CS335 : Assignment 2

Ankush Singh <anksingh@iitk.ac.in>
150107

12th February 2020

Answer 1

Original grammer is

$$S \to (L) \mid a$$
$$L \to L, S \mid LS \mid b$$

After removing left recursion, grammer becomes

$$S \to (L)|a$$
$$L \to bL'$$
$$L' \to , SL' \mid SL' \mid \epsilon$$

We don't need to left factor the grammer. The First and the Follow Sets for the nonterminals are:

Nonterminal	First Set	Follow Set
S	{ (,a }	{ \$,(,,,a }
L	{ b }	{) }
L'	{ (,,,ϵ, a}	{) }

We can find the predictive parser table for the grammar from the First and Follow sets as:

Nonterminal	a	b	(	)	,	\$
S	S $\to a$		S $\to (L)$			
L		L $\to bL'$				
L'	L' $\to SL'$		L' $\to SL'$	L' $\to \epsilon$	L' $\to , SL'$	

Since, there are no multiple entries in the table the grammar is unambiguous and hence it is LL(1). We can now design our parser easily using the parse-table.

Answer 2

Given grammar is,

$$S \to Lp \mid qLr \mid sr \mid qsp$$
$$L \to s$$

Augmented grammar with rule numbering is :

$$S' \to S$$
$$rule\#1 : S \to Lp$$
$$rule\#2 : S \to qLr$$
$$rule\#3 : S \to sr$$
$$rule\#4 : S \to qsp$$
$$rule\#5 : L \to s$$

Calculating the first and the follow sets as:

Nonterminal	First Set	Follow Set
S	{ s,q }	{ $ }
L	{ s }	{ p,r }

We would first generate the SLR parsing table for this grammar:
Computing the canonical collection of LR(0) items (or states):

$$I_o = Closure(S' \to .S) = \{$$
$$S' \to .S$$
$$S \to .Lp$$
$$S \to .qLr$$
$$S \to .sr$$
$$S \to .qsp$$
$$L \to .s$$
$$\}$$
$$I_1 = Goto(I_o, S) = \{$$
$$S' \to S.$$
$$\}$$
$$I_2 = Goto(I_o, L) = \{$$
$$S \to L.p$$
$$\}$$
$$I_3 = Goto(I_o, q) = \{$$
$$S \to q.Lr$$
$$L \to .s$$
$$S \to q.sp$$
$$\}$$
$$I_4 = Goto(I_o, s) = \{$$
$$S \to s.r$$
$$L \to s.$$
$$\}$$
$$I_5 = Goto(I_2, p) = \{$$
$$S \to Lp.$$
$$\}$$
$$I_6 = Goto(I_3, L) = \{$$
$$S \to qL.r$$
$$\}$$
$$I_7 = Goto(I_3, s) = \{$$
$$L \to s.$$
$$S \to qs.p$$
$$\}$$
$$I_8 = Goto(I_4, r) = \{$$
$$S \to sr.$$
$$\}$$
$$I_9 = Goto(I_6, r) = \{$$
$$S \to qLr.\}$$
$$I_{10} = Goto(I_7, p) = \{$$
$$S \to qsp.$$
$$\}$$

From the states we have the following Actions :

Action[I_o, q]=shift to I_3
Action[I_o, s]=shift to I_4
Action[I_2, p]=shift to I_5

Action$[I_3, s]$=shift to I_7
Action$[I_4, r]$=shift to I_8
Action$[I_6, r]$=shift to I_9
Action$[I_7, p]$=shift to I_{10}
Action$[I_4, p]$= reduce by rule 5: L→ s
Action$[I_4, r]$= reduce by rule 5: L→ s
Action$[I_5, \$]$= reduce by rule 1: S→ Lp
Action$[I_7, p]$= reduce by rule 5: L→ s
Action$[I_7, r]$= reduce by rule 5: L→ s
Action$[I_8, \$]$= reduce by rule 3: S→ sr
Action$[I_9, \$]$= reduce by rule 2: S→ qLr
Action$[I_{10}, \$]$= reduce by rule 4: S→ qsp And the following Gotos:

Goto$[0, S]$=1
Goto$[0, L]$=2
Goto$[3, L]$=6

The SLR parsing table we have found as:

	Action					**Goto**	
state	p	q	r	s	$	S	L
0		s3		s4		1	2
1					acc		
2	s5						
3				s7			6
4	r5		r5,s8				
5					r1		
6			s9				
7	r5,s10		r5				
8					r3		
9					r2		
10					r4		

