CS143: Parsing III

David L. Dill

Stanford University

Parsing

- LL(1) Parse Table Construction
- Computing FNE
- Computing FOLLOW
- LL(1) Odds and Ends
- Recursive Descent
- Bottom-up Parsing
- Shift-reduce Parsing

LL(1) Parse Table Construction

	n	+	*	(	)	$
E	E→TA			E→TA		
T	T→FB			T→FB		
F	F→n			F→(E)		
A		A→+TA			A→ε	A→ε
B		B→ε	B→*FB		B→ε	B→ε

Input

n+n*n$

stack

E$

TA$

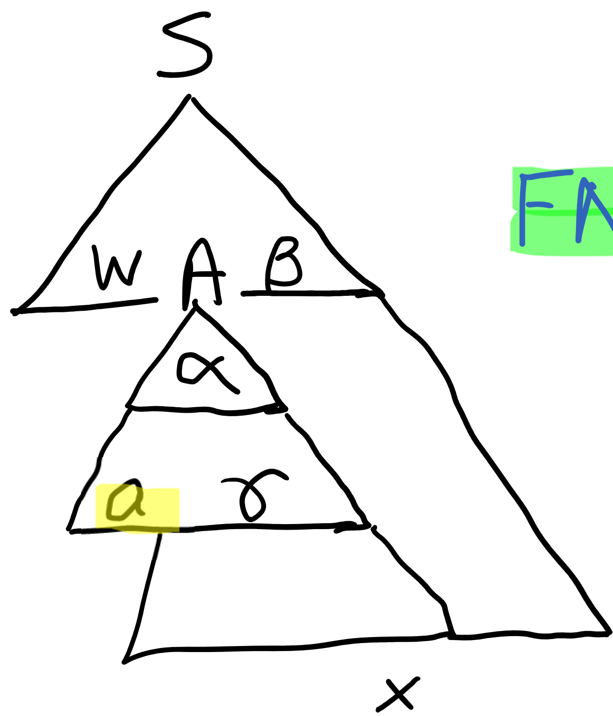
Table says what production to expand when TOS = A, nextin = a

How do we compute it?

$\text{TABLE}[A, a] = A \to \alpha$

$\text{Stack} = A\beta$ input $= wax$

next input



$\text{FNE}(\alpha) =$

$\{a \mid \alpha \overset{*}{\Longrightarrow} ax\}$

Later: how to compute it.

For every production $A \to \alpha$
and terminal symbol $a \in FNE(\alpha)$

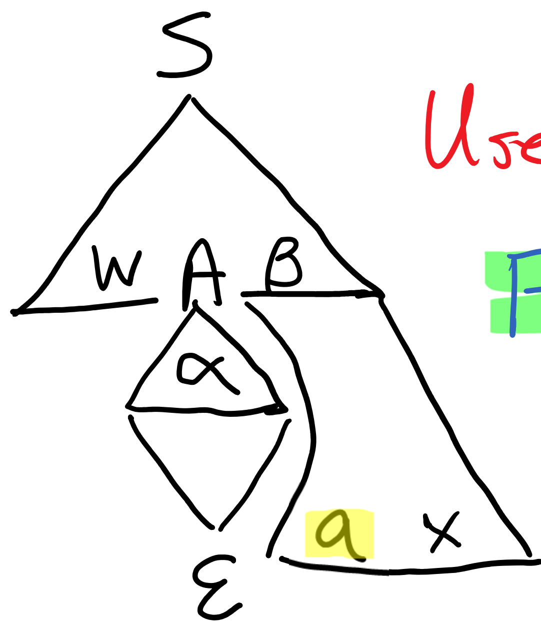
$$TABLE[A, a] = A \to \alpha$$

$$\text{TABLE}[A, a] = A \to \alpha$$

Stack $= A\beta$

input $= w\,a\,x$

↑ next input



Use when α is Nullable:

$$\text{FOLLOW}(A) = \{a \mid S\$ \overset{*}{\Rightarrow} wAax\}$$

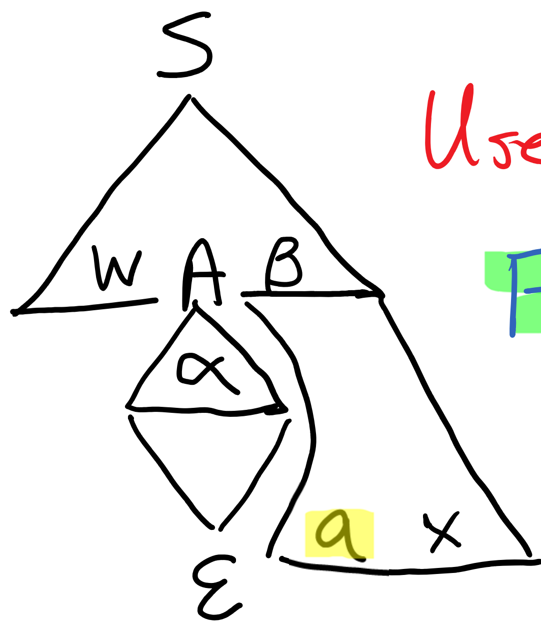
End-of-file

$\text{TABLE}[A, a] = A \to \alpha$

Stack $= A\beta$

input $= w\,a\,x$

↑ next input



Use when α is Nullable:

