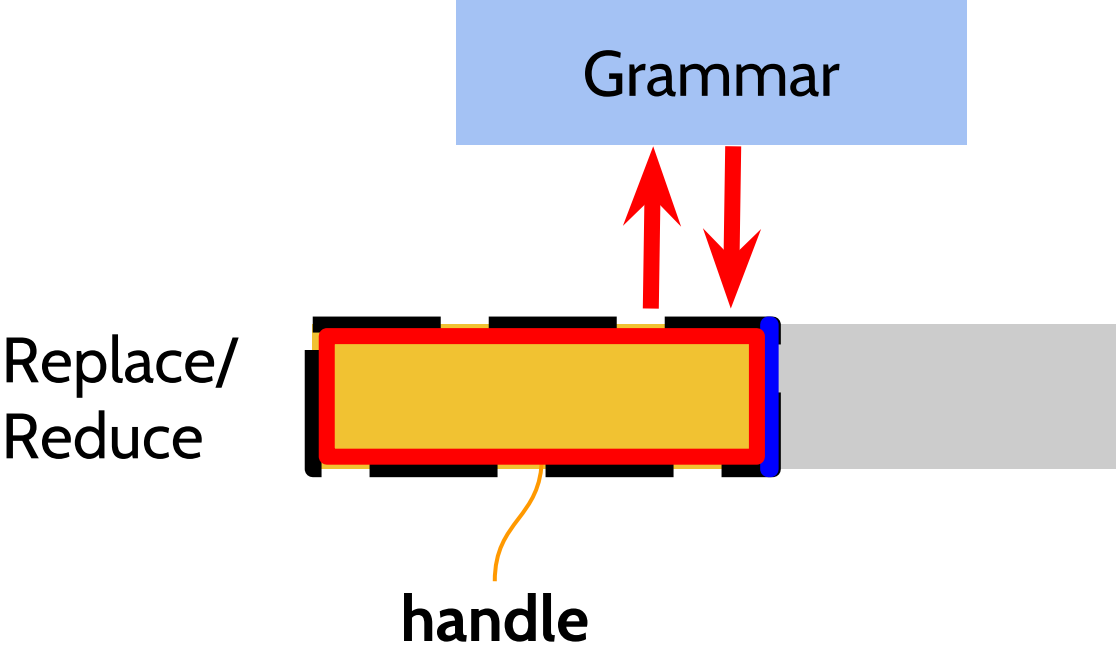
BUP in a nutshell



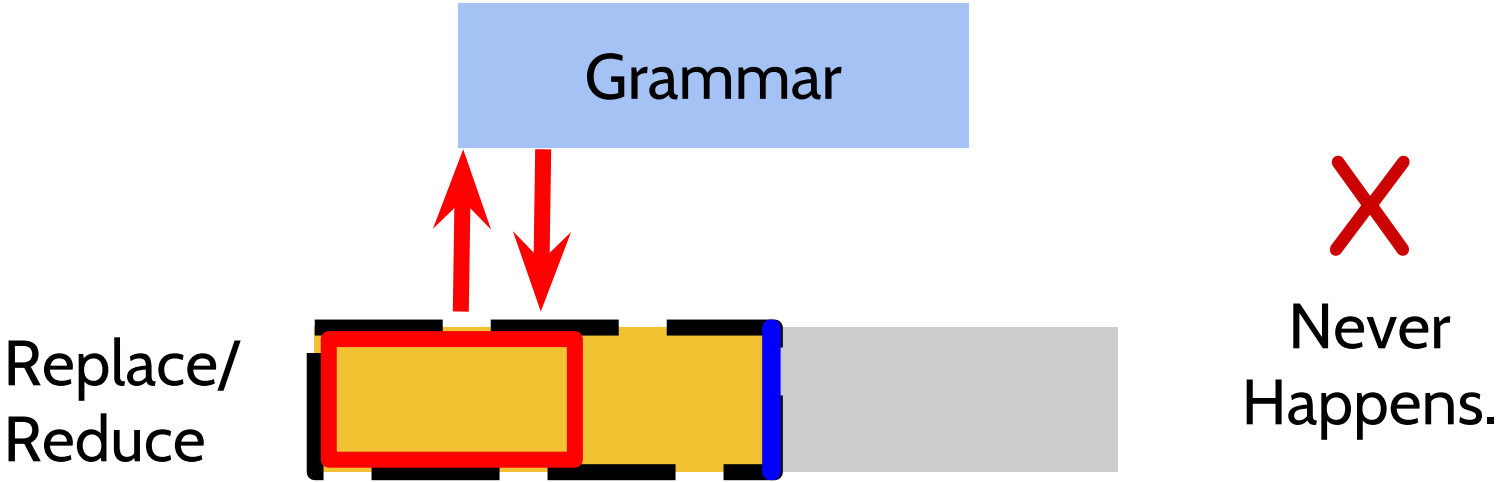
BUP in a nutshell



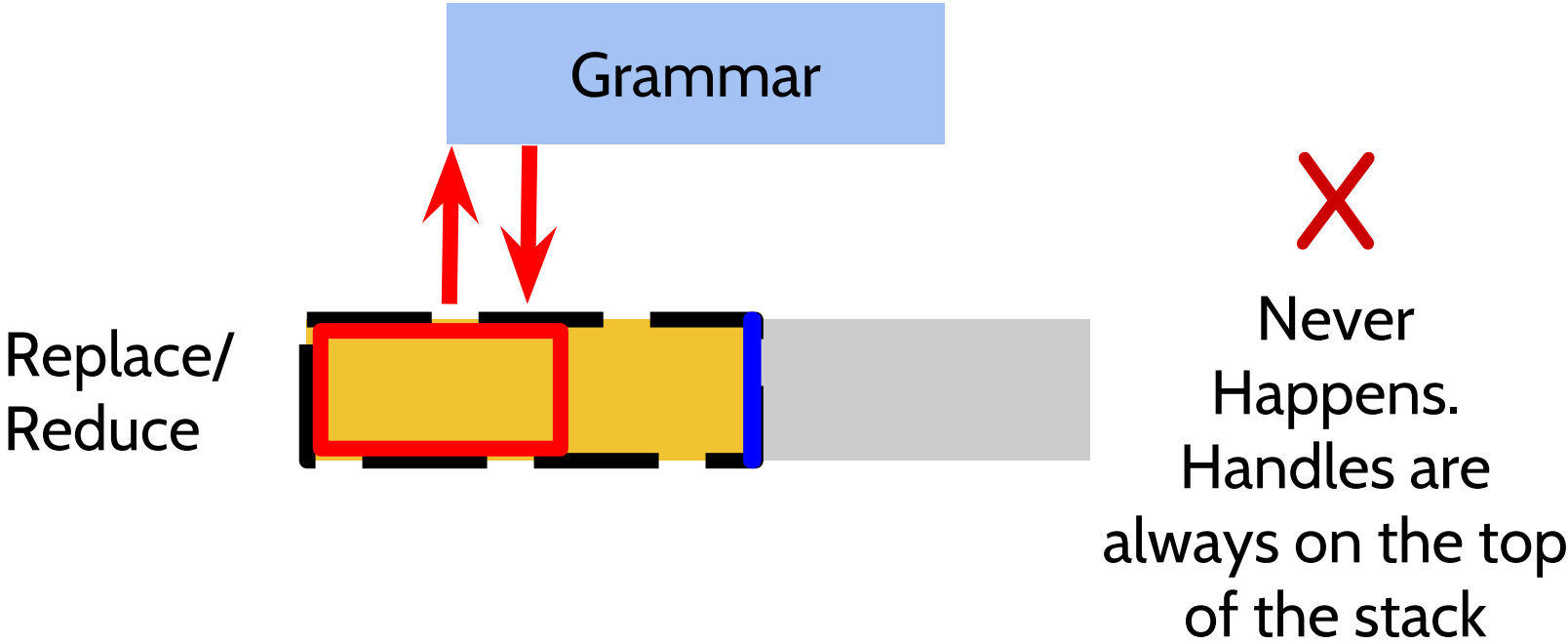
