6-B Deterministic Left Parsers

Let G = (N, T, P, S) is a ***context-free grammar****. A* ***non-deterministic Left Parser*** $L_P = ((T^* \times V^*), \rightarrow_L, (x, S), \{(\varepsilon, \varepsilon)\}, P, \tau)$ *with*

$$\rightarrow_L = \{(\varepsilon, A) \rightarrow_L^{A\rightarrow\alpha} (\varepsilon, \alpha) \in (T^* \times V^*)^2 \mid (A \rightarrow \alpha) \in P\}$$

guess *A* ***as*** α: ***non-deterministic***.

$$\cup \{(a, a) \rightarrow_L^{a} (\varepsilon, \varepsilon) \in (T^* \times V^*)^2 \mid a \in \Sigma\}$$

verify $a \in \Sigma$: ***deterministic***.

Consider ***two guess*** *actions* $B \rightarrow \beta \mid \beta'$ *where* $\beta \neq \beta'$.

$(\varepsilon, B) \rightarrow_L^{B\rightarrow\beta} (\varepsilon, \beta)$ *or* $(\varepsilon, B) \rightarrow_L^{B\rightarrow\beta'} (\varepsilon, \beta')$ *is* ***non-deterministic***.

Consider $L(\beta)$ *and* $L(\beta')$*!*

Def. $L: (N \cup T)^* \rightarrow 2^{T^*}$. $L(\alpha) = \{x \in T^* \mid \alpha \Rightarrow^* x, x \in T^*\}$ ***infinite***.

But $L(\beta)$ *and* $L(\beta')$ *may be* ***infinite***.

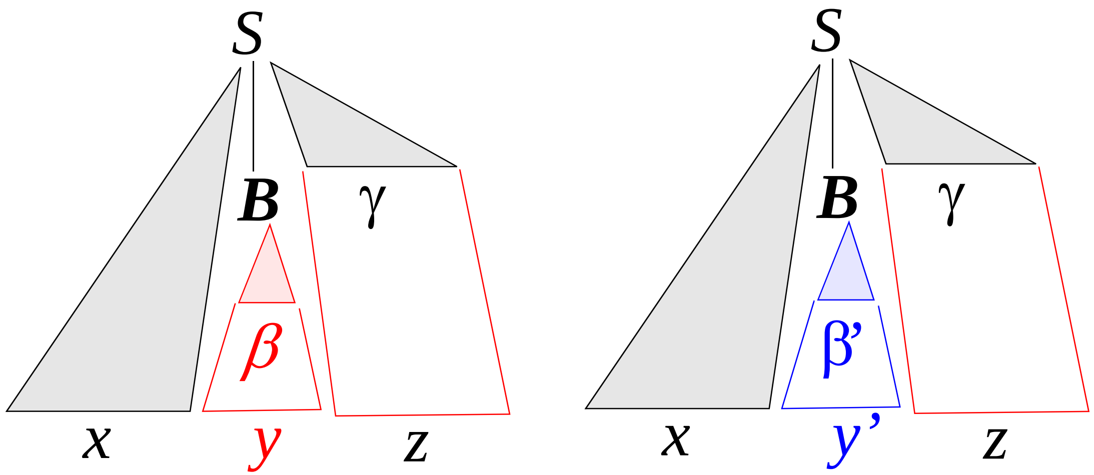
$\therefore\ First_k: (N \cup T)^* \rightarrow 2^{T^{\leq k}}$ where $T^{\leq k} = \{\varepsilon\} \cup T^1 \cup T^2 \cup \ldots T^k$.

$First_k(\alpha) = k{:}L(\alpha) = \{k{:}x \in T^{\leq k} \mid \alpha \Rightarrow^* x, x \in T^*\}$ ***finite***.

If $First_k(\beta) \cap First_k(\beta') = \emptyset$, then lookahead string y and y'

$(y, B) \rightarrow_L (y, \beta)\ y \in First_k(\beta)$ and $(y', B) \rightarrow_L (y', \beta')\ y' \in First_k(\beta')$

$\therefore$ ***deterministic***!



Computation *of* $First_k(\alpha)$ *for* $\alpha \in (N \cup T)^*$.