$\text{FOLLOW}(A)$

$= \{ a \mid S\$ \overset{*}{\Rightarrow} wAax \}$

"a" can be $

For every production $A \to \alpha$
where $\alpha \overset{*}{\Rightarrow} \varepsilon$
and terminal symbol $a \in \text{FOLLOW}(A)$

$$\text{TABLE}[A, a] = A \to \alpha$$

LL(1)

A CFG is LL(1) iff the above rules put at most one production in each table entry.

Computing FNE

Def: α is nullable if $\alpha \stackrel{*}{\Rightarrow} \varepsilon$ in CFG G

PREVIOUS LECTURE

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

A, B are the only nullable nonterminals

FNE – "FIRST, no epsilon"

$$FNE(\alpha) = \{a \mid a \in \Sigma \wedge \alpha \overset{*}{\Rightarrow} a\beta\}$$

Generic Approach to many Compiler Problems

Algorithm: Iterative application of rules

1. Start with initial value
2. Update only when rules require
3. Stop when no more updates required

FNE – "FIRST, no epsilon"

$FNE(\alpha) = \{a \mid a \in \Sigma \wedge \alpha \stackrel{*}{\Rightarrow} a\beta\}$

Iterative rules

$FNE(\alpha) = \emptyset$ initially

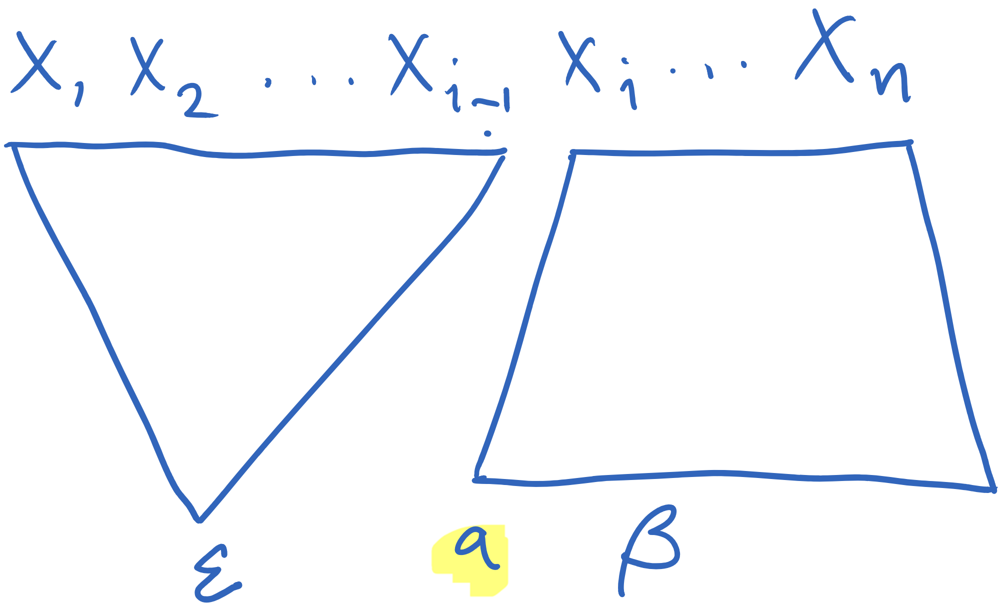
1. $FNE(a) = \{a\}$ for $a \in \Sigma$

2. $FNE(X_1 X_2 \ldots X_i \ldots X_n) \supseteq FNE(X_i)$ when $X_1 \ldots X_{i-1}$ is Nullable

add to $FNE(X_1 \ldots)$

2. $FNE(X_1 X_2 \ldots X_i \ldots X_n) \supseteq FNE(X_i)$

when $X_1 \ldots X_{i-1}$ is Nullable



$$X_1 X_2 \ldots X_{i-1} X_i \ldots X_n \overset{*}{\Rightarrow} a\beta$$

FNE – "FIRST, no epsilon"

