## Section: Parsing

Parsing: Deciding if $x \in \Sigma^{*}$ is in $L(G)$ for some CFG G.

Consider the CFG G:
$\mathrm{S} \rightarrow \mathrm{Aa}$
$\mathbf{A} \rightarrow \mathbf{A A}|\mathbf{A B a}| \lambda$
$\mathbf{B} \rightarrow \mathbf{B B a}|\mathbf{b}| \lambda$

## Is ba in $L(G)$ ? Running time?

New grammar $G^{\prime}$ is:
$\mathrm{S} \rightarrow \mathrm{Aa} \mid \mathbf{a}$
$\mathrm{A} \rightarrow \mathrm{AA}|\mathrm{ABa}| \mathrm{Aa}|\mathrm{Ba}| \mathbf{a}$ $\mathrm{B} \rightarrow \mathrm{BBa}|\mathrm{Ba}| \mathrm{a} \mid \mathrm{b}$

Is ba in $L(G)$ ? Running time?

## Top-down Parser:

- Start with $S$ and try to derive the string.

$$
\mathrm{S} \rightarrow \mathrm{aS} \mid \mathrm{b}
$$

- Examples: LL Parser, Recursive Descent


## Bottom-up Parser:

- Start with string, work backwards, and try to derive S .
- Examples: Shift-reduce,

Operator-Precedence, LR Parser

## The function FIRST:

$$
\begin{aligned}
& \mathbf{G}=(\mathbf{V}, \mathbf{T}, \mathbf{S}, \mathbf{P}) \\
& \mathbf{w}, \mathbf{v} \in(\mathbf{V} \cup \mathbf{T})^{*} \\
& \mathbf{a} \in \mathbf{T} \\
& \mathbf{X}, \mathbf{A}, \mathbf{B} \in \mathbf{V} \\
& \mathbf{X}_{I} \in(\mathbf{V} \cup \mathbf{T})^{+}
\end{aligned}
$$

Definition: $\operatorname{FIRST}(w)=$ the set of terminals that begin strings derived from w.

If $\mathbf{w} \stackrel{*}{\Rightarrow}$ av then a is in FIRST(w)<br>If $\mathbf{w} \stackrel{*}{\Rightarrow} \lambda$ then $\lambda$ is in $\operatorname{FIRST}(\mathrm{w})$

## To compute FIRST:

1. $\operatorname{FIRST}(\mathbf{a})=\{\mathbf{a}\}$
2. $\operatorname{FIRST}(\mathrm{X})$
(a) If $\mathrm{X} \rightarrow$ aw then
a is in $\operatorname{FIRST}(\mathrm{X})$
(b) IF $\mathrm{X} \rightarrow \lambda$ then
$\lambda$ is in $\operatorname{FIRST}(\mathrm{X})$
(c) If X $\rightarrow$ Aw and $\lambda \in \operatorname{FIRST}(\mathbf{A})$ then
Everything in FIRST(w) is in FIRST(X)
3. In general, $\operatorname{FIRST}\left(\mathrm{X}_{1} \mathrm{X}_{2} \mathrm{X}_{3} . . \mathrm{X}_{K}\right)=$

- $\operatorname{FIRST}\left(\mathrm{X}_{1}\right)$
- $\cup \operatorname{FIRST}\left(\mathrm{X}_{2}\right)$ if $\lambda$ is in $\operatorname{FIRST}\left(\mathrm{X}_{1}\right)$
- $\cup \operatorname{FIRST}\left(\mathbf{X}_{3}\right)$ if $\lambda$ is in FIRST( $\mathrm{X}_{1}$ )
and $\lambda$ is in $\operatorname{FIRST}\left(\mathrm{X}_{2}\right)$
- $\cup \operatorname{FIRST}\left(\mathbf{X}_{K}\right)$ if $\lambda$ is in $\operatorname{FIRST}\left(\mathrm{X}_{1}\right)$
and $\lambda$ is in $\operatorname{FIRST}\left(\mathrm{X}_{2}\right)$
... and $\lambda$ is in $\operatorname{FIRST}\left(\mathrm{X}_{K-1}\right)$
- $-\{\lambda\}$ if $\lambda \notin \operatorname{FIRST}\left(\mathbf{X}_{J}\right)$ for all $\mathbf{J}$


## Example:

$$
\begin{aligned}
& \mathbf{S} \rightarrow \mathbf{a S c} \mid \mathbf{B} \\
& \mathbf{B} \rightarrow \mathbf{b} \mid \lambda
\end{aligned}
$$

$\operatorname{FIRST}(B)=$
$\operatorname{FIRST}(S)=$
$\operatorname{FIRST}(S c)=$

## Example

$$
\begin{aligned}
& \mathbf{S} \rightarrow \mathbf{B C D} \mid \mathbf{a D} \\
& \mathbf{A} \rightarrow \mathbf{C E B} \mid \mathbf{a A} \\
& \mathbf{B} \rightarrow \mathbf{b} \mid \lambda \\
& \mathbf{C} \rightarrow \mathbf{d B} \mid \lambda \\
& \mathbf{D} \rightarrow \mathbf{c A} \mid \lambda \\
& \mathbf{E} \rightarrow \mathbf{e} \mid \mathbf{f E}
\end{aligned}
$$

$\operatorname{FIRST}(S)=$
$\operatorname{FIRST}(\mathrm{A})=$
$\operatorname{FIRST}(\mathrm{B})=$
$\operatorname{FIRST}(\mathrm{C})=$
$\operatorname{FIRST}(\mathrm{D})=$
$\operatorname{FIRST}(E)=$

Definition: FOLLOW (X) = set of terminals that can appear to the right of $X$ in some derivation.

If $S \xrightarrow{*}$ wAav then a is in FOLLOW(A)

To compute FOLLOW:

1. $\$$ is in FOLLOW $(S)$
2. If $\mathbf{A} \rightarrow \mathbf{w B v}$ and $\mathbf{v} \neq \lambda$ then

FIRST(v) $-\{\lambda\}$ is in FOLLOW(B)
3. IF $\mathrm{A} \rightarrow \mathrm{wB}$ OR
$\mathrm{A} \rightarrow \mathrm{wBv}$ and $\lambda$ is in FIRST(v)
then
FOLLOW(A) is in FOLLOW(B)
4. $\lambda$ is never in FOLLOW

## Example:

$$
\begin{aligned}
& \mathbf{S} \rightarrow \mathbf{a S c} \mid \mathbf{B} \\
& \mathbf{B} \rightarrow \mathbf{b} \mid \lambda
\end{aligned}
$$

$\operatorname{FOLLOW}(S)=$
FOLLOW(B) =

## Example:

$$
\begin{aligned}
& \mathbf{S} \rightarrow \mathbf{B C D} \mid \mathbf{a D} \\
& \mathbf{A} \rightarrow \mathbf{C E B} \mid \mathbf{a A} \\
& \mathbf{B} \rightarrow \mathbf{b} \mid \lambda \\
& \mathbf{C} \rightarrow \mathbf{d B} \mid \lambda \\
& \mathbf{D} \rightarrow \mathbf{c A} \mid \lambda \\
& \mathbf{E} \rightarrow \mathbf{e} \mid \mathbf{f E}
\end{aligned}
$$

$\operatorname{FOLLOW}(S)=$
FOLLOW(A) =
FOLLOW(B) =
FOLLOW $(\mathrm{C})=$
FOLLOW (D) $=$
FOLLOW $(E)=$

