Parsing

Homework 9 (LR), due 22 June 2020

Laura Kallmeyer

SS 2020, Heinrich-Heine-Universität Düsseldorf

Question 1 (Canonical LR(1) construction and parse table)

1. *Consider the CFG G with non-terminals $N = \{S\}$, terminals $T = \{a, b\}$, start symbol S and productions 1. $S \to Sa$, 2. $S \to aSb$, 3. $S \to \varepsilon$.*

   Consider the canonical LR(1)-construction for this grammar.

   (a) *How does the state $q_0 = closure(\{S' \to \bullet S\})$ look like?*

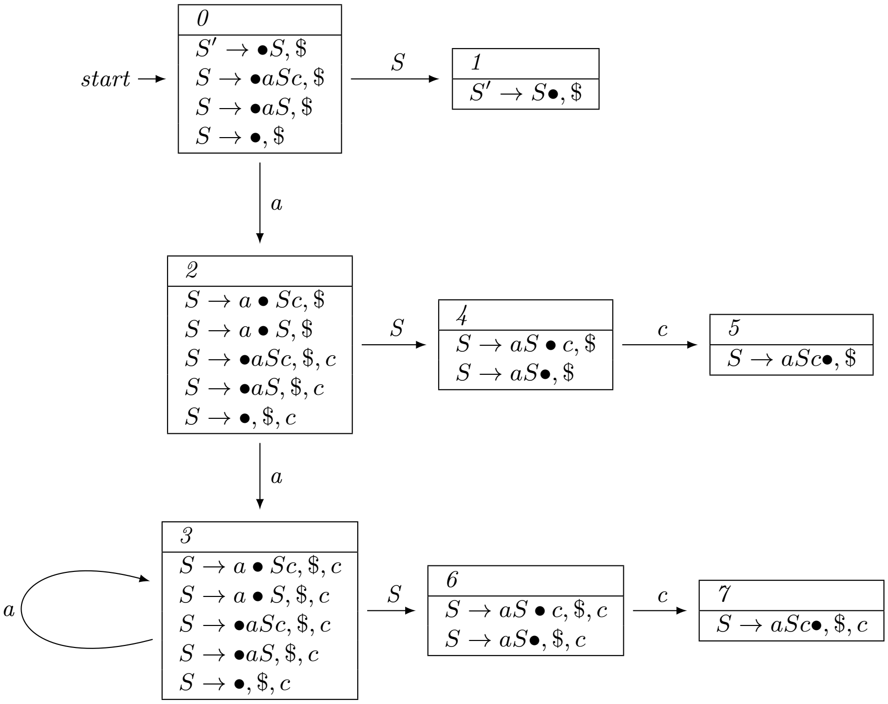
   (b) *Give the state $q_1 = goto - state(q_0, a)$.*

   (c) *What would be the entry of the field for q_0 and a in the LR(1) action table?*

2. *Now consider the CFG $\langle N, T, P, S\rangle$ with*

   $N = \{S\}$, $T = \{a, c\}$, start symbol S, and productions $1.S \to aSc$, $2.S \to aS$, $3.S \to \varepsilon$.

   The LR(1)-automaton is as follows:



   (a) *Give the LALR(1) automaton (states collapse that contain the same dotted productions, and all lookaheads are collected).*

   (b) *Read off the LALR(1) parse table.*

Question 2 (Canonical LR(1) parsing)

Consider the CFG $G_4 = \langle N, T, P, S\rangle$ with $N = \{S, A, B, C\}$, $T = \{a, b, c\}$ and productions $1.S \to ABC$, $2.A \to a$, $3.A \to aC$, $4.B \to b$, $5.B \to bC$, $6.C \to c$.

It's LR(1) parse table with the canonical LR algorithm is as follows:

	a	b	c	$	A	B	C	S
0	s1				2			3
1		r2	s11				10	
2		s5				4		
3				acc				
4			s7				6	
5			s8, r4				9	
6				r1				
7				r6				
8			r6					
9			r5					
10		r3						
11		r6						

1. *Is this grammar LR(1)?*

2. *Give the shift-reduce parsing trace (i.e., the sequence of pairs of stack and remaining input) that we obtain for the input* $w = abcc$.

   If the parser is not LR(1) and the input cannot be parsed deterministically, list all possibilities.