First and Follow Sets

Grammar

E → T X
X → + E
X → ε
T → int Y
T → (E)
Y → * T
Y → ε

First Set

First(a) = set of terminals that start string of terminals derived from a.

Apply following rules until no terminal or ε can be added

1. If t ∈ T, then **First**(t) = { t }.

2. If X ∈ N and X → ε exists (nullable), then add ε to **First**(X).

3. If X ∈ N and X → $Y_1Y_2Y_3 \ldots Y_m$, where Y_1, Y_2, Y_3, ... Y_m are non-terminals, then:

   for each *i* from 1 to *m*

   if $Y_1 \ldots Y_{i-1}$ are all nullable (or if *i* = 1)

   First(X) = **First**(X) ∪ **First**(Y_i)

By rule 1:

The **First** of a terminal is that terminal.

Symbol	First
(	(
)	)
+	+
*****	*
int	int
Y	
X	
T	
E	

By rule 2:

If X ∈ N and X → ε exists (nullable), then add ε to **First**(X).

X → ε
Y → ε

Symbol	First
(	(
)	)
+	+
*	*
int	int
Y	ε
X	ε
T	
E	

By rule 3:

If X ∈ N and X → $Y_1Y_2Y_3 \ldots Y_m$, where Y_1, Y_2, Y_3, ... Y_m are non-terminals, then:

for each *i* from 1 to *m*

if $Y_1 \ldots Y_{i-1}$ are all nullable (or if *i* = 1)

First(X) = **First**(X) ∪ **First**(Y_i)

E → T X
First(E) = **First**(E) ∪ **First**(T)

We need First(T)… first.

First(T) = {int, (}

T → int Y because First(int) = int
T → (E) because First(( ) = (

Now we can go back and do **First**(E)

First(E) = { int, (}

We don't consider the first of X because T is not nullable.

First(X) = { ε, + }

X → + E First(+) = +
X → ε handled at step 2

First(Y) = { ε, * }

Y → * T   First(*) = *
Y → ε   handled at step 2

Put it all together:

Symbol	First
(	(
)	)
+	+
*****	*
int	int
Y	ε, *
X	ε, +
T	int, (
E	int, (

Done.

Follow Set

Apply following rules until no terminal or e can be added

1. $ ∈ **Follow**(S), where S is the start symbol.

2. Look at the occurrence of a non-terminal on the right hand side of a production which is followed by something
   If A → a B b, then **First**(b) - {e} ⊆ **Follow**(B)

3. Look at N on the RHS that is not followed by anything,
   if (A → a B) or (A → a B b and ε ∈ **First**(b)),
   then **Follow**(A) ⊆ **Follow**(B)

By step 1:

$ ∈ **Follow**(S), where S is the start symbol.

a.) **Follow**(E) = { $ }   E is our start symbol

The table so far:

Symbol	First	Follow
(	(	N/A
)	)	
+	+	
*****	*	
int	int	
Y	ε, *	
X	ε, +	
T	int, (	
E	int, (	$

By step 2:

Look at the occurrence of a non-terminal on the right hand side of a production which is followed by something:

If A → a B b, then **First**(b) - { ε } ⊆ **Follow**(B)

This subset notation can be confusing. It's saying here that the Follow(B) contains at least the First(b) (with ε excluded) as a subset. There might be more things in Follow(B) because other steps added them in or other grammar rules were analyzed.

E → T X

a.) **Follow**(T) contains (at least) the First(X) = { ε, + } – { ε } = { + }

T → (E)

b.) **Follow**(E) contains (at least) the First( )) = {) }

So now **Follow**(E) is {), $ } (From step 1a)

The table so far:

Symbol	First	Follow
(	(	N/A
)	)	
+	+	
*****	*	
int	int	
Y	ε, *	
X	ε, +	
T	int, (	+
E	int, (	), $

By step 3:

Look at N on the RHS that is not followed by anything,

if (A → a B) or (A → a B b and ε ∈ **First**(b)),
then **Follow**(A) ⊆ **Follow**(B)

Again, be careful with the notation. It's saying here that the Follow(B) contains at least the Follow(A) as a subset. There might be more things in Follow(B) because other steps added them in or other grammar rules were analyzed.

E → T **X**
a.) **Follow**(X) contains (at least) Follow(E) = {), $ }   (from step 2b)

ε ∈ **First**(X) so:

b.) **Follow**(T) contains (at least) Follow(E) = {), $, +}   (from step 2a and 2b)

X → + **E**
c.) **Follow**(E) contains (at least) Follow(X) = {), $ }   (from step 3a)

T → int **Y**
d.) **Follow**(Y) contains (at least) Follow(T) = {), $, +}   (from step 3b)

Y → * **T**
e.) **Follow**(T) contains (at least) Follow(Y) = {), $ }   (from step 3d)

We do this whole process again until no more additions happen:

E → T **X**
f.) **Follow**(X) contains (at least) Follow(E) = {), $ }   (no change)

ε ∈ **First**(X) so:

g.) **Follow**(T) contains (at least) Follow(E) = {), $, +}   (no change)

X → + **E**
h.) **Follow**(E) contains (at least) Follow(X) = {), $ }   (no change)

T → int **Y**
i.) **Follow**(Y) contains (at least) Follow(T) = {), $, +}   (no change)

Y → * **T**
j.) **Follow**(T) contains (at least) Follow(Y) = {), $ }   (no change)

Done.

Results

Symbol	First	Follow
(	(	N/A
)	)	
+	+	
*****	*	
int	int	
Y	ε, *	), $, +
X	ε, +	), $
T	int, (	), $, +
E	int, (	), $