BUP in a nutshell



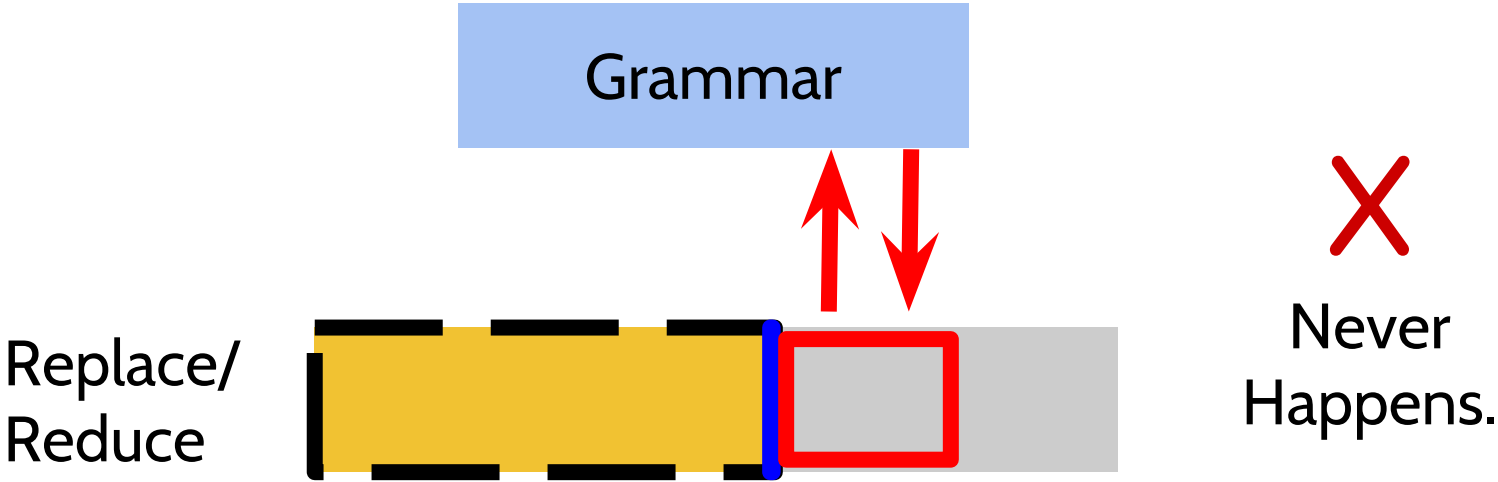
BUP in a nutshell



BUP in a nutshell



BUP in a nutshell



BUP
 comes down to
 knowing
 when to shift or
 reduce...

...and
exactly **what to**
replace, in your input
stack....

..which comes down
to recognizing which
part of your input
stack is a **handle**.

