

CS335: Compiler Design

Assignment 2

Yash Gupta (190997)

1 Problem 1

1.1 Part (i)

The given grammar is not LL(1) because of the following rules:

- $Type \rightarrow Type*$ introduces left recursion
- $ArgList \rightarrow Typeid, ArgList$ and $ArgList \rightarrow Typeid$ have common prefixes

1.2 Part (ii)

The nonterminals $Type'$ and $ArgList'$ can be introduced to make the grammar LL(1). The transformed grammar is as follows:

- (1) $Function \rightarrow Type\ id\ (Arguments)$
- (2) $Type \rightarrow id\ Type'$
- (3) $Type' \rightarrow *Type'$
- (4) $Type' \rightarrow \epsilon$
- (5) $Arguments \rightarrow ArgList$
- (6) $Arguments \rightarrow \epsilon$
- (7) $ArgList \rightarrow Type\ id\ ArgList'$
- (8) $ArgList' \rightarrow ,ArgList$
- (9) $ArgList' \rightarrow \epsilon$

1.3 Part (iii)

Nonterminal	FIRST	FOLLOW
Function	id	\$
Type	id	id
Type'	*, ϵ	id
Arguments	id, ϵ)
ArgList	id)
ArgList'	,, ϵ)

1.4 Part (iv)

Nonterminal	id	(()	*	,	\$
Function	$Function \rightarrow Type\ id\ (Arguments)$					
Type	$Type \rightarrow id\ Type'$					
Type'	$Type' \rightarrow \epsilon$			$Type' \rightarrow *Type'$		
Arguments	$Arguments \rightarrow ArgList$		$Arguments \rightarrow \epsilon$			
ArgList	$ArgList \rightarrow Type\ id\ ArgList'$					
ArgList'			$ArgList' \rightarrow \epsilon$		$ArgList' \rightarrow ,ArgList$	

2 Problem 2

2.1 Part (i)

2.1.1 Augmented grammar

- (0) $S' \rightarrow S$
- (1) $S \rightarrow LM$
- (2) $S \rightarrow Lp$
- (3) $S \rightarrow qLr$
- (4) $S \rightarrow sr$
- (5) $S \rightarrow qsp$
- (6) $L \rightarrow aMb$
- (7) $L \rightarrow s$
- (8) $L \rightarrow t$
- (9) $M \rightarrow t$

2.1.2 FIRST and FOLLOW sets

Nonterminal	FIRST	FOLLOW
S'	a, s, t, q	\$
S	a, s, t, q	\$
L	a, s, t	t, p, r
M	t	b, \$

2.1.3 LR(0) canonical collection

$$\begin{aligned} I_0 &= \text{Closure}(S' \rightarrow \cdot S) \\ &= \{S' \rightarrow \cdot S, \\ &\quad S \rightarrow \cdot LM, \\ &\quad S \rightarrow \cdot Lp, \\ &\quad S \rightarrow \cdot qLr, \\ &\quad S \rightarrow \cdot sr, \\ &\quad S \rightarrow \cdot qsp, \\ &\quad L \rightarrow \cdot aMb, \\ &\quad L \rightarrow \cdot s, \\ &\quad L \rightarrow \cdot t, \end{aligned}$$

$$\begin{aligned} I_1 &= \text{Goto}(I_0, S) \\ &= \{S' \rightarrow S \cdot\} \end{aligned}$$

$$\begin{aligned} I_2 &= \text{Goto}(I_0, L) \\ &= \{S \rightarrow L \cdot M, \\ &\quad S \rightarrow L \cdot p, \\ &\quad M \rightarrow \cdot t\} \end{aligned}$$

$$\begin{aligned} I_3 &= \text{Goto}(I_0, q) \\ &= \{S \rightarrow q \cdot Lr, \\ &\quad S \rightarrow q \cdot sp, \\ &\quad L \rightarrow \cdot aMb, \\ &\quad L \rightarrow \cdot s, \\ &\quad L \rightarrow \cdot t\} \end{aligned}$$

$$\begin{aligned} I_4 &= \text{Goto}(I_0, s) \\ &= \{S \rightarrow s \cdot r, \\ &\quad L \rightarrow s \cdot\} \end{aligned}$$

$$\begin{aligned} I_5 &= \text{Goto}(I_0, a) \\ &= \{L \rightarrow a \cdot Mb, \\ &\quad M \rightarrow \cdot t\} \end{aligned}$$

