Parsing

Homework 2 (Top-Down Parsing), due 03 May 2021

Laura Kallmeyer

SS 2021, Heinrich-Heine-Universität Düsseldorf

Question 1 (Unger recognition) *Consider the CFG G with $N = \{S, A, B\}, T = \{a, b\}$, start symbol S and productions*

$$S \to AB \quad A \to a \,|\, Aa \quad B \to b \,|\, aB$$

and the input $w = aab$.

How does the resulting chart like when we do Unger recognition with tabulation along the lines of the example on slide 18, assuming that we make a prediction of a span of length 1 for a terminal a in a righthand side only if the terminal in the input that this span consists of is actually a.

Solution:

j				
3	$\langle S, t\rangle$	$\langle B, t\rangle$	$\langle B, t\rangle$ $\langle b, t\rangle$	
2	$\langle A, t\rangle$	$\langle a, t\rangle$		
1	$\langle A, t\rangle$ $\langle a, t\rangle$			
	1	2	3	i

Question 2 (Top-Down parsing) *Consider again the CFG G with $N = \{S, A, B\}, T = \{a, b\}$, start symbol S and productions*

$$S \to AB \quad A \to a \,|\, Aa \quad B \to b \,|\, aB$$

and the input $w = aab$.

1. *Give all parse trees for this input.*
2. *Is G ambiguous?*
3. *Give all pairs of prediction stack, remaining input and analysis stack that arise in a top-down parsing for this input. (The productions are assumed to be numbered in the order above, i.e., $S \to aSa$ is the 1st S-production etc.) List them in a table with a unique number for each pair and indicate from which other pair a new pair was obtained. Assume we have a filter such, that we do not generate pairs where the stack is longer than the remaining input + 2.*

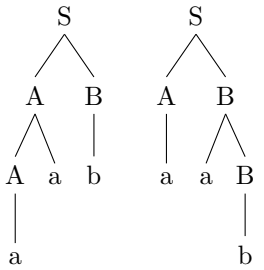
   Beginning of the table:

id	*stack*	*rem. input*	*analysis stack*	*obtained from*
1.	S	aab	–	*axiom*
2.	AB	aab	S_1	*from 1.*
3.	aB	aab	$S_1 A_1$	*from 2.*
4.	...	...		

4. *Give the resulting leftmost derivations.*
5. *Assuming we have a filter such that we do not generate pairs where the stack is longer than the remaining input. Which of the triples from the table in 3. would not be generated in this case?*
6. *Now assume that we pursue a depth-first strategy, which of the triples in the table in 3. would then remain? (We stop once the first successful parse is found.) Which is the leftmost derivation one encounters first?*

7. *Why is it a good idea, to have a Stack filter for our given CFG, when creating a table like in exercise 2? Write down your idea.*

Solution:

1. 

2. yes (see 1.)

3.

id	stack	rem. input	analysis stack	obtained from
1.	S	aab	–	axiom
2.	AB	aab	S_1	from 1.
3.	aB	aab	S_1A_1	from 2.
4.	B	ab	S_1A_1a	from 3.
5.	b	ab	$S_1A_1aB_1$	from 4.
6.	aB	ab	$S_1A_1aB_2$	from 4.
7.	B	b	$S_1A_1aB_2a$	from 6.
8.	aB	b	$S_1A_1aB_2aB_2$	from 7.
9.	b	b	$S_1A_1aB_2aB_1$	from 7.
10.	–	–	$S_1A_1aB_2aB_1b$	from 9.
11.	AaB	aab	S_1A_2	from 2.
12.	AaaB	aab	$S_1A_2A_2$	from 11.
13.	AaaaB	aab	$S_1A_2A_2A_2$	from 12.
14.	aaaaB	aab	$S_1A_2A_2A_2A_1$	from 13.
15.	aaaB	ab	$S_1A_2A_2A_2A_1a$	from 14.
16.	aaB	b	$S_1A_2A_2A_2A_1aa$	from 15.
17.	aaaB	aab	$S_1A_2A_2A_1$	from 12.
18.	aaB	ab	$S_1A_2A_2A_1a$	from 17.
19.	aB	b	$S_1A_2A_2A_1aa$	from 18.
20.	aaB	aab	$S_1A_2A_1$	from 11.
21.	aB	ab	$S_1A_2A_1a$	from 20.
22.	B	b	$S_1A_2A_1aa$	from 21.
23.	aB	b	$S_1A_2A_1aaB_2$	from 22.
24.	b	b	$S_1A_2A_1aaB_1$	from 22.
25.	–	–	$S_1A_2A_1aaB_1b$	from 24.

4. $S_1A_1B_2B_1$, $S_1A_2A_1B_1$

5. items to delete: 8, 12-19

6. Remaining items: 1-5., 6., 7., 9., 10., leftmost derivation is $S_1A_1B_2B_1$

7. Because the given CFG G is left-recursive (due to production $A \to Aa$) and the exercise is to give all pairs of the prediction stack, the stack can be filled indefinitely, i.e., one would run into a prediction loop with the problematic A-production. This is prevented by using a filter on the maximal length of the stack, as done in the exercise.