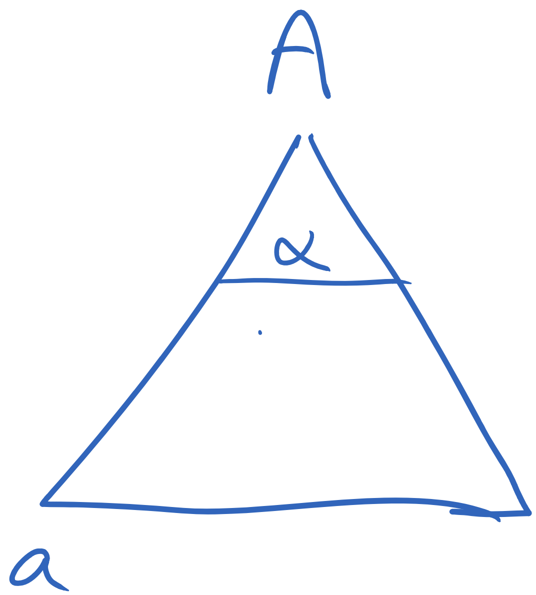
$$FNE(\alpha) = \{a \mid a \in \Sigma \wedge \alpha \overset{*}{\Rightarrow} a\beta\}$$

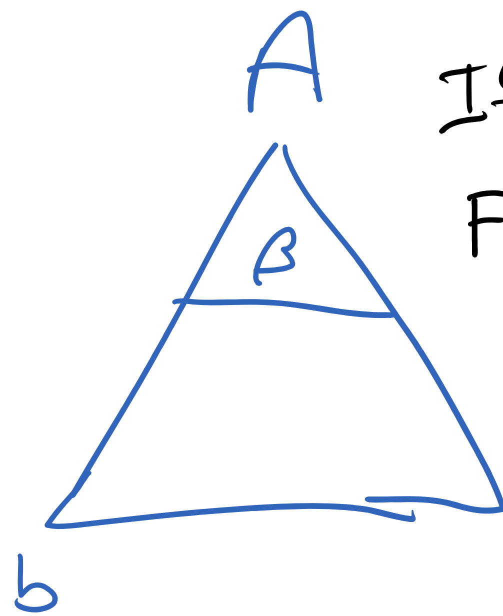
Iterative rules

$FNE(\alpha) = \emptyset$ initially

1. $FNE(a) = \{a\}$ for $a \in \Sigma$
2. $FNE(X_1 X_2 \ldots X_i \ldots X_n) \supseteq FNE(X_i)$ when $X_1 \ldots X_{i-1}$ is Nullable
3. $FNE(A) \supseteq FNE(\alpha)$ when $A \to \alpha$

3. $FNE(A) \supseteq FNE(\alpha)$ when $A \to \alpha$





Is $A \to \beta$,
$FNE(A) \supseteq FNE(\beta)$
also

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FNE

$+TA \quad \{+\}$

$*FB \quad \{*\}$

$(E) \quad \{(\}$

$n \quad \{n\}$

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FNE

$+TA$	$\{+\}$
$*FB$	$\{*\}$
(E)	$\{(\}$
n	$\{n\}$
F	$\{(, n\}$

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FNE	
$+TA$	$\{+\}$
$*FB$	$\{*\}$
(E)	$\{(\}$
n	$\{n\}$
F	$\{(, n\}$
B	$\{*\}$

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FNE

$+TA$	$\{+\}$
$*FB$	$\{*\}$
(E)	$\{(\}$
n	$\{n\}$
F	$\{(, n\}$
B	$\{*\}$
FB	$\{(, n\}$

$E \to TA$
$A \to +TA$
$A \to \varepsilon$
$T \to FB$
$B \to *FB$
$B \to \varepsilon$
$F \to (E)$
$F \to n$

FNE

$+TA$	$\{+\}$	T	$\{(, n\}$
$*FB$	$\{*\}$		
(E)	$\{(\}$		
n	$\{n\}$		
F	$\{(, n\}$		
B	$\{*\}$		
FB	$\{(, n\}$		

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FNE

$+TA$	$\{+\}$	T	$\{(, n\}$
$*FB$	$\{*\}$	A	$\{+\}$
(E)	$\{(\}$		
n	$\{n\}$		
F	$\{(, n\}$		
B	$\{*\}$		
FB	$\{(, n\}$		

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FNE

$+TA$	$\{+\}$	T	$\{(, n\}$
$*FB$	$\{*\}$	A	$\{+\}$
(E)	$\{(\}$	TA	$\{(, n\}$
n	$\{n\}$		
F	$\{(, n\}$		
B	$\{*\}$		
FB	$\{(, n\}$		

$E \rightarrow TA$

$A \rightarrow +TA$

$A \rightarrow \varepsilon$

$T \rightarrow FB$

$B \rightarrow *FB$

$B \rightarrow \varepsilon$

$F \rightarrow (E)$

$F \rightarrow n$

FNE

$+TA$	$\{+\}$	T	$\{(, n\}$
$*FB$	$\{*\}$	A	$\{+\}$
(E)	$\{(\}$	TA	$\{(, n\}$
n	$\{n\}$	E	$\{(, n\}$
F	$\{(, n\}$		
B	$\{*\}$		
FB	$\{(, n\}$		

	n	+	*	(	)	$
E	$E \to TA$			$E \to TA$		
T	$T \to FB$			$T \to FB$		
F	$F \to n$			$F \to (E)$		
A		$A \to +TA$			$A \to \varepsilon$	$A \to \varepsilon$
B		$B \to \varepsilon$	$B \to *FB$		$B \to \varepsilon$	$B \to \varepsilon$