$$\begin{aligned} I_6 &= \text{Goto}(I_0, t) \\ &= \{L \rightarrow t \cdot\} \end{aligned}$$

$$\begin{aligned}
I_7 &= \text{Goto}(I_2, M) \\
&= \{S \rightarrow LM\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_8 &= \text{Goto}(I_2, p) \\
&= \{S \rightarrow Lp\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_9 &= \text{Goto}(I_2, t) \\
&= \{M \rightarrow t\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_{10} &= \text{Goto}(I_3, L) \\
&= \{S \rightarrow qL \cdot r\}
\end{aligned}$$

$$\begin{aligned}
I_{11} &= \text{Goto}(I_3, s) \\
&= \{S \rightarrow qs \cdot p, \\
&\quad L \rightarrow s\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_{12} &= \text{Goto}(I_4, r) \\
&= \{S \rightarrow sr\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_{13} &= \text{Goto}(I_5, M) \\
&= \{L \rightarrow aM \cdot b\}
\end{aligned}$$

$$\begin{aligned}
I_{14} &= \text{Goto}(I_{10}, r) \\
&= \{S \rightarrow qLr\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_{15} &= \text{Goto}(I_{11}, p) \\
&= \{S \rightarrow qsp\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_{16} &= \text{Goto}(I_{13}, b) \\
&= \{L \rightarrow aMb\cdot\}
\end{aligned}$$

$$\begin{aligned}
I_5 &= \text{Goto}(I_3, a) \\
I_6 &= \text{Goto}(I_3, t) \\
I_9 &= \text{Goto}(I_5, t)
\end{aligned}$$

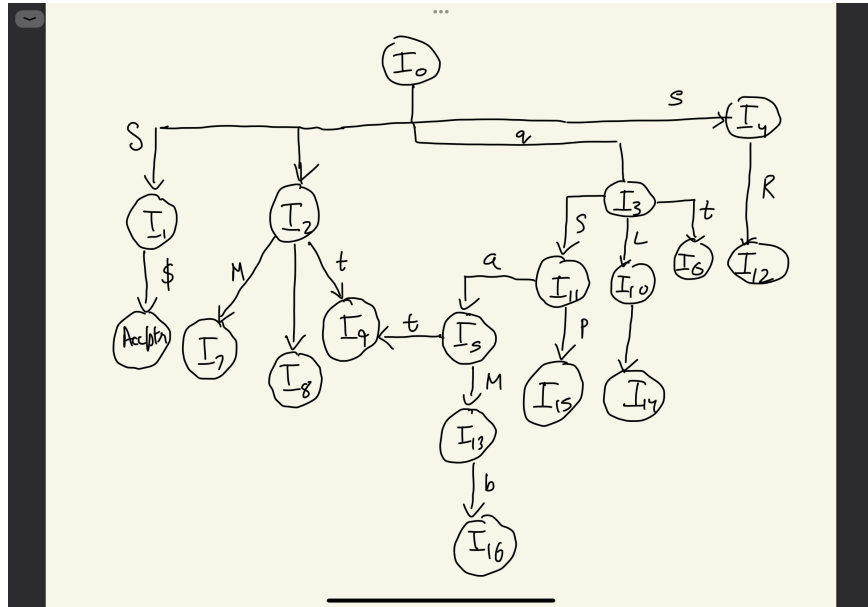


Figure 1: LR(0) automaton

2.1.4 LR(0) automaton

2.1.5 SLR parsing table

State	ACTION								GOTO		
	a	b	p	q	r	s	t	\$	S	L	M
0	s5			s3		s4	s6		1	2	
1								acc			
2			s8				s9				7
3	s5					s11	s6			10	
4			r7		s12/r7		r7				
5							s9				13
6			r8		r8		r8				
7								r1			
8								r2			
9		r9						r9			
10					s14						
11			s15/r7		r7		r7				
12								r4			
13		s16									
14								r3			
15								r5			
16			r6		r6		r6				

2.1.6 Conclusion

As there are shift-reduce conflicts (in **bold**) in the SLR parsing table, the given CFG is not SLR(1).