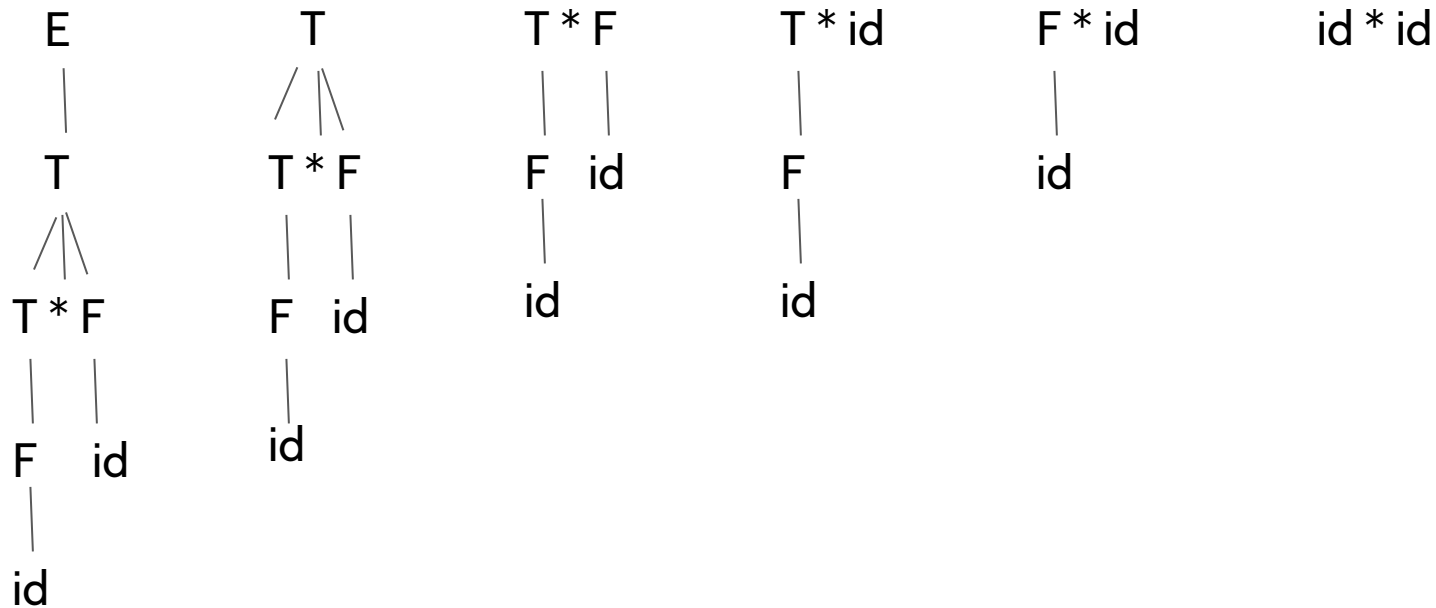
LR

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \leftarrow id * id$

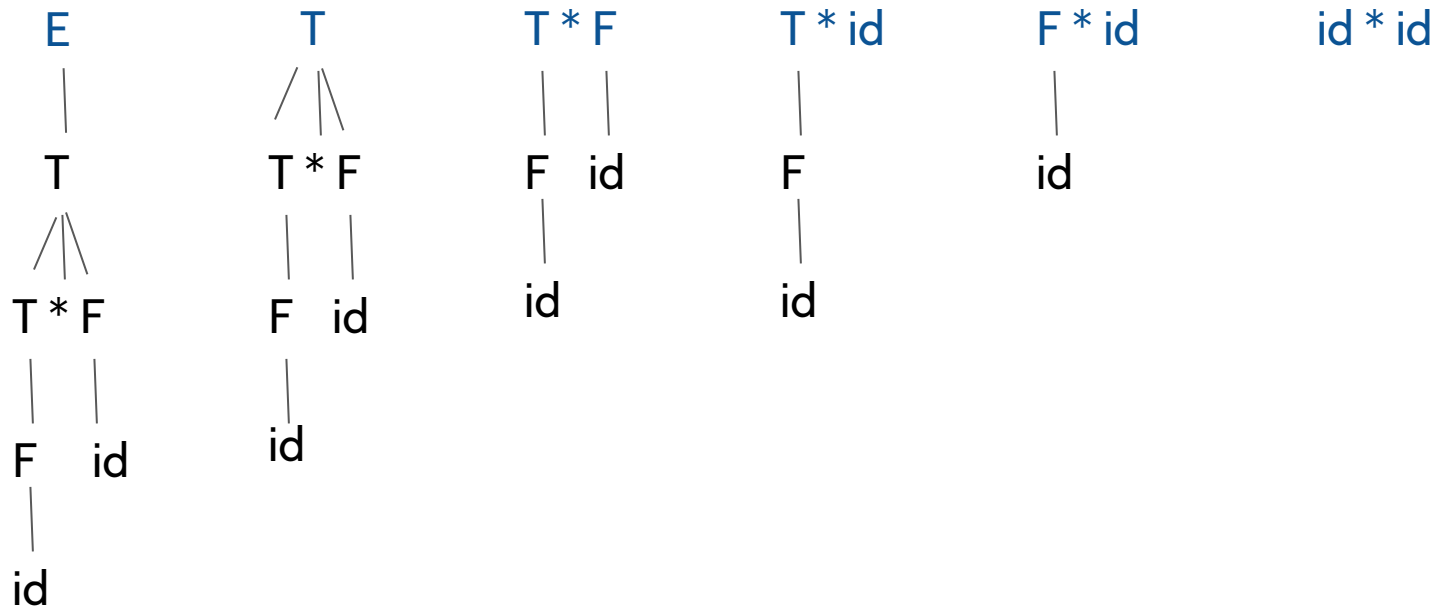


$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid \text{id}$

$E \leftarrow \text{id} * \text{id}$

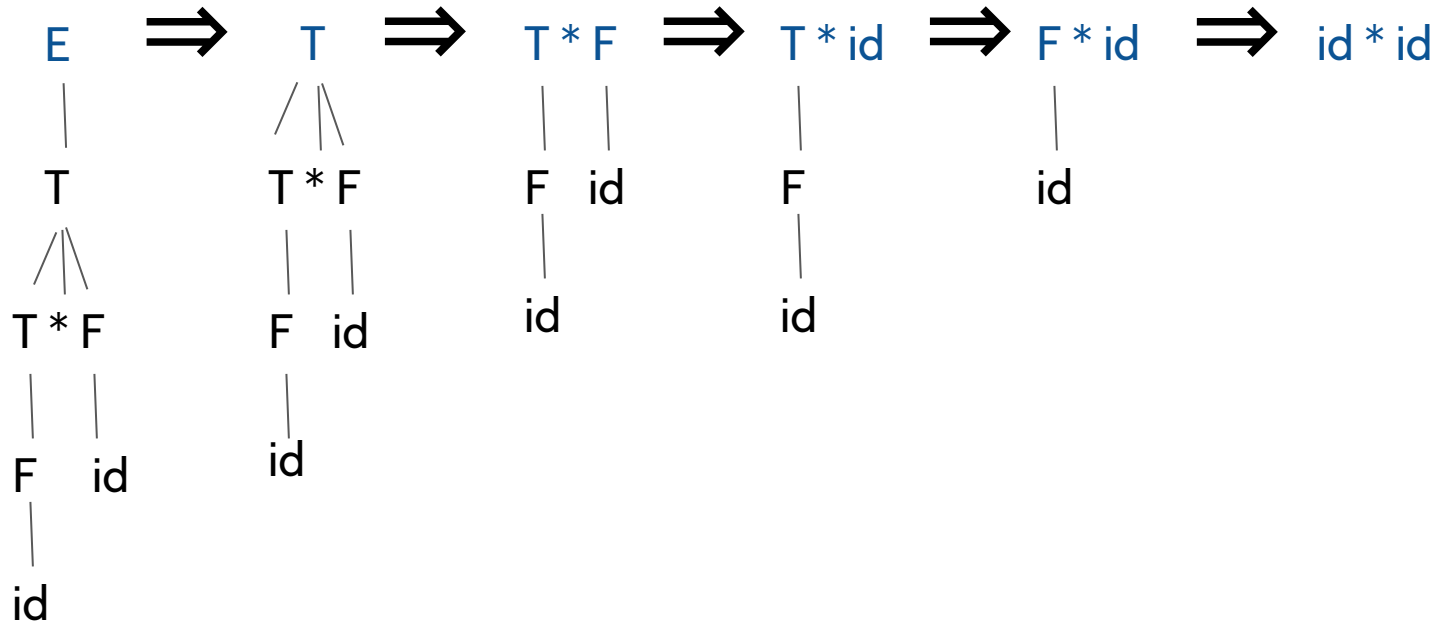


$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid \text{id}$

$E \xleftarrow{\hspace{15em}} \text{id} * \text{id}$



$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E ← id * id

$E \Rightarrow T \Rightarrow T * F \Rightarrow T * id \Rightarrow F * id \Rightarrow id * id$

$$\begin{array}{c} | \\ \mathbf{T} \\ \swarrow \downarrow \searrow \\ \mathbf{T} * \mathbf{F} \\ | \quad | \\ \mathbf{F} \quad \mathbf{id} \\ | \\ \mathbf{id} \end{array}$$
$$\begin{array}{c} \diagup \quad \diagdown \\ \text{T} * \text{F} \\ \text{F} \quad \text{id} \\ \text{id} \end{array}$$

F	id
id	

$$\begin{array}{c} | \\ \mathbf{F} \\ | \\ \mathbf{id} \end{array}$$

id

A rightmost derivation

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \xleftarrow{\hspace{10cm}} id * id$

$E \Rightarrow T \Rightarrow T * F \Rightarrow T * id \Rightarrow F * id \Rightarrow id * id$

T
/
 $T * F$
/
 $F \quad id$
/
 id

$T * F$
/
 $F \quad id$
/
 id

$F \quad id$
/
 id

F
/
 id

id

So BUP is a rightmost derivation in reverse.