$TA \quad \{ (, n \}$

For every production $A \to \alpha$ and terminal symbol $a \in FNE(\alpha)$

$$\text{TABLE}[A, a] = A \to \alpha$$

	n	+	*	(	)	$
E	$E \to TA$			$E \to TA$		
T	$T \to FB$			$T \to FB$		
F	$F \to n$			$F \to (E)$		
A		$A \to +TA$			$A \to \varepsilon$	$A \to \varepsilon$
B		$B \to \varepsilon$	$B \to *FB$		$B \to \varepsilon$	$B \to \varepsilon$

TA $\{ (, n \}$

FB $\{ (, n \}$

For every production $A \to \alpha$
and terminal symbol $a \in FNE(\alpha)$

$$TABLE[A, a] = A \to \alpha$$

	n	+	*	(	)	$
E	E → TA			E → TA		
T	T → FB			T → FB		
F	F → n			F → (E)		
A		A → +TA			A → ε	A → ε
B		B → ε	B → *FB		B → ε	B → ε

TA { (, n }

FB { (, n }

For every production $A \to \alpha$
and terminal symbol $a \in FNE(\alpha)$

$$\text{TABLE}[A, a] = A \to \alpha$$

Computing Follow

$$\mathrm{FOLLOW}(A) = \{ a \mid S\$ \overset{*}{\Rightarrow} \alpha A a \beta \}$$

Only needs to be defined for single non terminals.

$$\text{FOLLOW}(A) = \{a \mid S\$ \stackrel{*}{\Rightarrow} \alpha A a \beta\}$$

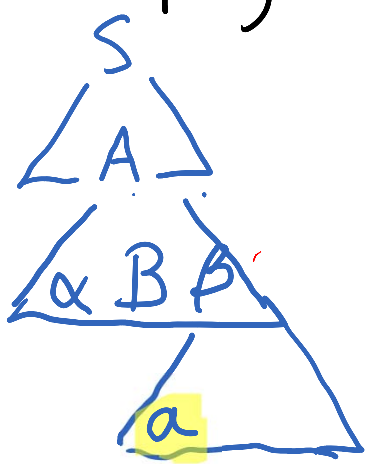
If A is last symbol, then

$a = \$$

$$\text{FOLLOW}(A) = \{a \mid S\$ \xRightarrow{*} \alpha A a \beta\}$$

Init: $\text{FOLLOW}(A) = \emptyset$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$



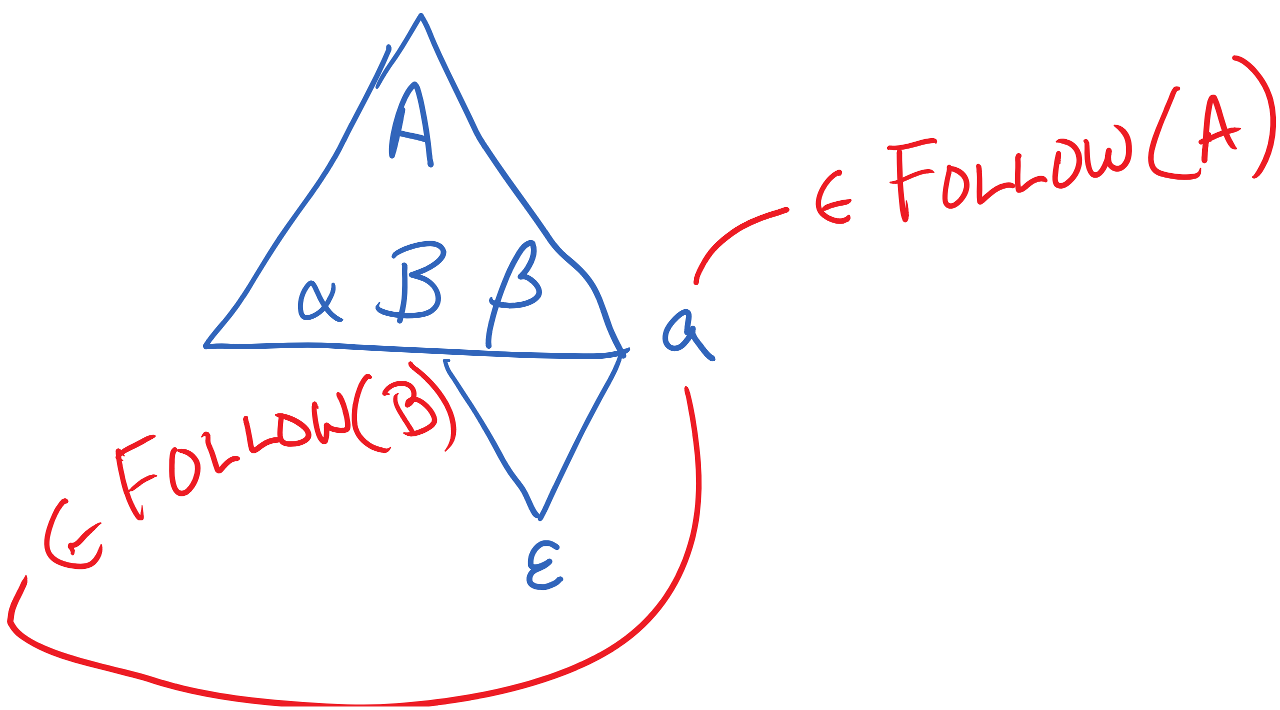
$$\text{FOLLOW}(A) = \{a \mid S\$ \xRightarrow{*} \alpha A a \beta\}$$

