

# CS335: COMPILER DESIGN - Assignment 2

**Harsh Bihany (210406)**

# Problem 1

Consider the following context-free grammar. The terminal symbols are **id** ( ) \* , and \$. *Function* is the start nonterminal.

$$\begin{aligned}Function &\rightarrow Type \textbf{id} (Arguments) \\Type &\rightarrow \textbf{id} \\Type &\rightarrow Type * \\Arguments &\rightarrow ArgList \\Arguments &\rightarrow \epsilon \\ArgList &\rightarrow Type \textbf{id} , ArgList \\ArgList &\rightarrow Type \textbf{id}\end{aligned}$$

- (i) Explain why the grammar is not LL(1)?

A grammar is *left recursive* if it has a non-terminal  $A$  such that there exists a derivation  $A \xrightarrow{*} A\alpha$  for some string  $\alpha$ . In the following grammar we have one such derivation  $Type \rightarrow Type*$ .

Further a left recursive grammar cannot be LL(1). Hence the above grammar is not LL(1).

- (ii) Perform required transformations on the grammar to make it LL(1). You may introduce new nonterminals. Show your transformed grammar.

To make the grammar LL(1), we need two things. Elimination of direct and indirect left recursion and left factoring to eliminate backtracking if needed.

Removing the direct left recursion mentioned earlier. We get

$$\begin{aligned}Function &\rightarrow Type \textbf{id} (Arguments) \\Type &\rightarrow \textbf{id} Type' \\Type' &\rightarrow *Type' \\Type' &\rightarrow \epsilon \\Arguments &\rightarrow ArgList \\Arguments &\rightarrow \epsilon \\ArgList &\rightarrow Type \textbf{id} , ArgList \\ArgList &\rightarrow Type \textbf{id}\end{aligned}$$

Now, after left factoring (required for the nonterminal *ArgList*) we get,

$$Function \rightarrow Type \textbf{id} (Arguments) \quad (1)$$

$$Type \rightarrow \textbf{id} Type' \quad (2)$$

$$Type' \rightarrow *Type' \quad (3)$$

$$Type' \rightarrow \epsilon \quad (4)$$

$$Arguments \rightarrow ArgList \quad (5)$$

$$Arguments \rightarrow \epsilon \quad (6)$$

$$ArgList \rightarrow Type \textbf{id} More \quad (7)$$

$$More \rightarrow , ArgList \quad (8)$$

$$More \rightarrow \epsilon \quad (9)$$

*Remark.* The productions are labeled so that they can be referred subsequently.

(iii) Show FIRST and FOLLOW set for the nonterminals in your transformed grammar.

$$\begin{array}{ll} \text{FIRST}(\mathbf{T}) = \{\mathbf{T}\} \quad \forall \mathbf{T} \in \{\textbf{id}, (, ), *, ,, \$, \epsilon\} & \text{FOLLOW}(Function) = \{\$\} \\ \text{FIRST}(Function) = \{\textbf{id}\} & \text{FOLLOW}(Type) = \{\textbf{id}\} \\ \text{FIRST}(Type) = \{\textbf{id}\} & \text{FOLLOW}(Type') = \{\textbf{id}\} \\ \text{FIRST}(Type') = \{*, \epsilon\} & \text{FOLLOW}(Arguments) = \{\}\} \\ \text{FIRST}(Arguments) = \{\textbf{id}, \epsilon\} & \text{FOLLOW}(ArgList) = \{\}\} \\ \text{FIRST}(ArgList) = \{\textbf{id}\} & \text{FOLLOW}(More) = \{\}\} \\ \text{FIRST}(More) = \{,, \epsilon\} & \end{array}$$

(iv) Show the predictive LL(1) parsing table for your transformed grammar. You do not need to show the working of the parser with an example input string.

NON-TERMINALS	INPUTS					
	\$	(	)	id	*	,
<i>Function</i>				(1)		
<i>Type</i>				(2)		
<i>Type'</i>				(4)	(3)	
<i>Arguments</i>			(6)	(5)		
<i>ArgList</i>				(7)		
<i>More</i>			(9)			(8)

Table 1.1: Predictive LL(1) Parsing Table

# Problem 2

Consider the following context-free grammar. The terminal symbols are  $a, b, p, q, r, s$ , and  $t$ .  $S$  is the start nonterminal.

$$\begin{aligned} S &\rightarrow LM \mid Lp \mid qLr \mid sr \mid qsp \\ L &\rightarrow aMb \mid s \mid t \\ M &\rightarrow t \end{aligned}$$

(i) Is the given CFG SLR(1)?

