

Einführung in die Computerlinguistik

Hausaufgabe zu symbolischem Parsing, Abgabe 29.06.2020

Laura Kallmeyer

Sommer 2020, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 Betrachten Sie folgende CFG:

$G = \langle \{S, A, B\}, \{a, b, d\}, \{S \rightarrow aS \mid dB, B \rightarrow bB \mid b\}, S \rangle$

Geben Sie für die Eingabe *adbb* alle Paare von Resteingabe und Stack (ohne Analysestack), die sich beim Top-Down Parsing ergeben. Geben Sie dabei für jedes Paar außerdem an, aus welchem anderen Paar und mit welcher Operation es entstanden ist, ähnlich wie in dem Beispiel auf Folie 6.

Gehen Sie davon aus, dass jede Möglichkeit verfolgt wird, bis man nicht mehr weiterkommt.

Nehmen Sie außerdem an, dass nur Kombinationen erzeugt werden, bei denen der Stack nicht länger als die Resteingabe ist.

Lösung:

1. *adbb* *S* Initialisierung
2. *adbb* *aS* predict 1.
3. *adbb* *dB* predict 1.
4. *dbb* *S* scan 2.
5. *dbb* *aS* predict 4.
6. *dbb* *dB* predict 4.
7. *bb* *B* scan 6.
8. *bb* *bB* predict 7.
9. *bb* *b* predict 7.
10. *b* *B* scan 8.
11. *b* ε scan 9.
12. *b* *b* predict 10.
13. ε ε scan 12.

Aufgabe 2 Betrachten Sie nochmals die Grammatik aus Aufgabe 1 und die Eingabe *adbb*.

1. Wie sieht die Folge von Paaren von Stack und Resteingabe aus, die sich für *adbb* mit einem Shift-Reduce Parser ergibt? Geben Sie auch hier an, welche Operation zu dem jeweiligen Paar geführt hat und aus welchem anderen Paar es entstanden ist.

Es sollen in jeder Konfiguration sämtliche reduce Möglichkeiten durchgeführt werden und immer auch shift, egal ob diese Wege zum erfolgreichen Parse führen oder nicht.

2. Wie sieht die Rechtsableitung aus, die sich ergibt, wenn man die verwendeten Produktionen in umgekehrter Reihenfolge anwendet?

Lösung:

	1	<i>adbb</i>		
	2	<i>a</i>	<i>dbb</i>	shift 1
	3	<i>ad</i>	<i>bb</i>	shift 2
	4	<i>adb</i>	<i>b</i>	shift 3
	5	<i>adB</i>	<i>b</i>	reduce 4, $B \rightarrow b$
	6	<i>aS</i>	<i>b</i>	reduce 5, $S \rightarrow dB$
	7	<i>S</i>	<i>b</i>	reduce 6, $S \rightarrow aS$
1.	8	<i>adbb</i>	ε	shift 4
	9	<i>adBb</i>	ε	shift 5
	10	<i>aSb</i>	ε	shift 6
	11	<i>Sb</i>	ε	shift 7
	12	<i>adbB</i>	ε	reduce 8 $B \rightarrow b$
	13	<i>adBB</i>	ε	reduce 9 $B \rightarrow b$
	14	<i>aSB</i>	ε	reduce 10 $B \rightarrow b$
	15	<i>SB</i>	ε	reduce 11 $B \rightarrow b$
	16	<i>adB</i>	ε	reduce 12 $B \rightarrow bB$
	17	<i>aS</i>	ε	reduce 16 $S \rightarrow dB$
	18	<i>S</i>	ε	reduce 17 $S \rightarrow aS$

2. $S \Rightarrow aS \Rightarrow adB \Rightarrow adbB \Rightarrow adbb$

Aufgabe 3 Betrachten Sie nun folgende CFG:

$G = \langle \{S, A, B, C, D, F\}, \{a, b\}, \{S \rightarrow CD \mid AF, C \rightarrow AS \mid SB, D \rightarrow b, F \rightarrow a \mid b, A \rightarrow a, B \rightarrow b\}, S \rangle$

1. Wie sieht die Chart aus, die sich bei einem CYK-Parsing der Eingabe *aaabb* mit dieser CFG ergibt? (Nur Recognition.)
2. Gehört das Wort zur Sprache? Begründen Sie Ihre Antwort anhand der Chart.

Lösung:

	<i>l</i>					
	5	<i>C</i>				
	4	<i>S</i>	<i>S</i>			
1.	3	<i>C</i>	<i>C</i>	<i>C</i>		
	2	<i>S</i>	<i>S</i>	<i>S</i>		
	1	<i>A, F</i>	<i>A, F</i>	<i>A, F</i>	<i>F, B, D</i>	<i>F, B, D</i>
		1	2	3	4	5
						<i>i</i>

2. Nein da es im Feld mit $i = 1, l = 5$ kein *S* gibt.