$First_k(\alpha) = First_k(X_1X_2...X_n)$ where $\alpha = X_1X_2...X_n$.

$= First_k(X_1) \oplus_k First_k(X_2) \oplus_k ... \oplus_k First_k(X_n)$ where

$First_k(a) = \{a\}, a \in T$(**basis**).

$\oplus_k: 2^{T^{\leq k}} \times 2^{T^{\leq k}} \rightarrow 2^{T^{\leq k}}$.

$A \oplus_k B = k:AB = \{k:xy \in T^{\leq k} | x \in A, y \in B\}$

for $a \in T$ ***do*** $First_k(a) = \{a\}$ ***od; for*** $A \in N$ ***do*** $First_k(A) := \emptyset$ ***od;***

repeat

for $A \in N$ ***do***

for $A \rightarrow X_1X_2...X_n \in P$ ***do***

for *i:=1 to n* ***do*** $First_k(A) := First_k(A) \oplus_k First_k(X_i)$ ***od***

od od

until $First_k(A)$ *does not change*

Example G_{Uexp}: $E \rightarrow E + T \mid T$

$T \rightarrow T * F \mid F$

$F \rightarrow a \mid (E)$

F: $First_1(a) = \{a\}$ $First_1((E)) = \{(\}$

$\therefore (a, \mathbf{F}) \rightarrow (a, a)$ $((, \mathbf{F}) \rightarrow ((, (E))$ ***deterministic!***

T: $First_1(T * F) = \{a, (\}$ $First_1(F) = \{a, (\}$

$\therefore (a, \mathbf{T}) \rightarrow (a, T * F)$ $(a, \mathbf{T}) \rightarrow (a, F)$ ***non-deterministic!***

$((, \mathbf{T}) \rightarrow ((, T * F)$ $((, \mathbf{T}) \rightarrow ((, F)$ ***non-deterministic!***

E: $First_1(E + T) = \{a, (\}$ $First_1(T) = \{a, (\}$

$\therefore (a, \mathbf{E}) \rightarrow (a, E + T)$ $(a, \mathbf{E}) \rightarrow (a, T)$ ***non-deterministic!***

$((, \mathbf{E}) \rightarrow ((, E + T)$ $((, \mathbf{E}) \rightarrow ((, T)$ ***non-deterministic!***

$\therefore$ *Non-deterministic for* **T** *and* **E**!

A grammar rule $A \rightarrow A\alpha \in P$ *is said to be* ***left recursive***.

If $A \rightarrow A\alpha \in P$, $A \Rightarrow^* A\alpha^*$ *does* ***not terminate***! $\therefore\ {}^{\exists}A \rightarrow \beta \in P$

$A \rightarrow A\alpha \mid \beta$, *then* $A \Rightarrow^* \beta\alpha^*$.

$First_k(A\alpha) \cap First_k(\beta) \neq \emptyset$, *since* $First_k(A\alpha) \supseteq First_k(A) \supseteq First_k(\beta)$.

$\therefore$ *If a grammar G has a* ***left recursive*** *rule,*

the ***left parser*** *for G is* ***non-deterministic***!

What can I do?

Change the ***left recursion*** *to* ***right recursion***.

$A \rightarrow A\alpha \mid \beta \quad \equiv A \rightarrow \beta A', A' \rightarrow \alpha A' \mid \varepsilon. \quad (A \Rightarrow \beta A' \Rightarrow^* \beta\alpha^*).$

G_{Dexp}: $E \rightarrow T\ E'$

$E' \rightarrow +\ T\ E' \mid \varepsilon$

$T \rightarrow F\ T'$

$T' \rightarrow *\ F\ T' \mid \varepsilon$

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

F: $First_1(a) = \{a\}$, $First_1((\ E\ )) = \{(\}$

$(a, \mathbf{F}) \rightarrow (a, a)$, $((, \mathbf{F}) \rightarrow ((, (\ E\ ))$ **deterministic!**

T': $First_1(*\ F\ T') = \{*\}$ $First_1(\varepsilon) = \{\varepsilon\}$

$(*, \mathbf{T'}) \rightarrow (a, *\ F\ T')$ $(\varepsilon, \mathbf{T'}) \rightarrow (\varepsilon, \varepsilon)$ **non-deterministic!**