2.2 Part (ii)

2.2.1 LR(1) collection

$$\begin{aligned} I_0 &= \text{Closure}(S' \rightarrow \cdot S, \$) \\ &= \{S' \rightarrow \cdot S, \$, \\ &\quad S \rightarrow \cdot LM, \$, \\ &\quad S \rightarrow \cdot Lp, \$, \\ &\quad S \rightarrow \cdot qLr, \$, \\ &\quad S \rightarrow \cdot sr, \$, \\ &\quad S \rightarrow \cdot qsp, \$, \\ &\quad L \rightarrow \cdot aMb, t/p, \\ &\quad L \rightarrow \cdot s, t/p, \\ &\quad L \rightarrow \cdot t, t/p, \end{aligned}$$

$$\begin{aligned} I_1 &= \text{Goto}(I_0, S) \\ &= \{S' \rightarrow S \cdot, \$\} \end{aligned}$$

$$\begin{aligned} I_2 &= \text{Goto}(I_0, L) \\ &= \{S \rightarrow L \cdot M, \$, \\ &\quad S \rightarrow L \cdot p, \$, \\ &\quad M \rightarrow \cdot t, \$\} \end{aligned}$$

$$\begin{aligned} I_3 &= \text{Goto}(I_0, q) \\ &= \{S \rightarrow q \cdot Lr, \$, \\ &\quad S \rightarrow q \cdot sp, \$, \\ &\quad L \rightarrow \cdot aMb, r, \\ &\quad L \rightarrow \cdot s, r, \\ &\quad L \rightarrow \cdot t, r\} \end{aligned}$$

$$\begin{aligned} I_4 &= \text{Goto}(I_0, s) \\ &= \{S \rightarrow s \cdot r, \$, \\ &\quad L \rightarrow s \cdot, t/p\} \end{aligned}$$

$$\begin{aligned} I_5 &= \text{Goto}(I_0, a) \\ &= \{L \rightarrow a \cdot Mb, t/p, \\ &\quad M \rightarrow \cdot t, b\} \end{aligned}$$

$$\begin{aligned}
I_6 &= \text{Goto}(I_0, t) \\
&= \{L \rightarrow t\cdot, t/p\}
\end{aligned}$$

$$\begin{aligned}
I_7 &= \text{Goto}(I_2, M) \\
&= \{S \rightarrow LM\cdot, \$\}
\end{aligned}$$

$$\begin{aligned}
I_8 &= \text{Goto}(I_2, p) \\
&= \{S \rightarrow Lp\cdot, \$\}
\end{aligned}$$

$$\begin{aligned}
I_9 &= \text{Goto}(I_2, t) \\
&= \{M \rightarrow t\cdot, \$\}
\end{aligned}$$

$$\begin{aligned}
I_{10} &= \text{Goto}(I_3, L) \\
&= \{S \rightarrow qL\cdot r, \$\}
\end{aligned}$$

$$\begin{aligned}
I_{11} &= \text{Goto}(I_3, s) \\
&= \{S \rightarrow qs\cdot p, \$, \\
&\quad L \rightarrow s\cdot, r\}
\end{aligned}$$

$$\begin{aligned}
I_{12} &= \text{Goto}(I_3, a) \\
&= \{L \rightarrow a\cdot Mb, r, \\
&\quad M \rightarrow \cdot t, b\}
\end{aligned}$$

$$\begin{aligned}
I_{13} &= \text{Goto}(I_3, t) \\
&= \{L \rightarrow t\cdot, r\}
\end{aligned}$$

$$\begin{aligned}
I_{14} &= \text{Goto}(I_4, r) \\
&= \{S \rightarrow sr\cdot, \$\}
\end{aligned}$$

$$\begin{aligned}
I_{15} &= \text{Goto}(I_5, M) \\
&= \{L \rightarrow aM\cdot b, t/p\}
\end{aligned}$$

$$\begin{aligned}
I_{16} &= \text{Goto}(I_5, t) \\
&= \{M \rightarrow t\cdot, b\}
\end{aligned}$$

$$\begin{aligned}
I_{17} &= \text{Goto}(I_{10}, r) \\
&= \{S \rightarrow qLr\cdot, \$\}
\end{aligned}$$