After augmentation of the grammar and numbering the productions:

$$\begin{aligned} S' &\rightarrow S & (0) \\ S &\rightarrow LM & (1) \\ S &\rightarrow Lp & (2) \\ S &\rightarrow qLr & (3) \\ S &\rightarrow sr & (4) \\ S &\rightarrow qsp & (5) \\ L &\rightarrow aMb & (6) \\ L &\rightarrow s & (7) \\ L &\rightarrow t & (8) \\ M &\rightarrow t & (9) \end{aligned}$$

The *FIRST* and *FOLLOW* sets are as follows:

$$\begin{aligned} \text{FIRST}(\mathbf{T}) &= \{\mathbf{T}\} \quad \forall \mathbf{T} \in \{a, b, p, q, r, s, t\} \\ \text{FIRST}(S) &= \{a, q, s, t\} \\ \text{FIRST}(L) &= \{a, s, t\} \\ \text{FIRST}(M) &= \{t\} \end{aligned}$$

$$\begin{aligned} \text{FOLLOW}(S') &= \{\$ \} \\ \text{FOLLOW}(S) &= \{\$ \} \\ \text{FOLLOW}(L) &= \{p, r, t\} \\ \text{FOLLOW}(M) &= \{\$, b\} \end{aligned}$$

Now onto the *LR*(0) canonical collection

$$\begin{aligned}
I_0 &= \text{Closure}(\{[S' \rightarrow \cdot S]\}) \\
&= \{[S' \rightarrow \cdot S], \\
&\quad [S \rightarrow \cdot LM], \\
&\quad [S \rightarrow \cdot Lp], \\
&\quad [S \rightarrow \cdot qLr], \\
&\quad [S \rightarrow \cdot sr], \\
&\quad [S \rightarrow \cdot qsp], \\
&\quad [L \rightarrow \cdot aMb], \\
&\quad [L \rightarrow \cdot s], \\
&\quad [L \rightarrow \cdot t], \\
&\quad \} \\
I_1 &= \text{Goto}(I_0, S) \\
&= \{[S' \rightarrow S \cdot], \\
&\quad \} \\
I_2 &= \text{Goto}(I_0, L) \\
&= \{[S \rightarrow L \cdot M], \\
&\quad [S \rightarrow L \cdot p], \\
&\quad [M \rightarrow \cdot t], \\
&\quad \} \\
I_3 &= \text{Goto}(I_0, q) \\
&= \{[S \rightarrow q \cdot Lr], \\
&\quad [S \rightarrow q \cdot sp], \\
&\quad [L \rightarrow \cdot aMb], \\
&\quad [L \rightarrow \cdot s], \\
&\quad [L \rightarrow \cdot t], \\
&\quad \} \\
I_4 &= \text{Goto}(I_0, s) \\
&= \{[S \rightarrow s \cdot r], \\
&\quad [L \rightarrow s \cdot], \\
&\quad \} \\
I_5 &= \text{Goto}(I_0, a) \\
&= \{[L \rightarrow a \cdot Mb], \\
&\quad [M \rightarrow \cdot t], \\
&\quad \} \\
I_6 &= \text{Goto}(I_3, t) \\
I_{13} &= \text{Goto}(I_2, t) \\
I_6 &= \text{Goto}(I_0, t) \\
&= \{[L \rightarrow t \cdot], \\
&\quad \} \\
I_7 &= \text{Goto}(I_2, M) \\
&= \{[S \rightarrow LM \cdot], \\
&\quad \} \\
I_8 &= \text{Goto}(I_2, p) \\
&= \{[S \rightarrow Lp \cdot], \\
&\quad \} \\
I_9 &= \text{Goto}(I_3, L) \\
&= \{[S \rightarrow qL \cdot r], \\
&\quad \} \\
I_{10} &= \text{Goto}(I_3, s) \\
&= \{[S \rightarrow qs \cdot p], \\
&\quad [L \rightarrow s \cdot], \\
&\quad \} \\
I_{11} &= \text{Goto}(I_4, r) \\
&= \{[S \rightarrow sr \cdot], \\
&\quad \} \\
I_{12} &= \text{Goto}(I_5, M) \\
&= \{[L \rightarrow aM \cdot b], \\
&\quad \} \\
I_{13} &= \text{Goto}(I_5, t) \\
&= \{[M \rightarrow t \cdot], \\
&\quad \} \\
I_{14} &= \text{Goto}(I_9, r) \\
&= \{[S \rightarrow qLr \cdot], \\
&\quad \} \\
I_{15} &= \text{Goto}(I_{10}, p) \\
&= \{[S \rightarrow qsp \cdot], \\
&\quad \} \\
I_{16} &= \text{Goto}(I_{12}, b) \\
&= \{[L \rightarrow aMb \cdot], \\
&\quad \} \\
I_5 &= \text{Goto}(I_3, a)
\end{aligned}$$

