

Compiler Construction 2016

— Series 5 —

Hand in until June 14th before the exercise class.



Credit: <https://xkcd.com/1443/>

General Remarks

- It is allowed to hand in your solutions for the theoretical part via email as a **separately attached PDF file**.
- Please hand in your solutions in groups of 3 or 4.
- If you hand in solutions via email please make sure that the size is reasonable (< 1 MB). Solutions that our printer rejects to print within 30min due to large scanned images will not be corrected.

Exercise 1

(2 Points)

A language $L \in \Sigma^*$ is called prefix-free, if $L \cap L\Sigma^+ = \emptyset$, i.e. if no proper prefix of a word in L is in L , too.

Show that the following holds for all non prefix-free languages L : $L \notin \mathcal{L}(LR(0))$.

Solution: _____

Towards a contradiction assume L is not prefix-free, but $L \in LR(0)$. Then there exists $v \in \Sigma^*$ and $w \in \Sigma^+$ such that $v \in L$ and $vw \in L$.

Since $L \in LR(0)$ we know by Theorem 9.10 that there exists a deterministic bottom-up parsing automaton accepting L .

Consider its run on input word v :

$$(v, I_0, \varepsilon) \vdash^* (\varepsilon, I_0 I, z) \vdash (\varepsilon, \varepsilon, z0), \quad \text{where } act(I) = \text{accept}$$

Furthermore, consider the run on input word vw :

$$(vw, I_0, \varepsilon) \vdash^* (w, I_0 I, z) \vdash^* (\varepsilon, I_0 I', zz') \vdash (\varepsilon, \varepsilon, zz'0), \quad \text{where } act(I') = \text{accept}$$

However, this means that $act(I) \neq \text{accept}$ which is a contradiction to the run on input word v .

Exercise 2

(2 Points)

- (a) Show that there exists an $LR(0)$ grammar that is not an $LL(1)$ grammar.
- (b) Show that there are regular languages for which no $LR(0)$ grammars exist.

Solution: _____

- (a) Consider the following grammar G :

$$\begin{aligned} S &\rightarrow A \mid B \\ A &\rightarrow aAb \mid ab \\ B &\rightarrow aaBc \mid aac \end{aligned}$$

This grammar is not in $LL(1)$, because $a \in la(S \rightarrow A) \cap la(S \rightarrow B)$. It is, however, an $LR(0)$ grammar. To show this, we show that each item set $I \in LR(0)(G)$ contains no conflicting items.

$$\begin{aligned} I_0 &:= LR(0)(\varepsilon) : [S' \rightarrow \cdot S] [S \rightarrow \cdot A] [S \rightarrow \cdot B] \\ &\quad [A \rightarrow \cdot aAb] [A \rightarrow \cdot ab] [B \rightarrow \cdot aaBc] \\ &\quad [B \rightarrow \cdot aac] \\ I_1 &:= LR(0)(S) : [S' \rightarrow S \cdot] \\ I_2 &:= LR(0)(A) : [S \rightarrow A \cdot] \\ I_3 &:= LR(0)(B) : [S \rightarrow B \cdot] \\ I_4 &:= LR(0)(a) : [A \rightarrow a \cdot Ab] [A \rightarrow \cdot aAb] [A \rightarrow \cdot ab] \\ &\quad [A \rightarrow a \cdot b] [B \rightarrow a \cdot aBc] [B \rightarrow a \cdot ac] \\ I_5 &:= LR(0)(aa) : [A \rightarrow a \cdot Ab] [A \rightarrow \cdot aAb] [A \rightarrow \cdot ab] \\ &\quad [A \rightarrow a \cdot b] [B \rightarrow aa \cdot Bc] [B \rightarrow \cdot aaBc] \\ &\quad [B \rightarrow \cdot aac] [B \rightarrow aa \cdot c] \\ I_6 &:= LR(0)(ab) : [A \rightarrow ab \cdot] \\ I_7 &:= LR(0)(aA) : [A \rightarrow aA \cdot b] \\ I_8 &:= LR(0)(aac) : [B \rightarrow aac \cdot] \\ I_9 &:= LR(0)(aaB) : [B \rightarrow aaB \cdot c] \\ I_{10} &:= LR(0)(aAb) : [A \rightarrow aAb \cdot] \\ I_{11} &:= LR(0)(aaBc) : [B \rightarrow aaBc \cdot] \end{aligned}$$