We observe that there are multiple entries in the SLR parsing table which proves that grammar is ambiguous and not SLR(1).
Now, let us make the LALR parsing table:
First, we would be computing the canonical collection of LR(1) items:

$$
\begin{gathered}
I_o = Closure(S' \rightarrow .S, \$) = \{ \\
S' \rightarrow .S, \$ \\
S \rightarrow .Lp, \$ \\
S \rightarrow .qLr, \$ \\
S \rightarrow .sr, \$ \\
S \rightarrow .qsp, \$ \\
L \rightarrow .s, p \\
\} \\
I_1 = Goto(I_o, S) = \{ \\
S' \rightarrow S., \$ \\
\} \\
I_2 = Goto(I_o, L) = \{ \\
S \rightarrow L.p, \$ \\
\} \\
I_3 = Goto(I_o, q) = \{ \\
S \rightarrow q.Lr, \$ \\
L \rightarrow .s, r \\
S \rightarrow q.sp, \$ \\
\} \\
I_4 = Goto(I_o, s) = \{ \\
S \rightarrow s.r, \$ \\
L \rightarrow s., p \\
\} \\
I_5 = Goto(I_2, p) = \{ \\
S \rightarrow Lp., \$ \\
\} \\
I_6 = Goto(I_3, L) = \{ \\
S \rightarrow qL.r, \$ \\
\} \\
I_7 = Goto(I_3, s) = \{ \\
L \rightarrow s., r \\
S \rightarrow qs.p, \$ \\
\} \\
I_8 = Goto(I_4, r) = \{ \\
S \rightarrow sr., \$ \\
\} \\
I_9 = Goto(I_6, r) = \{ \\
S \rightarrow qLr., \$ \\
\} \\
I_{10} = Goto(I_7, p) = \{ \\
S \rightarrow qsp., \$ \\
\}
\end{gathered}
$$

From the states, we have the following actions:

Action$[I_o, q]$=shift to I_3
Action$[I_o, s]$=shift to I_4
Action$[I_2, p]$=shift to I_5
Action$[I_3, s]$=shift to I_7
Action$[I_4, r]$=shift to I_8

$$\text{Action}[I_6, r]\text{=shift to } I_9$$
$$\text{Action}[I_7, p]\text{=shift to } I_{10}$$
$$\text{Action}[I_4, p]\text{= reduce by rule 5: L}\rightarrow s$$
$$\text{Action}[I_5, \$]\text{= reduce by rule 1: S}\rightarrow Lp$$
$$\text{Action}[I_7, r]\text{= reduce by rule 5: L}\rightarrow s$$
$$\text{Action}[I_8, \$]\text{= reduce by rule 3: S}\rightarrow sr$$
$$\text{Action}[I_9, \$]\text{= reduce by rule 2: S}\rightarrow qLr$$
$$\text{Action}[I_{10}, \$]\text{= reduce by rule 4: S}\rightarrow qsp$$

And the following Gotos:

$$\text{Goto}[0, S]\text{=1}$$
$$\text{Goto}[0, L]\text{=2}$$
$$\text{Goto}[3, L]\text{=6}$$

The LALR parsing table we get is:

	Action					**Goto**	
state	p	q	r	s	$	S	L
0		s3		s4		1	2
1					acc		
2	s5						
3				s7			6
4	r5		s8				
5					r1		
6			s9				
7	s10		r5				
8					r3		
9					r2		
10					r4		

We observed that the LALR parsing table has no multiple entries, implying the grammar is unambiguous and is LALR(1).
Hence, the given grammar is LALR(1) and not SLR(1).

Answer 3

Given grammar is

$$R' \rightarrow R$$
$$rule\#1 : R \rightarrow R|R$$
$$rule\#2 : R \rightarrow RR$$
$$rule\#3 : R \rightarrow R*$$
$$rule\#4 : R \rightarrow (R)$$
$$rule\#5 : R \rightarrow a$$
$$rule\#6 : R \rightarrow b$$