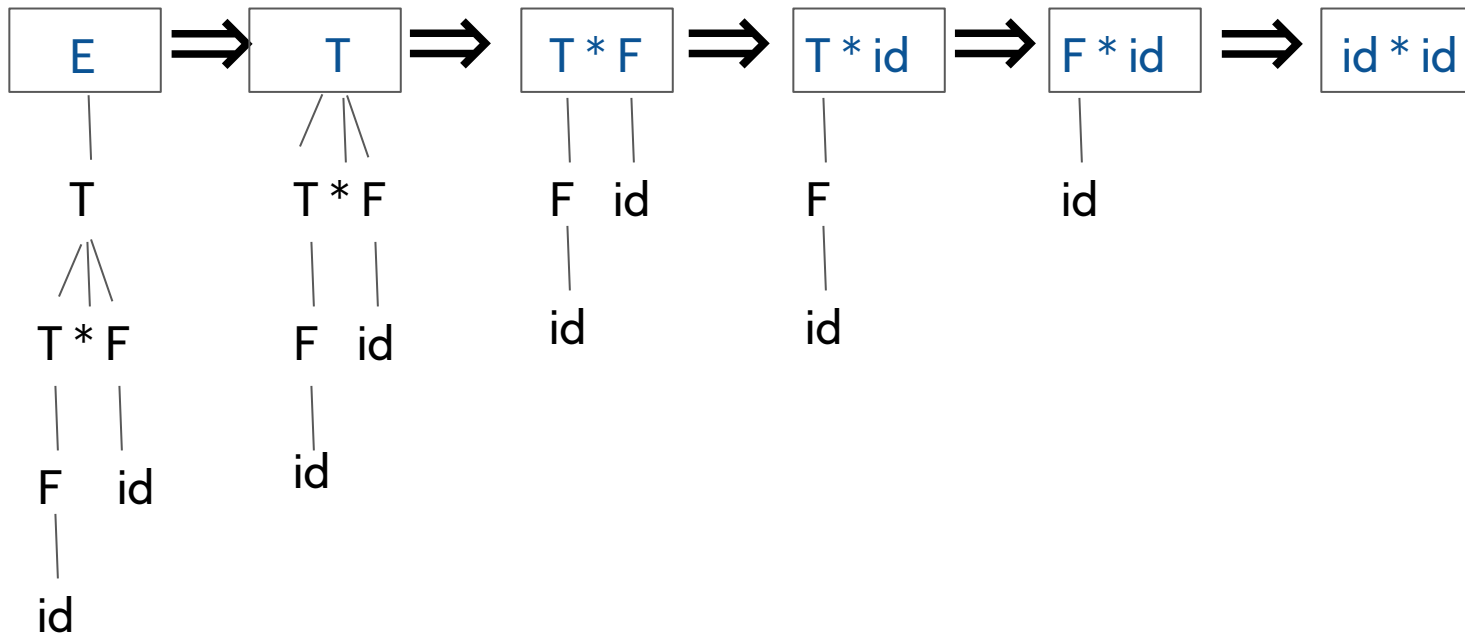
Another
Handle
Def

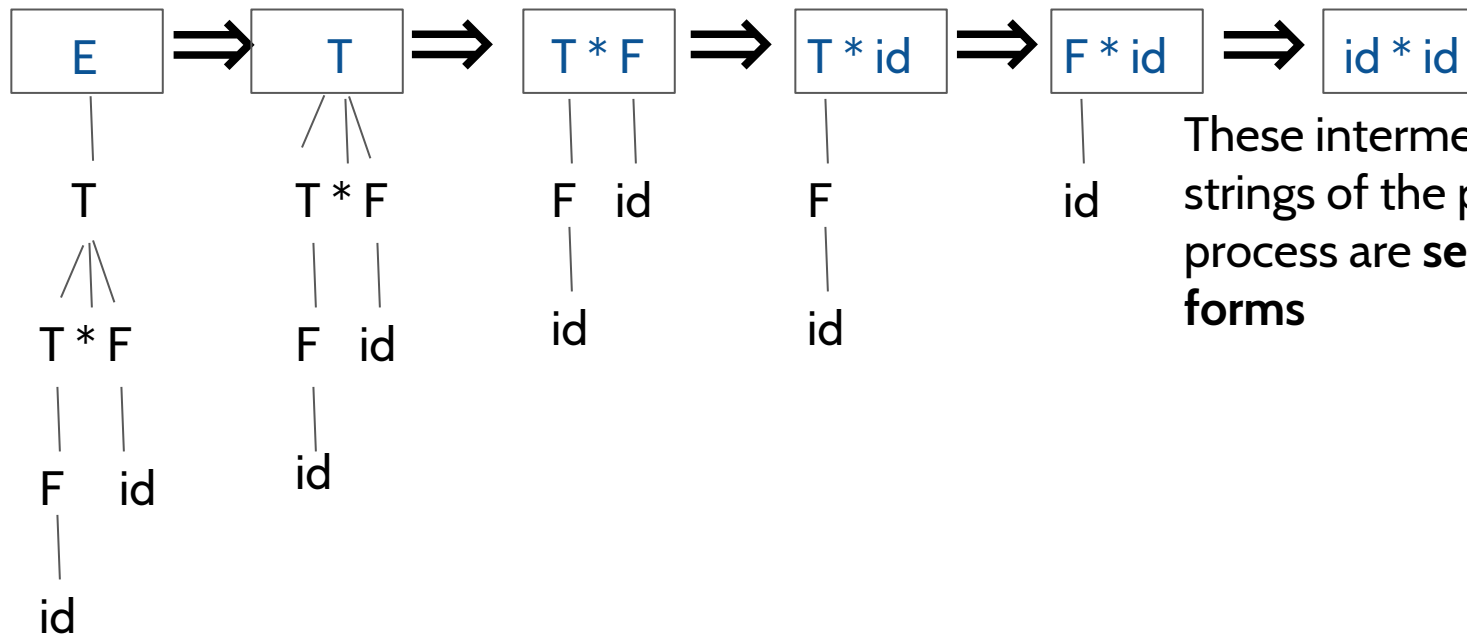
$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \xleftarrow{\hspace{15em}} id * id$