Figure 2.1: LR(0) Canonical Collection

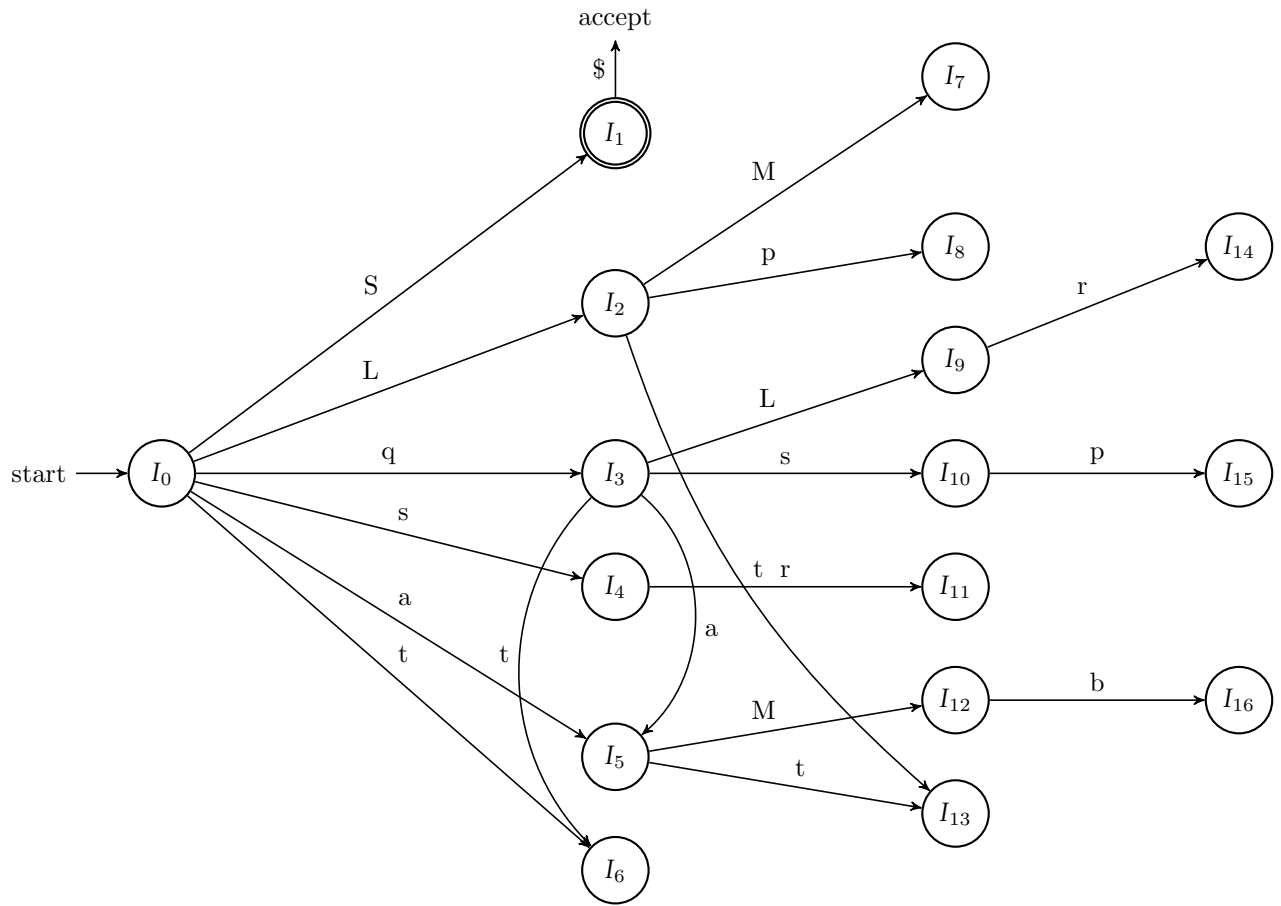


Figure 2.2: LR(0) automaton