First (R) =(,a,b
Follow(R) ={*,|,(,),a,b,$}
Computing the canonical collection of LR(0) items (or the states):

$$I_o = Closure(R' \to .R) = \{$$
$$R' \to .R$$
$$R \to .R|R$$
$$R \to .RR$$
$$R \to .R*$$
$$R \to .(R)$$
$$R \to .a$$
$$R \to .b$$
$$\}$$
$$I_1 = Goto(I_o, R) = \{$$
$$R' \to R.$$
$$R \to R.|R$$
$$R \to R.R$$
$$R \to R.*$$
$$R \to .R|R$$
$$R \to .RR$$
$$R \to .R*$$
$$R \to .(R)$$
$$R \to .a$$
$$R \to .b$$
$$\}$$
$$I_2 = Goto(I_o, () = \{$$
$$R \to (.R)$$
$$R \to .R|R$$
$$R \to .RR$$
$$R \to .R*$$
$$R \to .(R)$$
$$R \to .a$$
$$R \to .b$$
$$\}$$
$$I_3 = Goto(I_o, a) = \{$$
$$R \to a.$$
$$\}$$
$$I_4 = Goto(I_o, b) = \{$$
$$R \to b.$$
$$\}$$
$$I_5 = Goto(I_1, |) = \{$$
$$R \to R|.R$$
$$R \to .R|R$$
$$R \to .RR$$
$$R \to .R*$$
$$R \to .(R)$$
$$R \to .a$$
$$R \to .b$$
$$\}$$

$$I_6 = Goto(I_1, R) = \{$$
$$R \rightarrow RR.$$
$$R \rightarrow R.|R$$
$$R \rightarrow R.R$$
$$R \rightarrow R.*$$
$$R \rightarrow .R|R$$
$$R \rightarrow .RR$$
$$R \rightarrow .R*$$
$$R \rightarrow .(R)$$
$$R \rightarrow .a$$
$$R \rightarrow .b$$
$$\}$$
$$I_7 = Goto(I_1, *) = \{$$
$$R \rightarrow R * .$$
$$\}$$
$$I_8 = Goto(I_2, R) = \{$$
$$R \rightarrow (R.)$$
$$R \rightarrow R.|R$$
$$R \rightarrow R.R$$
$$R \rightarrow R.*$$
$$R \rightarrow .R|R$$
$$R \rightarrow .RR$$
$$R \rightarrow .R*$$
$$R \rightarrow .(R)$$
$$R \rightarrow .a$$
$$R \rightarrow .b$$
$$\}$$
$$I_9 = Goto(I_5, R) = \{$$
$$R \rightarrow R|R.$$
$$R \rightarrow R.|R$$
$$R \rightarrow R.R$$
$$R \rightarrow R.*$$
$$R \rightarrow .R|R$$
$$R \rightarrow .RR$$
$$R \rightarrow .R*$$
$$R \rightarrow .(R)$$
$$R \rightarrow .a$$
$$R \rightarrow .b$$
$$\}$$
$$I_{10} = Goto(I_8,)) = \{$$
$$R \rightarrow (R).$$
$$\}$$

We have the following actions defined :

Action$[I_o, (]$=shift to I_2
Action$[I_o, a]$=shift to I_3
Action$[I_o, b]$=shift to I_4
Action$[I_1, |]$=shift to I_5
Action$[I_1, *]$=shift to I_7
Action$[I_8,)]$=shift to I_{10}
Action$[I_3, |]$= reduce by rule 5: R$\rightarrow a$