$$F \rightarrow (E) \mid id$$


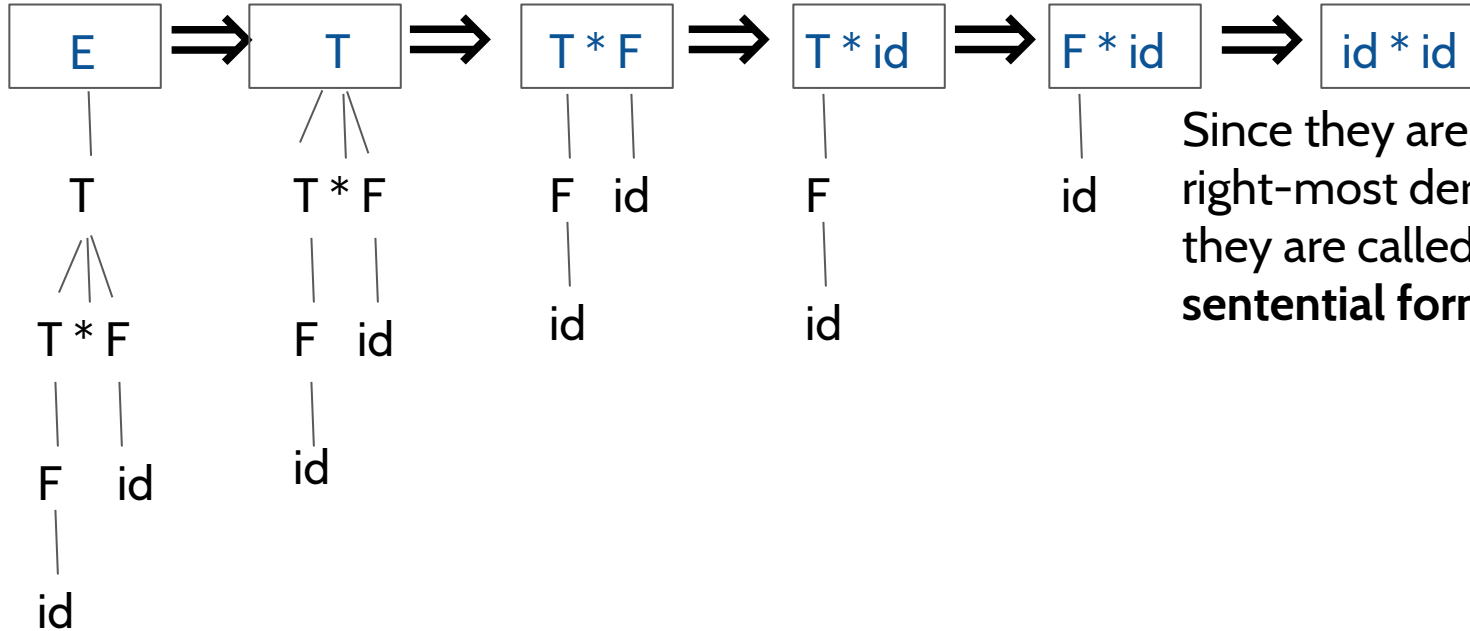
These intermediate strings of the parsing process are **sentential forms**

$$E \rightarrow E + T \mid T$$

$$T \rightarrow T * F \mid F$$

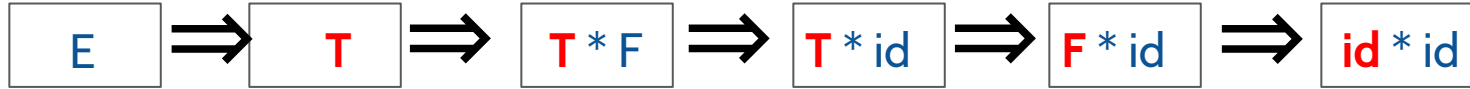
$$F \rightarrow (E) \mid id$$

$E \xleftarrow{\hspace{15em}} id * id$



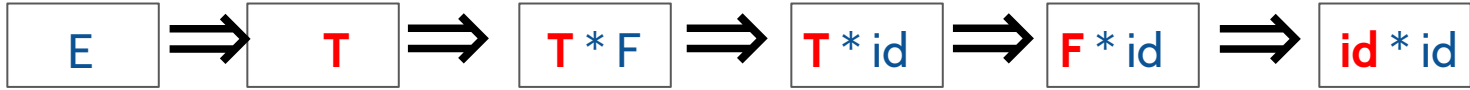
$E \rightarrow E + T \mid T$
 $T \rightarrow T * F \mid F$
 $F \rightarrow (E) \mid id$

$E \leftarrow id * id$



$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E

$$\text{id} * \text{id}$$


The handles

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E

$$\text{id} * \text{id}$$


F -> id or, simply, **id**
is a handle of **id * id**

$E \rightarrow E + T \mid T$
 $T \rightarrow T * F \mid F$
 $F \rightarrow (E) \mid id$

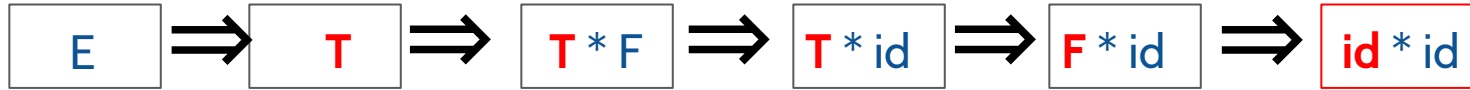
E ←———— id * id



F \rightarrow **id** or, simply, **id**
is a handle of **id * id**
Why?