T: $First_1(F\ T') = \{*\}$

$(*, \mathbf{T}) \rightarrow (*, *\ F\ T')$ **deterministic(unique rule)!**

E': $First_1(*\ F\ T') = \{*\}$ $First_1(\varepsilon) = \{\varepsilon\}$

$(*, \mathbf{E'}) \rightarrow (*, *\ F\ T')$ $(\varepsilon, \mathbf{E'}) \rightarrow (\varepsilon, \varepsilon)$ **non-deterministic!**

E: $First_1(+\ T\ E') = \{+\}$

$(+, \mathbf{E}) \rightarrow (+, +\ T\ E')$ **deterministic(unique rule)!**

*We are **happy** except for $E' \rightarrow \varepsilon$ and $T' \rightarrow \varepsilon$.*
What can I do?

Consider $Follow_k(E')$ and $Follow_k(T')$ or k:z in the figure!

$Follow_k: N \to 2^{T^{\leq k}}$.

$$Follow_k(A) = \{k{:}z \in T^{\leq k} \mid S \Rightarrow^* \alpha A z,\ z \in T^*\}$$ ***finite***.

$Follow_k(T') = \{+,), \varepsilon\}$, $Follow_k(E') = \{), \varepsilon\}$

T': $First_1(*\ F\ T') = \{*\}$

$First_1(\varepsilon) \oplus_1 Follow_1(T') = \{+,), \varepsilon\}$ two sets are ***disjoint***!

$(*, T') \to (*, *\ F\ T')$

$(+, T') \to (+, \varepsilon)$, $(), T') \to (), \varepsilon)$, $(\varepsilon, T') \to (\varepsilon, \varepsilon)$ ***deterministic!***

$(\$, T') \to (\$, \varepsilon)$

E': $First_1(+\ T\ E')=\{+\}$

$First_1(\varepsilon) \oplus_1 Follow_1(E') = \{), \varepsilon\}$ two sets are ***disjoint***!

$(+, E') \to (+, +\ T\ E')$

$(), E') \to (), \varepsilon,)$, $(\varepsilon, E') \to (\varepsilon, \varepsilon)$ ***deterministic!***

$(\$, T') \to (\$, \varepsilon)$

Parsing table for G_{Dexp}.: $SLL_1PT[\mathbf{N}, T] \rightarrow 2^P$.

G_{Dexp}:	$E \rightarrow T\,E'$	$\{a, (\}$	
	$E' \rightarrow +\,T\,E' \mid \varepsilon$	$\{+\}$	$\{), \varepsilon\}$
	$T \rightarrow F\,T'$	$\{a, (\}$	
	$T' \rightarrow *\,F\,T' \mid \varepsilon$	$\{*\}$	$\{+,), \varepsilon\}$
	$F \rightarrow a \mid (\,E\,)$	$\{a\}$	$\{(\}$

	a	(	*	+	)	$\varepsilon(\$)$
E	$E \rightarrow T\,E'$	$E \rightarrow T\,E'$				
E'				$E' \rightarrow +\,T\,E'$	$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
T	$T \rightarrow F\,T'$	$T \rightarrow F\,T'$				
T'			$T' \rightarrow *\,F\,T'$	$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
F	$F \rightarrow a$	$F \rightarrow (\,E\,)$				

Parsing table for G_{Dexp}.: $SLL_1PT[\mathbf{N}, T] \rightharpoonup P$. $\therefore$ ***deterministic***!

Ex.	*a*	(	*	+	)	ε($)
E	$E \to T E'$	$E \to T E'$				
E'				$E' \to + T E'$	$E' \to \varepsilon$	$E' \to \varepsilon$
T	$T \to F T'$	$T \to F T'$				
T'			$T' \to * F T'$	$T' \to \varepsilon$	$T' \to \varepsilon$	$T' \to \varepsilon$
F	$F \to a$	$F \to (E)$				

$$(a * a + a, E) \Rightarrow_L^{(a,E)\to(a,TE')} (a * a + a, T E') \Rightarrow_L^{(a,T)\to(a,FT')} (a * a + a, F T' E')$$