Init: $\text{FOLLOW}(A) = \emptyset$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \xRightarrow{*} \varepsilon$, $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

If $A \to \alpha B \beta$ and $\beta \xRightarrow{*} \varepsilon$,

$\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$



$E \to T A$

$A \to + T A$

$A \to \varepsilon$

$T \to F B$

$B \to * F B$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \overset{*}{\Rightarrow} \varepsilon$, $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	$\{+\}$
B	$\{*\}$

FOLLOW	
E	$\{$
T	$\{$
F	$\{$
A	$\{$
B	$\{$

$E \to T A$

$A \to + T A$

$A \to \varepsilon$

$T \to F B$

$B \to * F B$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \stackrel{*}{\Longrightarrow} \varepsilon$,
   $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	$\{ + \}$
B	$\{ * \}$

FOLLOW	
E	$\{ \$$
T	$\{$
F	$\{$
A	$\{$
B	$\{$

$E \to T\,A$

$A \to +\,T\,A$

$A \to \varepsilon$

$T \to F\,B$

$B \to *\,F\,B$

$B \to \varepsilon$

$F \to (\,E\,)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \xRightarrow{*} \varepsilon$,
   $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	$\{+\}$
B	$\{*\}$

FOLLOW	
E	$\{\$,)$
T	$\{$
F	$\{$
A	$\{$
B	$\{$

$E \to T A$

$A \to + T A$

$A \to \varepsilon$

$T \to F B$

$B \to * F B$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \overset{*}{\Rightarrow} \varepsilon$, $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	{ + }
B	{ * }

FOLLOW	
E	{ $,)
T	{ +
F	{
A	{
B	{

$E \to T\,A$

$A \to +\,T\,A$

$A \to \varepsilon$

$T \to F\,B$

$B \to *\,F\,B$

$B \to \varepsilon$

$F \to (\,E\,)$

$F \to n$

1. $\$ \in$ FOLLOW(S)
2. If $A \to \alpha B \beta$, FNE(β) $\subseteq$ FOLLOW(B)
3. If $A \to \alpha B \beta$ and $\beta \stackrel{*}{\Rightarrow} \varepsilon$, FOLLOW(B) $\supseteq$ FOLLOW(A)

	FNE
A	{ + }
B	{ * }

	FOLLOW
E	{ \$,)
T	{ +, \$,)
F	{
A	{
B	{

$E \to T A$

$A \to + T A$

$A \to \varepsilon$

$T \to F B$

$B \to * F B$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

1. $\$ \in$ FOLLOW(S)
2. If $A \to \alpha B \beta$, FNE$(\beta) \subseteq$ FOLLOW(B)
3. If $A \to \alpha B \beta$ and $\beta \overset{*}{\Rightarrow} \varepsilon$, FOLLOW$(B) \supseteq$ FOLLOW(A)

FNE	
A	{ + }
B	{ * }

FOLLOW	
E	{ \$,)
T	{ +, \$,)
F	{ *
A	{
B	{

$E \to T\,A$

$A \to +\,T\,A$

$A \to \varepsilon$

$T \to F\,B$

$B \to *\,F\,B$

$B \to \varepsilon$

$F \to (\,E\,)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \overset{*}{\Rightarrow} \varepsilon$, $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	{ + }
B	{ * }

FOLLOW	
E	{ $,)
T	{ +, $,)
F	{ *, +, $,)
A	{
B	{

$E \to T A$

$A \to + T A$

$A \to \varepsilon$

$T \to F B$

$B \to * F B$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \stackrel{*}{\Rightarrow} \varepsilon$, $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE		FOLLOW	
A	$\{+\}$	E	$\{\$,)$
B	$\{*\}$	T	$\{+, \$,)$
		F	$\{*, +, \$,)$
		A	$\{\$,)$
		B	$\{$