Blank cells in the following table lead to the empty item set.

goto	S	A	B	a	b	c
I_0	I_1	I_2	I_3	I_4		
I_1						
I_2						
I_3						
I_4		I_7		I_5	I_6	
I_5		I_7	I_9	I_4	I_6	I_8
I_6						
I_7					I_{10}	
I_8						
I_9						I_{11}
I_{10}						
I_{11}						

By inspection of each item set I_i , $1 \leq i \leq 11$, it follows that no item set contains conflicting items. Hence, G is an $LR(0)$ grammar.

- (b) By Exercise 1 we know that each language which is recognized by an $LR(0)$ grammar is prefix-free. Thus, any regular non prefix-free language cannot be recognized by an $LR(0)$ grammar, e.g. $L = \{a, ab\}$.

Exercise 3

(2 Points)

Consider the following grammar G :

$$\begin{aligned} S &\rightarrow A_1b_1 \mid A_2b_2 \mid A_3b_3 \\ A_1 &\rightarrow a_2A_1 \mid a_3A_1 \mid a_2 \mid a_3 \\ A_2 &\rightarrow a_1A_2 \mid a_3A_2 \mid a_1 \mid a_3 \\ A_3 &\rightarrow a_1A_3 \mid a_2A_3 \mid a_1 \mid a_2 \end{aligned}$$

- (a) Compute all $LR(0)$ sets of G .
(b) Prove or disprove: G is an $SLR(1)$ grammar.

Solution: _____

- (a) The $LR(0)$ sets:

$$\begin{aligned} I_0 &:= LR(0)(\varepsilon) : [S' \rightarrow \cdot S] [S \rightarrow \cdot A_1b_1] [S \rightarrow \cdot A_2b_2] [S \rightarrow \cdot A_3b_3] \\ &\quad [A_1 \rightarrow \cdot a_2A_1] [A_1 \rightarrow \cdot a_3A_1] [A_1 \rightarrow \cdot a_2] [A_1 \rightarrow \cdot a_3] \\ &\quad [A_2 \rightarrow \cdot a_1A_2] [A_2 \rightarrow \cdot a_3A_2] [A_2 \rightarrow \cdot a_1] [A_2 \rightarrow \cdot a_3] \\ &\quad [A_3 \rightarrow \cdot a_1A_3] [A_3 \rightarrow \cdot a_2A_3] [A_3 \rightarrow \cdot a_1] [A_3 \rightarrow \cdot a_2] \\ I_1 &:= LR(0)(S) : [S' \rightarrow S \cdot] \\ I_2 &:= LR(0)(A_1) : [S \rightarrow A_1 \cdot b_1] \\ I_3 &:= LR(0)(A_2) : [S \rightarrow A_2 \cdot b_2] \\ I_4 &:= LR(0)(A_3) : [S \rightarrow A_3 \cdot b_3] \\ I_5 &:= LR(0)(a_1) : [A_2 \rightarrow a_1 \cdot A_2] [A_2 \rightarrow \cdot a_1A_2] [A_2 \rightarrow \cdot a_3A_2] [A_2 \rightarrow \cdot a_1] [A_2 \rightarrow \cdot a_3] [A_2 \rightarrow a_1 \cdot] \\ &\quad [A_3 \rightarrow a_1 \cdot A_3] [A_3 \rightarrow \cdot a_1A_3] [A_3 \rightarrow \cdot a_2A_3] [A_3 \rightarrow \cdot a_1] [A_3 \rightarrow \cdot a_2] [A_3 \rightarrow a_1 \cdot] \\ I_6 &:= LR(0)(a_2) : [A_1 \rightarrow a_2 \cdot A_1] [A_1 \rightarrow \cdot a_2A_1] [A_1 \rightarrow \cdot a_3A_1] [A_1 \rightarrow \cdot a_2] [A_1 \rightarrow \cdot a_3] [A_1 \rightarrow a_2 \cdot] \\ &\quad [A_3 \rightarrow a_2 \cdot A_3] [A_3 \rightarrow \cdot a_1A_3] [A_3 \rightarrow \cdot a_2A_3] [A_3 \rightarrow \cdot a_1] [A_3 \rightarrow \cdot a_2] [A_3 \rightarrow a_2 \cdot] \\ I_7 &:= LR(0)(a_3) : [A_1 \rightarrow a_3 \cdot A_1] [A_1 \rightarrow \cdot a_2A_1] [A_1 \rightarrow \cdot a_3A_1] [A_1 \rightarrow \cdot a_2] [A_1 \rightarrow \cdot a_3] [A_1 \rightarrow a_3 \cdot] \\ &\quad [A_2 \rightarrow a_2 \cdot A_2] [A_2 \rightarrow \cdot a_1A_2] [A_2 \rightarrow \cdot a_3A_2] [A_2 \rightarrow \cdot a_1] [A_2 \rightarrow \cdot a_3] [A_2 \rightarrow a_3 \cdot] \\ I_8 &:= LR(0)(A_1b_1) : [S \rightarrow A_1b_1 \cdot] \\ I_9 &:= LR(0)(A_2b_2) : [S \rightarrow A_2b_2 \cdot] \\ I_{10} &:= LR(0)(A_3b_3) : [S \rightarrow A_3b_3 \cdot] \\ I_{11} &:= LR(0)(a_1A_2) : [A_2 \rightarrow a_1A_2 \cdot] \\ I_{12} &:= LR(0)(a_1A_3) : [A_3 \rightarrow a_1A_3 \cdot] \\ I_{13} &:= LR(0)(a_1a_2) : [A_3 \rightarrow a_2 \cdot A_3] [A_3 \rightarrow \cdot a_1A_3] [A_3 \rightarrow \cdot a_2A_3] [A_3 \rightarrow \cdot a_1] [A_3 \rightarrow \cdot a_2] [A_3 \rightarrow a_2 \cdot] \\ I_{14} &:= LR(0)(a_1a_3) : [A_2 \rightarrow a_3 \cdot A_2] [A_2 \rightarrow \cdot a_1A_2] [A_2 \rightarrow \cdot a_3A_2] [A_2 \rightarrow \cdot a_1] [A_2 \rightarrow \cdot a_3] [A_2 \rightarrow a_3 \cdot] \\ I_{15} &:= LR(0)(a_2A_1) : [A_1 \rightarrow a_2A_1 \cdot] \\ I_{16} &:= LR(0)(a_2A_3) : [A_3 \rightarrow a_2A_3 \cdot] \end{aligned}$$

$I_{17} := LR(0)(a_2a_1) : [A_3 \rightarrow a_1 \cdot A_3] [A_3 \rightarrow \cdot a_1 A_3] [A_3 \rightarrow \cdot a_2 A_3] [A_3 \rightarrow \cdot a_1] [A_3 \rightarrow \cdot a_2] [A_3 \rightarrow a_1 \cdot]$
 $I_{18} := LR(0)(a_2a_3) : [A_1 \rightarrow a_3 \cdot A_1] [A_1 \rightarrow \cdot a_2 A_1] [A_1 \rightarrow \cdot a_3 A_1] [A_1 \rightarrow \cdot a_2] [A_1 \rightarrow \cdot a_3] [A_1 \rightarrow a_3 \cdot]$
 $I_{19} := LR(0)(a_3A_1) : [A_1 \rightarrow a_3A_1 \cdot]$
 $I_{20} := LR(0)(a_3A_2) : [A_2 \rightarrow a_3A_2 \cdot]$
 $I_{21} := LR(0)(a_3a_1) : [A_2 \rightarrow a_1 \cdot A_2] [A_2 \rightarrow \cdot a_1 A_2] [A_2 \rightarrow \cdot a_3 A_2] [A_2 \rightarrow \cdot a_1] [A_2 \rightarrow \cdot a_3] [A_2 \rightarrow a_1 \cdot]$
 $I_{22} := LR(0)(a_3a_2) : [A_1 \rightarrow a_2 \cdot A_1] [A_1 \rightarrow \cdot a_2 A_1] [A_1 \rightarrow \cdot a_3 A_1] [A_1 \rightarrow \cdot a_2] [A_1 \rightarrow \cdot a_3] [A_1 \rightarrow a_2 \cdot]$