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid \text{id}$$

$E \leftarrow \text{id} * \text{id}$



$F \rightarrow \text{id}$ or, simply, id
is a handle of $\text{id} * \text{id}$
Why?

Because replacing id at this position produces the previous right sentential form in a rightmost derivation of $\text{id} * \text{id}$

Yet
another
Handle
Def

$$\begin{aligned} E &\rightarrow E + T \mid T \\ T &\rightarrow T * F \mid F \\ F &\rightarrow (E) \mid \text{id} \end{aligned}$$

id * **id**

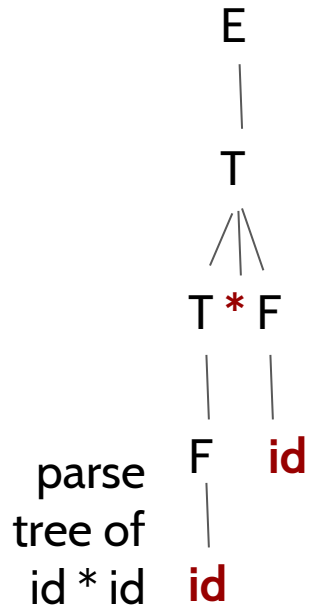
F \rightarrow **id** or, simply, **id**
is a handle of **id * id**

*“The handle of a parse tree T is the
leftmost complete cluster of leaf nodes.”*

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

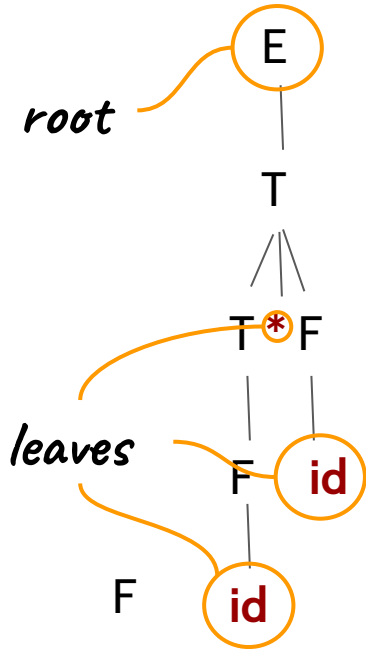
$F \rightarrow (E) \mid \text{id}$



id * id

$F \rightarrow \text{id}$ or, simply, **id**
is a handle of **id * id**

*“The handle of a parse tree T is the
leftmost complete cluster of leaf nodes.”*

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid \text{id}$$


id * **id**

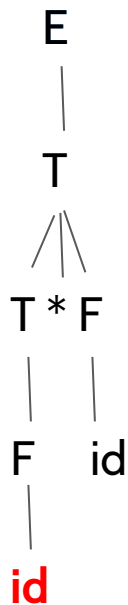
F \rightarrow **id** or, simply, **id**
is a handle of **id** * **id**

“The handle of a parse tree T is the leftmost complete cluster of leaf nodes.”

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid \text{id}$



id * id

F \rightarrow **id** or, simply, **id**
is a handle of **id** * **id**

*“The handle of a parse tree T is the
leftmost complete cluster of leaf nodes.”*

$E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow (E) \mid id$

$E \xleftarrow{\hspace{15em}} id * id$

$E \Rightarrow T \Rightarrow T * F \Rightarrow T * id \Rightarrow F * id \Rightarrow id * id$

$E \rightarrow E + T \mid T$
 $T \rightarrow T * F \mid F$
 $F \rightarrow (E) \mid \text{id}$

E ←———— id * id

E \Rightarrow T \Rightarrow T * F \Rightarrow T * id \Rightarrow F * id \Rightarrow id * id

So, how to use handles to do BUP?

$$E \rightarrow E + T \mid T$$
$$T \rightarrow T * F \mid F$$
$$F \rightarrow (E) \mid id$$

E ← id * id

$E \Rightarrow T \Rightarrow T * F \Rightarrow T * id \Rightarrow F * id \Rightarrow id * id$

Prune Handles