$E \to T\,A$

$A \to +\,T\,A$

$A \to \varepsilon$

$T \to F\,B$

$B \to *\,F\,B$

$B \to \varepsilon$

$F \to (\,E\,)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \overset{*}{\Longrightarrow} \varepsilon$,
   $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	$\{+\}$
B	$\{*\}$

FOLLOW	
E	$\{\$,)\}$
T	$\{+, \$,)\}$
F	$\{*, +, \$,)\}$
A	$\{\$,)\}$
B	$\{+, \$,)\}$

$E \to T\,A$

$A \to +\,T\,A$

$A \to \varepsilon$

$T \to F\,B$

$B \to *\,F\,B$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

1. $\$ \in \text{FOLLOW}(S)$
2. If $A \to \alpha B \beta$, $\text{FNE}(\beta) \subseteq \text{FOLLOW}(B)$
3. If $A \to \alpha B \beta$ and $\beta \overset{*}{\Longrightarrow} \varepsilon$, $\text{FOLLOW}(B) \supseteq \text{FOLLOW}(A)$

FNE	
A	$\{+\}$
B	$\{*\}$

FOLLOW	
E	$\{\$,)\}$
T	$\{+, \$,)\}$
F	$\{*, +, \$,)\}$
A	$\{\$,)\}$
B	$\{+, \$,)\}$

	n	+	*	(	)	$
E	$E \to TA$			$E \to TA$		
T	$T \to FB$			$T \to FB$		
F	$F \to n$			$F \to (E)$		
A		$A \to +TA$			$A \to \varepsilon$	$A \to \varepsilon$
B		$B \to \varepsilon$	$B \to *FB$		$B \to \varepsilon$	$B \to \varepsilon$

A $\{ \$,) \}$

B $\{ +, \$,) \}$

For every production $A \to \alpha$
where $\alpha \overset{*}{\Longrightarrow} \varepsilon$
and terminal symbol $a \in \text{FOLLOW}(A)$

$\text{TABLE}[A, a] = A \to \alpha$

	n	+	*	(	)	$
E	$E \to TA$			$E \to TA$		
T	$T \to FB$			$T \to FB$		
F	$F \to n$			$F \to (E)$		
A		$A \to +TA$			$A \to \varepsilon$	$A \to \varepsilon$
B		$B \to \varepsilon$	$B \to *FB$		$B \to \varepsilon$	$B \to \varepsilon$

A { $,) }

B { +, $,) }

For every production $A \to \alpha$
where $\alpha \overset{*}{\Rightarrow} \varepsilon$
and terminal symbol $a \in$ FOLLOW(A)

TABLE[A, a] = $A \to \alpha$

Announcements

- PA1 due
- WA1 due – but extended with no penalty until Monday
- PA2 assigned
- WA2 assigned
- Notes on parsing posted.

LL(1) summary

Preprocessing:

- Remove left recursion
- Left factor

Parser generation

- Compute Nullable
- Compute FNE
- Compute Follow
- Build table

LL(1) Odds and Ends

FIRST Sets

FNE is nonstandard

Standard: FIRST sets

$$\text{FIRST}(\alpha) = \begin{cases} \text{FNE}(\alpha) \cup \{\varepsilon\} & \text{if } \alpha \overset{*}{\Rightarrow} \varepsilon \\ \text{FNE}(\alpha) & \text{if } \alpha \overset{*}{\not\Rightarrow} \varepsilon \end{cases}$$

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

FIRST

$+TA$	$\{+\}$	T	$\{(, n\}$
$*FB$	$\{*\}$	A	$\{+, \varepsilon\}$
(E)	$\{(\}$	TA	$\{(, n\}$
n	$\{n\}$	E	$\{(, n\}$
F	$\{(, n\}$		
B	$\{*, \varepsilon\}$		
FB	$\{(, n\}$		

Ambiguity and LL(1)

Thm No ambiguous CFG is LL(1)

proof idea: There are two distinct leftmost derivations for the same terminal string.

Ambiguity and LL(1)

Thm No ambiguous CFG is LL(1)

proof idea: There are two distinct leftmost derivations for the same terminal string. The derivations contain two different productions, $A \to \alpha$ and $A \to \beta$, that go into the same table entry.

Short answer questions.

A CFG with no useless productions is <u>not</u> LL(1) if:

1. It is left recursive
2. It has left factors
3. It is ambiguous.

Recursive Descent Parsing

Recursive Descent

Manual LL(1)-ish parsing

Why?

Custom heuristics for production choice (e.g. extra lookahead) when CFG is "not quite" LL(1).