goto	S	A_1	A_2	A_3	a_1	a_2	a_3	b_1	b_2	b_3
I_0	I_1	I_2	I_3	I_4	I_5	I_6	I_7			
I_1										
I_2								I_8		
I_3									I_9	
I_4										I_{10}
I_5			I_{11}	I_{12}	I_5	I_{13}	I_{14}			
I_6		I_{15}		I_{16}	I_{17}	I_6	I_{18}			
I_7		I_{19}	I_{20}		I_{21}	I_{22}	I_7			
I_8										
I_9										
I_{10}										
I_{11}										
I_{12}										
I_{13}				I_{16}	I_{17}	I_{13}				
I_{14}			I_{20}		I_{21}		I_{14}			
I_{15}										
I_{16}										
I_{17}				I_{12}	I_{17}	I_{13}				
I_{18}		I_{19}				I_{22}	I_{18}			
I_{19}										
I_{20}										
I_{21}			I_{11}		I_{21}		I_{14}			
I_{22}	I_{15}					I_{22}	I_{18}			

- (b) We first note that $fo(S) = \{\varepsilon\}$ and, for each $1 \leq i \leq 3$, we have $fo(A_i) = \{b_i\}$. Then the action function of G is given by (errors are omitted as in the lecture):

act	a_1	a_2	a_3	b_1	b_2	b_3	ε
I_0	shift	shift	shift				
I_1							accept
I_2				shift			
I_3					shift		
I_4						shift	
I_5	shift	shift	shift		red 10	red 14	
I_6	shift	shift	shift	red 6		red 15	
I_7	shift	shift	shift	red 7	red 11		
I_8							red 1
I_9							red 2
I_{10}							red 3
I_{11}					red 8		
I_{12}						red 12	
I_{13}	shift	shift				red 15	
I_{14}	shift		shift		red 7		
I_{15}				red 4			
I_{16}						red 13	
I_{17}	shift	shift			red 14		
I_{18}		shift	shift		red 7		
I_{19}				red 5			
I_{20}					red 9		
I_{21}	shift		shift		red 10		
I_{22}		shift	shift	red 6			

Since the action function is well-defined, $G \in SLR(1)$.

Exercise 4

(4 Points)

Consider the grammar $G = (N, \Sigma, P, S')$ covering some boolean expressions:

- $N := \{S', S\}$
- $\Sigma := \{true, false, \wedge, \neg, (,)\}$
- $S' \rightarrow S$
- $S \rightarrow (S \wedge S) \mid \neg S \mid true \mid false$

- Compute all $LR(0)$ sets of G .
- Specify the (deterministic) $LR(0)$ parsing automaton of G . Especially specify the parsing table. (Do not forget to give a numbering to the grammar rules.)
- Provide a run of the automaton on the input $((\neg true \wedge false) \wedge false)$.

Solution:

0. $S' \rightarrow S$
1. $S \rightarrow (S \wedge S)$
2. $S \rightarrow \neg S$
3. $S \rightarrow true$
4. $S \rightarrow false$

- The $LR(0)$ sets:

$$\begin{aligned}
 I_0 &:= LR(0)(\varepsilon) : & [S' \rightarrow \cdot S], [S \rightarrow \cdot (S \wedge S)], [S \rightarrow \cdot \neg S], [S \rightarrow \cdot true], [S \rightarrow \cdot false] \\
 I_1 &:= LR(0)(S) : & [S' \rightarrow S \cdot] \\
 I_2 &:= LR(0)(() : & [S \rightarrow (\cdot S \wedge S)], [S \rightarrow \cdot (S \wedge S)], [S \rightarrow \cdot \neg S], [S \rightarrow \cdot true], [S \rightarrow \cdot false]
 \end{aligned}$$