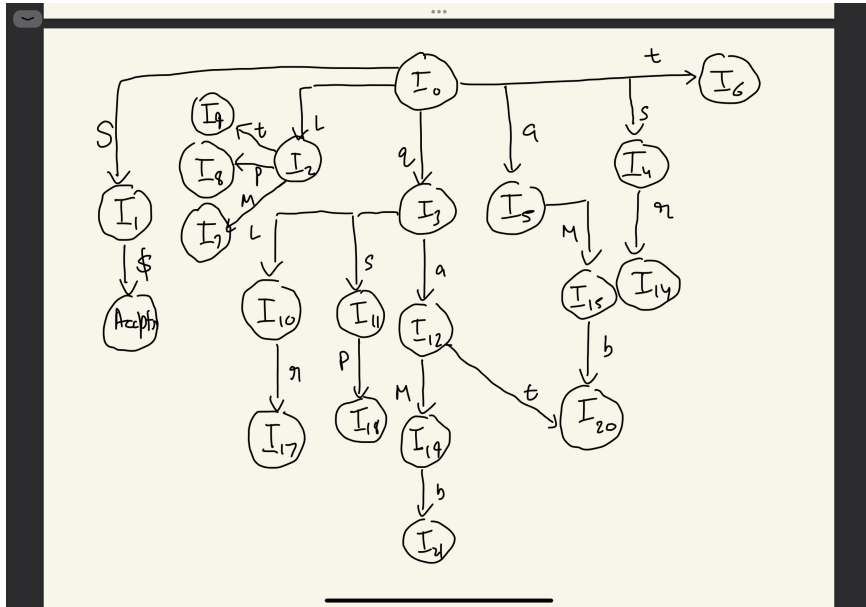


Figure 2: LR(1) automaton

$$\begin{aligned} I_{18} &= \text{Goto}(I_{11}, p) \\ &= \{S \rightarrow qsp\cdot, \$\} \end{aligned}$$

$$\begin{aligned} I_{19} &= \text{Goto}(I_{12}, M) \\ &= \{L \rightarrow aM \cdot b, r\} \end{aligned}$$

$$\begin{aligned} I_{20} &= \text{Goto}(I_{15}, b) \\ &= \{L \rightarrow aMb\cdot, t/p\} \end{aligned}$$

$$\begin{aligned} I_{21} &= \text{Goto}(I_{19}, b) \\ &= \{L \rightarrow aMb\cdot, r\} \end{aligned}$$

$$I_{16} = \text{Goto}(I_{12}, t)$$

2.2.2 LR(1) automaton

2.2.3 LALR collection

Merging sets of LR(1) items that have the same core:

$$\begin{aligned} I_{5,12} &= \text{Goto}(I_0, a) \\ &= \{L \rightarrow a \cdot Mb, t/p/r, \\ &\quad M \rightarrow \cdot t, b\} \end{aligned}$$

$$\begin{aligned} I_{6,13} &= \text{Goto}(I_0, t) \\ &= \{L \rightarrow t \cdot, t/p/r\} \end{aligned}$$

$$\begin{aligned} I_{9,16} &= \text{Goto}(I_2, t) \\ &= \{M \rightarrow t \cdot, \$ / b\} \end{aligned}$$

$$\begin{aligned} I_{15,19} &= \text{Goto}(I_{5,12}, M) \\ &= \{L \rightarrow aM \cdot b, t/p/r\} \end{aligned}$$

$$\begin{aligned} I_{20,21} &= \text{Goto}(I_{15,19}, b) \\ &= \{L \rightarrow aMb \cdot, t/p/r\} \end{aligned}$$

So, there will be 17 states ($I_0, I_1, I_2, I_3, I_4, I_{5,12}, I_{6,13}, I_7, I_8, I_{9,16}, I_{10}, I_{11}, I_{14}, I_{15,19}, I_{17}, I_{18}, I_{20,21}$) in the LALR collection.

2.2.4 LALR(1) parsing table

State	ACTION								GOTO		
	a	b	p	q	r	s	t	\$	S	L	M
0	s5,12			s3		s4	s6		1	2	
1								acc			
2			s8				s9,16				7
3	s5,12					s11	s6,13			10	
4			r7		s14		r7				
5,12							s9,16				15,19
6,13			r8		r8		r8				
7							r1				
8							r2				
9,16		r9						r9			
10					s17						
11			s18		r7						
14								r4			
15,19		s20,12									
17								r3			
18								r5			
20,21			r6		r6		r6				

2.2.5 Conclusion

As there are no conflicts in the LALR(1) parsing table, the given CFG is LALR(1).

3 Problem 3

3.1 Compilation commands

```
bison -d prob3.y
flex prob3.l
gcc -o prob3 lex.yy.c prob3.tab.c
```

3.2 Execution instructions

Let's say the input file is `testcase.txt`. Run `prob3` using the command:
`./prob3 < testcase.txt`

3.3 Other details

- Stray opening/closing tag is reported as syntax error along with the token having the stray tag.