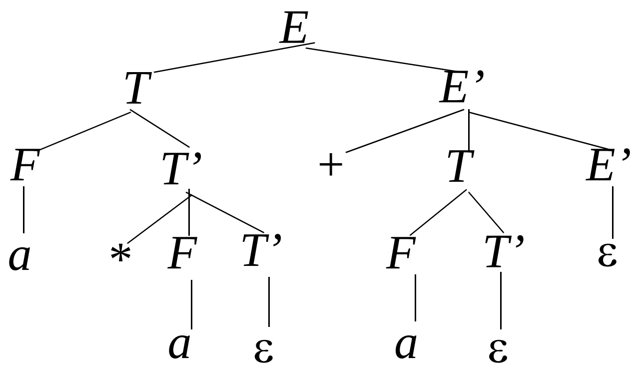
$$\Rightarrow_L^{(a,F)\to(a,a)} (a * a + a, a T' E') \Rightarrow_L^{(a,a)\to(\varepsilon,\varepsilon)} (* a + a, T' E')$$

$$\Rightarrow_L^{(*,T')\to(*,*FT')} (* a + a, * F T' E') \Rightarrow_L^{(*,*)\to(\varepsilon,\varepsilon)} (a + a, F T' E') \Rightarrow_L^{(a,F)\to(a,a)} (a + a, a T' E')$$

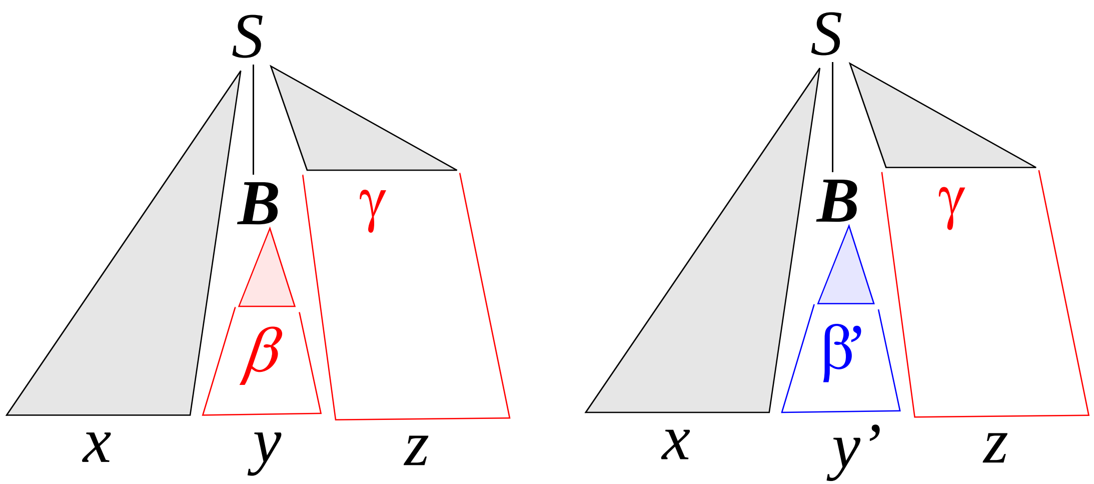
$$\Rightarrow_L^{(a,a)\to(\varepsilon,\varepsilon)} (+ a, T' E') \Rightarrow_L^{(+,T')\to(+,\varepsilon)} (+ a, E') \Rightarrow_L^{(+,E')\to(+,+TE')} (+ a, + T' E')$$

$$\Rightarrow_L^{(+,+)\to(\varepsilon,\varepsilon)} (a, T' E') \Rightarrow_L^{(a,T')\to(a,FT')} (a, F T' E') \Rightarrow_L^{(a,F)\to(a,a)} (a, a T' E')$$

$$\Rightarrow_L^{(a,a)\to(\varepsilon,\varepsilon)} (\varepsilon, T' E') \Rightarrow_L^{(\varepsilon,T')\to(\varepsilon,\varepsilon)} (\varepsilon, E') \Rightarrow_L^{(\varepsilon,E')\to(\varepsilon,\varepsilon)} (\varepsilon, \varepsilon).$$