Custom error reporting & recovery (+ more work and bugs)

Recursive Descent

Same preconditioning

Same Nullable, FNE, Follow

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

```
parse_E():
    t1 = parse_T();
    t2 = parse_A();
    return makeTree('E', t1, t2)
```

$E \to TA$

$A \to +TA$

$A \to \varepsilon$

$T \to FB$

$B \to *FB$

$B \to \varepsilon$

$F \to (E)$

$F \to n$

```
parse_A():
    tok = getlex();
    if (tok == '+')
    then  -- A → +TA
        t1 = parse_T()
        t2 = parse_A()
        return makeTree('A', '+', t1, t2)
    else
        -- A → ε
        return makeTree('A')
etc.
```

Bottom-up Parsing

$E \to E + T$

$E \to T$

$T \to T * F$

$T \to F$

$F \to (E)$

$F \to n$

$n + n * n$

$$E \rightarrow E + T$$

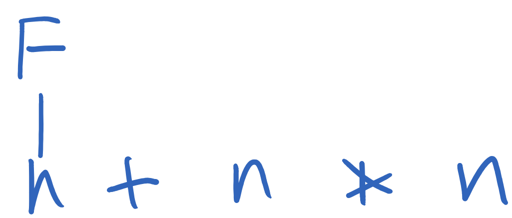
$$E \rightarrow T$$

$$T \rightarrow T * F$$

$$T \rightarrow F$$

$$F \rightarrow (E)$$

$$F \rightarrow n$$



$E \to E + T$

$E \to T$

$T \to T * F$

$T \to F$

$F \to (E)$

$F \to n$

```
T
|
F
|
n + n * n
```

$E \to E + T$

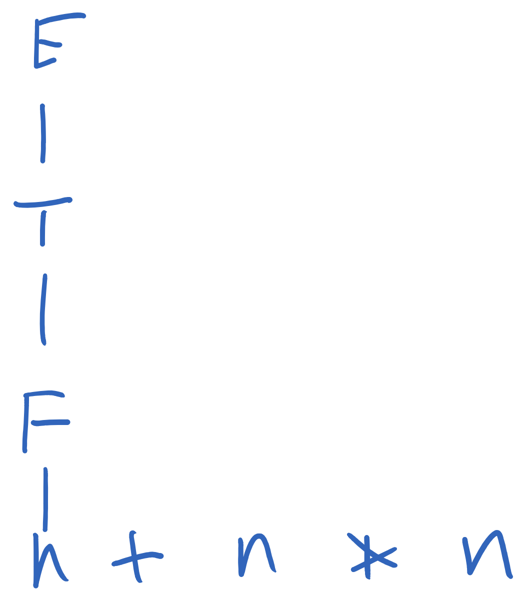