$I_3 := LR(0)(\neg) :$	$[S \rightarrow \neg \cdot S], [S \rightarrow \cdot (S \wedge S)], [S \rightarrow \cdot \neg S], [S \rightarrow \cdot true], [S \rightarrow \cdot false]$
$I_4 := LR(0)(true) :$	$[S \rightarrow true \cdot]$
$I_5 := LR(0)(false) :$	$[S \rightarrow false \cdot]$
$I_6 := LR(0)((S) :$	$[S \rightarrow (S \cdot \wedge S)]$
$I_7 := LR(0)(\neg S) :$	$[S \rightarrow \neg S \cdot]$
$I_8 := LR(0)((S \wedge) :$	$[S \rightarrow (S \wedge \cdot S)], [S \rightarrow \cdot (S \wedge S)], [S \rightarrow \cdot \neg S], [S \rightarrow \cdot true], [S \rightarrow \cdot false]$
$I_9 := LR(0)((S \wedge S) :$	$[S \rightarrow (S \wedge S \cdot)]$
$I_{10} := LR(0)((S \wedge S)) :$	$[S \rightarrow (S \wedge S) \cdot]$
$I_{11} := LR(0)(\gamma) :$	$\emptyset$ (for e.g. $\gamma = true\ S$)

(b) The (deterministic) $LR(0)$ parsing automaton of G :

- Input alphabet: Σ
- Pushdown alphabet: $\Gamma = \{I_0, \dots, I_{11}\}$
- Output alphabet: $\Delta = \{0, 1, 2, 3, 4, \text{error}\}$
- Configurations: $\Sigma^* \times \Gamma^* \times \Delta^*$
- Initial configuration: (w, I_0, ε) for $w \in \Sigma^*$
- Final configuration: $\{\varepsilon\} \times \{\varepsilon\} \times \Delta^*$
- Transitions: parsing table as follows (Empty entries stand for I_{11}).

LR(0)(G)	act	goto					
		S	(	$\neg$	true	false	$\wedge$)
I_0	shift	I_1	I_2	I_3	I_4	I_5	
I_1	accept						
I_2	shift	I_6	I_2	I_3	I_4	I_5	
I_3	shift	I_7	I_2	I_3	I_4	I_5	
I_4	red 3						
I_5	red 4						
I_6	shift						I_8
I_7	red 2						
I_8	shift	I_9	I_2	I_3	I_4	I_5	
I_9	shift						I_{10}
I_{10}	red 1						
I_{11}	error						

(c) An accepting run on the input $((\neg true \wedge false) \wedge false)$ is as follows:

(	$((\neg true \wedge false) \wedge false), I_0,$	ε	)
$\vdash$ (	$(\neg true \wedge false) \wedge false), I_0 I_2,$	ε	)
$\vdash$ (	$\neg true \wedge false) \wedge false), I_0 I_2 I_2,$	ε	)
$\vdash$ (	$true \wedge false) \wedge false), I_0 I_2 I_2 I_3,$	ε	)
$\vdash$ (	$\wedge false) \wedge false), I_0 I_2 I_2 I_3 I_4,$	ε	)
$\vdash$ (	$\wedge false) \wedge false), I_0 I_2 I_2 I_3 I_7,$	3	)
$\vdash$ (	$\wedge false) \wedge false), I_0 I_2 I_2 I_6,$	32	)
$\vdash$ (	$false) \wedge false), I_0 I_2 I_2 I_6 I_8,$	32	)
$\vdash$ (	) $\wedge false), I_0 I_2 I_2 I_6 I_8 I_5,$	32	)
$\vdash$ (	) $\wedge false), I_0 I_2 I_2 I_6 I_8 I_9,$	324	)
$\vdash$ (	$\wedge false), I_0 I_2 I_2 I_6 I_8 I_9 I_{10},$	324	)
$\vdash$ (	$\wedge false), I_0 I_2 I_6,$	3241	)
$\vdash$ (	$false), I_0 I_2 I_6 I_8,$	3241	)
$\vdash$ (	), $I_0 I_2 I_6 I_8 I_5,$	3241	)



$\vdash ($	$), I_0 I_2 I_6 I_8 I_9,$	32414	$)$
$\vdash ($	$\varepsilon, I_0 I_2 I_6 I_8 I_9 I_{10},$	32414	$)$
$\vdash ($	$\varepsilon, I_0 I_1,$	324141	$)$
$\vdash ($	$\varepsilon, \varepsilon,$	3241410	$)$