Adding lookahead string k:yz $\in T^{\leq k}$ *for* ***guess B as*** β.



$$(x, B) \rightarrow_L (x, \beta) \in (T^{\leq k} \times V^*) \times (T^{\leq k} \times V^*) \text{ for } B \rightarrow \beta \in P \text{ where}$$

$$x \in First_k(\beta) \oplus_k Follow(B) = k{:}yz.$$

LL(k) Parser

Left-to-right Scan in Leftmost derivation with k-lookahead symbols

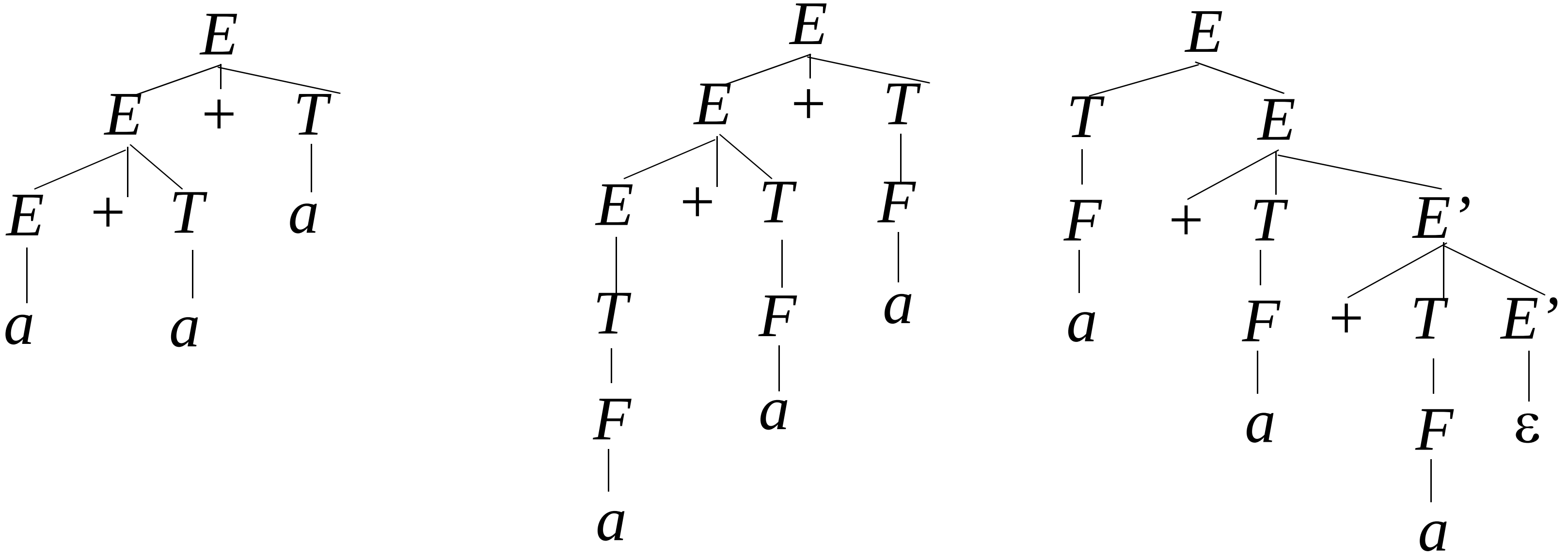
Strong LL(k) Parser(SLL(k) Parser)

SLL(k) grammars $\subset$ *LL(k) grammars*

But SLL(1) grammars = LL(1) grammars

*Compare three grammars for a + a + a or a * a * a*

G_{exp}:	$E \rightarrow E + T \mid T * F \mid a \mid (E)$	$E \rightarrow E + T \mid T$	$E \rightarrow T E'$
			$E' \rightarrow + T E' \mid \varepsilon$
	$T \rightarrow T * F \mid a \mid (E)$	$T \rightarrow T * F \mid F$	$T \rightarrow F T'$
			$T' \rightarrow * F T' \mid \varepsilon$
	$F \rightarrow a \mid (E)$	$F \rightarrow a \mid (E)$	$F \rightarrow a \mid (E)$



good! *unit production* *+ is right associative*

How about right parsers?