EECS4302 Winter 2020
Compilers and Interpreters
Instructor: Chen-Wei Wang
Solution to Exercise on Top-Down, Predictive Parser

---

1. Consider the following grammar:

$$\begin{array}{lll} L & \rightarrow & R\ \texttt{a} \\ & | & Q\ \texttt{ba} \end{array} \quad \begin{array}{lll} R & \rightarrow & \texttt{aba} \\ & | & \texttt{caba} \\ & | & R\ \texttt{bc} \end{array} \quad \begin{array}{lll} Q & \rightarrow & \texttt{bbc} \\ & | & \texttt{bc} \end{array}$$

Is it suitable for a **top-down predictive** parser?
- If so, show that it satisfies the LL(1) condition.
- If not, identify the problem(s) and correct it (them). Also show that the revised grammar satisfies the LL(1) condition.

Solution:

- The given grammar contains a **direct** left-recursion on the non-terminal R:

$$\begin{array}{lll} R & \rightarrow & \texttt{aba} \\ & | & \texttt{caba} \\ & | & R\ \texttt{bc} \end{array}$$

By removing the left-recursion, we rewrite the above productions as:

$$\begin{array}{lll} R & \rightarrow & \texttt{aba}R' \\ & | & \texttt{caba}R' \\ R' & \rightarrow & \texttt{bc}R' \\ & | & \epsilon \end{array}$$

Here is the revised grammar:

$$\begin{array}{lll} L & \rightarrow & R\ \texttt{a} \\ & | & Q\ \texttt{ba} \end{array} \quad \begin{array}{lll} R & \rightarrow & \texttt{aba}R' \\ & | & \texttt{caba}R' \\ R' & \rightarrow & \texttt{bc}R' \\ & | & \epsilon \end{array} \quad \begin{array}{lll} Q & \rightarrow & \texttt{bbc} \\ & | & \texttt{bc} \end{array}$$

- However, the revised grammar still fails the LL(1) condition: Each of the productions $A \rightarrow \gamma_1 \mid \gamma_2 \mid \cdots \mid \gamma_n$ satisfying

$$\forall i, j : 1 \leq i, j \leq n \land i \neq j \bullet \textbf{First}^+(\gamma_i) \cap \textbf{First}^+(\gamma_j) = \varnothing$$

Specifically, this production fails the above LL(1) condition by having a common prefix `b` among the RHS of multiple production rules:

$$\begin{array}{lll} Q & \rightarrow & \texttt{bbc} \\ & | & \texttt{bc} \end{array}$$

We then apply left factoring to remove the common prefix:

$$\begin{array}{lll} Q & \rightarrow & \texttt{b}Q' \\ Q' & \rightarrow & bc \\ & | & \texttt{c} \end{array}$$

Here is the revised grammar:

$$
\begin{array}{lcl}
L & \rightarrow & R\ \mathtt{a} \\
 & | & Q\ \mathtt{ba}
\end{array}
\qquad
\begin{array}{lcl}
R & \rightarrow & \mathtt{aba}R' \\
 & | & \mathtt{caba}R' \\
R' & \rightarrow & \mathtt{bc}R' \\
 & | & \epsilon
\end{array}
\qquad
\begin{array}{lcl}
Q & \rightarrow & \mathtt{b}Q' \\
Q' & \rightarrow & bc \\
 & | & \mathtt{c}
\end{array}
$$

- To show that this revised grammar is LL(1), we must then show that the **First**$^+$ sets of the alternative RHSs for each non-terminal are **disjoint**:

Non-Terminal	Alternative	**First**$^+$ Set	Intersection
L	Ra	{a, c}	$\varnothing$
	Qba	{b}	
R	abaR'	{a}	$\varnothing$
	cabaR'	{c}	
R'	bcR'	{b}	$\varnothing$
	ϵ	**Follow**(R) = {a}	
Q'	bc	{b}	$\varnothing$
	c	{c}	