Action$[I_3, *]$= reduce by rule 5: R→ a
Action$[I_3, (]$= reduce by rule 5: R→ a
Action$[I_3,)]$= reduce by rule 5: R→ a
Action$[I_3, a]$= reduce by rule 5: R→ a
Action$[I_3, b]$= reduce by rule 5: R→ a
Action$[I_3, \$]$= reduce by rule 5: R→ a
Action$[I_4, |]$= reduce by rule 6: R→ b
Action$[I_4, *]$= reduce by rule 6: R→ b
Action$[I_4, (]$= reduce by rule 6: R→ b
Action$[I_4,)]$= reduce by rule 6: R→ b
Action$[I_4, a]$= reduce by rule 6: R→ b
Action$[I_4, b]$= reduce by rule 6: R→ b
Action$[I_4, \$]$= reduce by rule 6: R→ b
Action$[I_6, |]$= reduce by rule 2: R→ RR
Action$[I_6, *]$= reduce by rule 2: R→ RR
Action$[I_6, (]$= reduce by rule 2: R→ RR
Action$[I_6,)]$= reduce by rule 2: R→ RR
Action$[I_6, a]$= reduce by rule 2: R→ RR
Action$[I_6, b]$= reduce by rule 2: R→ RR
Action$[I_6, \$]$= reduce by rule 2: R→ RR
Action$[I_7, |]$= reduce by rule 3: R→ $R*$
Action$[I_7, *]$= reduce by rule 3: R→ $R*$
Action$[I_7, a]$= reduce by rule 3: R→ $R*$
Action$[I_7, b]$= reduce by rule 3: R→ $R*$
Action$[I_7, (]$= reduce by rule 3: R→ $R*$
Action$[I_7,)]$= reduce by rule 3: R→ $R*$
Action$[I_7, \$]$= reduce by rule 3: R→ $R*$
Action$[I_9, |]$= reduce by rule 1: R→ $R|R$
Action$[I_9, *]$= reduce by rule 1: R→ $R|R$
Action$[I_9, (]$= reduce by rule 1: R→ $R|R$
Action$[I_9,)]$= reduce by rule 1: R→ $R|R$
Action$[I_9, a]$= reduce by rule 1: R→ $R|R$
Action$[I_9, b]$= reduce by rule 1: R→ $R|R$
Action$[I_9, \$]$= reduce by rule 1: R→ $R|R$
Action$[I_{10}, |]$= reduce by rule 4: R→ (R)
Action$[I_{10}, *]$= reduce by rule 4: R→ (R)
Action$[I_{10}, (]$= reduce by rule 4: R→ (R)
Action$[I_{10},)]$= reduce by rule 4: R→ (R)
Action$[I_{10}, a]$= reduce by rule 4: R→ (R)
Action$[I_{10}, b]$= reduce by rule 4: R→ (R)
Action$[I_{10}, \$]$= reduce by rule 4: R→ (R)

We have the following Gotos derived:

Goto$[0, R]$=1
Goto$[2, R]$=8
Goto$[5, R]$=9
Goto$[1, R]$=6
Goto$[1, (]$=2
Goto$[1, a]$=3
Goto$[1, b]$=4
Goto$[2, (]$=2
Goto$[2, a]$=3
Goto$[2, b]$=4
Goto$[5, (]$=2
Goto$[5, a]$=3
Goto$[5, b]$=4
Goto$[6, R]$=6
Goto$[6, |]$=5
Goto$[6, *]$=7
Goto$[6, (]$=2
Goto$[6, a]$=3
Goto$[6, b]$=4
Goto$[8, R]$=6

Goto[8, |]=5
Goto[8, (]=2
Goto[8, a]=3
Goto[8, b]=4
Goto[9, R]=6
Goto[9, |]=5
Goto[9, (]=2
Goto[9, a]=3
Goto[9, b]=4

The SLR parsing table that we get is :

	Action							**Goto**
state	a	b	(	)	*	|	$	R
0	s3	s4	s2					1
1	s3	s4	s2		s7	s5	acc	6
2	s3	s4	s2					8
3	r5	r5	r5	r5	r5	r5	r5	
4	r6	r6	r6	r6	r6	r6	r6	
5	s3	s4	s2					9
6	s3,r2	s4,r2	s2,r2	r2	s7,r2	s5,r2	r2	6
7	r3	r3	r3	r3	r3	r3	r3	
8	s3	s4	s2	s10	s7	s5		6
9	s3 ,r1	s4,r1	r1,s2	r1	s7,r1	s5,r1	r1	6
10	r4	r4	r4	r4	r4	r4	r4	

Total number of shift-reduce conflicts are **10**
Conflicting Actions are :

1 Action[I_6, a]= shift to I_3 for R→ $.a$
or reduce by rule # 2 R→ $RR.$
2 Action[I_6, b]= shift to I_4 for R→ $.b$
or reduce by rule # 2 R→ $RR.$
3 Action[I_6, (]= shift to I_2 for R→ $.(R)$
or reduce by rule # 2 R→ $RR.$
4 Action[$I_6, *$]= shift to I_7 for R→ $R.*$
or reduce by rule # 2 R→ $RR.$
5 Action[I_6, |]= shift to I_5 for R→ $R.|R$
or reduce by rule # 2 R→ $RR.$
6 Action[I_9, a]= shift to I_3 for R→ $.a$
or reduce by rule # 1 R→ $R|R.$
7 Action[I_9, b]= shift to I_4 for R→ $.b$
or reduce by rule # 1 R→ $R|R.$
8 Action[I_9, (]= shift to I_2 for R→ $.(R)$
or reduce by rule # 1 R→ $R|R.$
9 Action[$I_9, *$]= shift to I_7 for R→ $R.*$
or reduce by rule # 1 R→ $R|R.$
10 Action[I_9, |]= shift to I_5 for R→ $R.|R$
or reduce by rule # 1 R→ $R|R.$

Since we have to resolve the conflicts such that all the regular expressions are resolved properly, we would be assuming only regular language input. The priority of the regular expressions is() $\rightarrow * \rightarrow RR \rightarrow |$.
So, we would be resolving the conflicts accordingly.