$E \to T$

$T \to T * F$

$T \to F$

$F \to (E)$

$F \to n$



$E \to E + T$

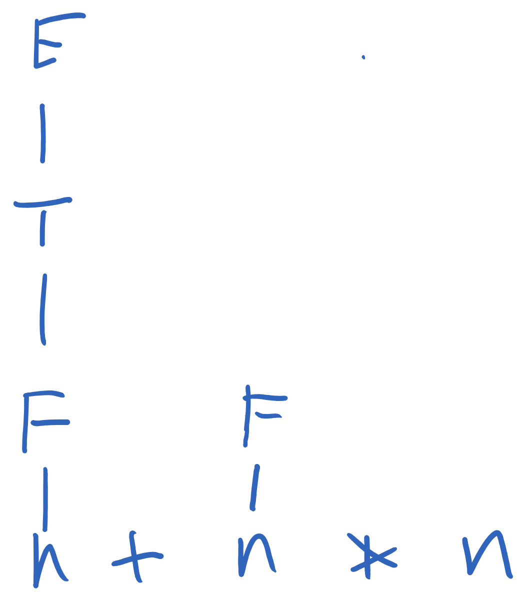
$E \to T$

$T \to T * F$

$T \to F$

$F \to (E)$

$F \to n$



$E \rightarrow E + T$

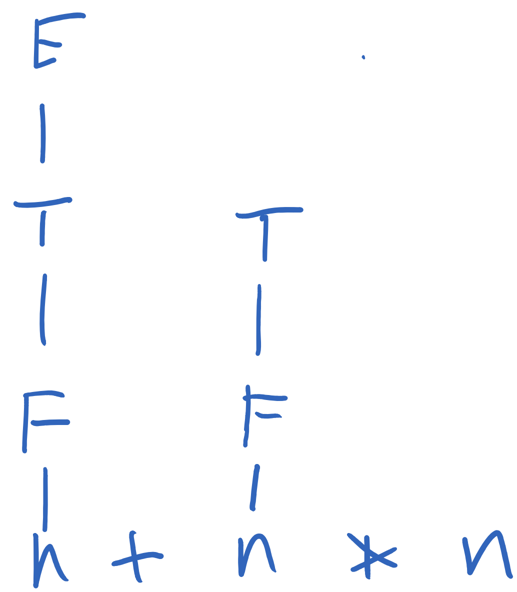
$E \rightarrow T$

$T \rightarrow T * F$

$T \rightarrow F$

$F \rightarrow (E)$

$F \rightarrow n$



$$E \rightarrow E + T$$

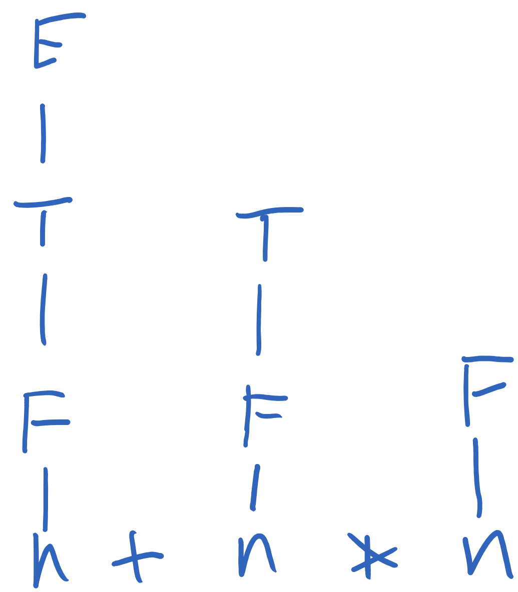
$$E \rightarrow T$$

$$T \rightarrow T * F$$

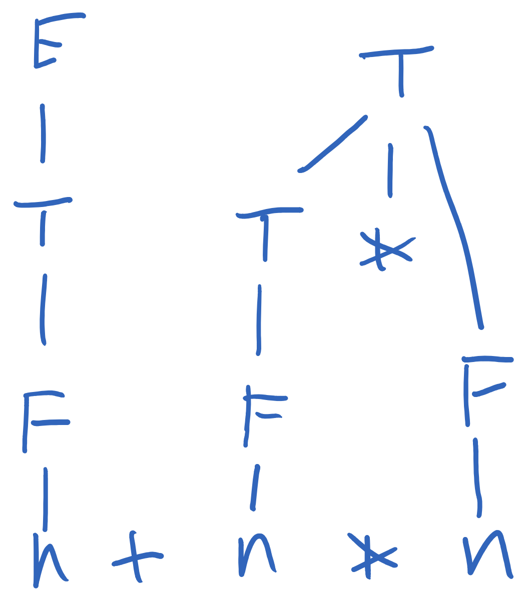
$$T \rightarrow F$$

$$F \rightarrow (E)$$

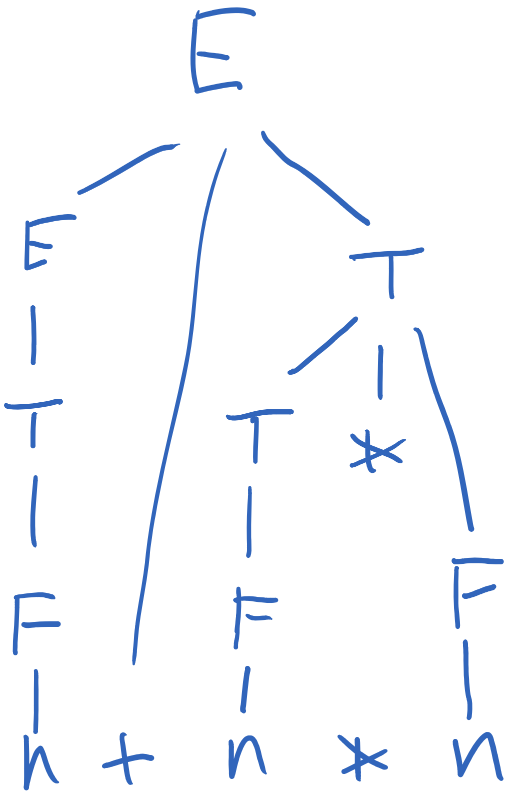
$$F \rightarrow n$$



$$E \to E + T$$
$$E \to T$$
$$T \to T * F$$
$$T \to F$$
$$F \to (E)$$
$$F \to n$$



$$E \to E + T$$
$$E \to T$$
$$T \to T * F$$
$$T \to F$$
$$F \to (E)$$
$$F \to n$$



Shift-Reduce Parsing

$S \to (S) \mid a$

Stack (top)	input	action
\$	((a))\$	shift
\$(	(a))\$	shift
\$((	a))\$	shift
\$((a	))\$	reduce $S \to a$
\$((S	))\$	shift
\$((S)	)\$	reduce $S \to (S)$
\$(S	)\$	shift
\$(S)	\$	reduce $S \to (S)$
\$S	\$	accept