Disambiguation Rules

Conflict	Resolved To
Action$[I_6, a]$= shift to I_3 for R→ $.a$ or reduce by rule # 2 R→ $RR.$	reduce by rule 2 R→ RR
Action$[I_6, b]$= shift to I_4 for R→ $.b$ or reduce by rule 2 R→ $RR.$	reduce by rule 2 R→ RR
Action$[I_6, (]$= shift to I_2 for R→ $.(R)$ or reduce by rule 2 R→ $RR.$	shift to I_2 for R→ $.(R)$
Action$[I_6, *]$= shift to I_7 for R→ $R.*$ or reduce by rule 2 R→ $RR.$	shift to I_7 for R→ $R.*$
Action$[I_6, \|]$= shift to I_5 for R→ $R.\|R$ or reduce by rule 2 R→ $RR.$	reduce by rule 2 R→ RR
Action$[I_9, a]$= shift to I_3 for R→ $.a$ or reduce by rule 1 R→ $R\|R.$	shift to I_3 for R→ $.a$
Action$[I_9, b]$= shift to I_4 for R→ $.b$ or reduce by rule 1 R→ $R\|R.$	shift to I_4 for R→ $.b$
Action$[I_9, (]$= shift to I_2 for R→ $.(R)$ or reduce by rule 1 R→ $R\|R.$	shift to I_2 for R→ $.(R)$
Action$[I_9, *]$= shift to I_7 for R→ $R.*$ or reduce by rule 1 R→ $R\|R.$	shift to I_7 for R→ $R.*$
Action$[I_9, \|]$= shift to I_5 for R→ $R.\|R$ or reduce by rule 1 R→ $R\|R.$	reduce by rule 1 R→ $R.\|R$

Answer 4

I am using ply for building the parser. I have defined a grammar for the language of the dissertation. The grammar is shown on the next page

All the code for the fourth part is under main folder [**assign2**] of the submission directory. Just go in that directory.
The files lex.py and yacc.py as always are part of the PLY toolkit.
There is a lexer file lexer.py which contains the syntactic defintion of this structure. The parser is in file parser.py which implements the grammar of the structure.
RUN instructions:
for printing the result, type this command :
python parser.py samplethesis.txt
or
python3 parser.py samplethesis.txt

Grammar

$$
\begin{gathered}
S' \rightarrow start \\
start \rightarrow thesis \\
newlines \rightarrow newline\ newlines \\
newlines \rightarrow newline \\
paragraph \rightarrow statements \\
paragraphs \rightarrow paragraph\ newlines\ paragraphs \\
paragraphs \rightarrow paragraph\ newlines \\
paragraphs \rightarrow paragraph \\
statements \rightarrow statement\ statements \\
statements \rightarrow statement \\
statement \rightarrow sentence\ excl \\
statement \rightarrow sentence\ dot \\
statement \rightarrow sentence\ qm \\
sentence \rightarrow sentenceword\ comma\ sentence \\
sentence \rightarrow sentenceword\ semicolon\ sentence \\
sentence \rightarrow sentenceword\ sentence \\
sentence \rightarrow sentenceword \\
sentenceword \rightarrow words \\
sentenceword \rightarrow numbers \\
sections \rightarrow section\ sections \\
sections \rightarrow section \\
section \rightarrow Section\ numbers\ colon\ headings\ paragraphs \\
headings \rightarrow headingword\ headings \\
headings \rightarrow headingword\ newlines \\
headingword \rightarrow words \\
headingword \rightarrow numbers \\
chapters \rightarrow chapter\ chapters \\
chapters \rightarrow chapter \\
chapter \rightarrow chapter1 \\
chapter \rightarrow chapter2 \\
chapter1 \rightarrow Chapter\ numbers\ colon\ headings\ paragraphs \\
chapter1 \rightarrow Chapter\ numbers\ colon\ headings\ paragraphs\ sections \\
chapter2 \rightarrow Chapter\ numbers\ colon\ headings\ sections \\
thesis \rightarrow Title\ colon\ headings\ chapters
\end{gathered}
$$