

# Einführung in die Computerlinguistik

## Hausaufgabe zu symbolischem Parsing, Abgabe 18.06.2019

Laura Kallmeyer

Sommer 2019, Heinrich-Heine-Universität Düsseldorf

**Aufgabe 1** Betrachten Sie folgende CFG:

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

1. 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.

2. Woran erkennt der Parser, dass *adbb* zu der von der Grammatik generierten Sprache gehört?

Lösung:

1. 

|     |               |               |                 |
|-----|---------------|---------------|-----------------|
| 1.  | <i>adbb</i>   | <i>S</i>      | Initialisierung |
| 2.  | <i>adbb</i>   | <i>aSB</i>    | predict 1.      |
| 3.  | <i>adbb</i>   | <i>d</i>      | predict 1.      |
| 4.  | <i>adbb</i>   | <i>c</i>      | predict 1.      |
| 5.  | <i>dbb</i>    | <i>SB</i>     | scan 2.         |
| 6.  | <i>dbb</i>    | <i>dB</i>     | predict 5.      |
| 7.  | <i>dbb</i>    | <i>cB</i>     | predict 5.      |
| 8.  | <i>bb</i>     | <i>B</i>      | scan 6.         |
| 9.  | <i>bb</i>     | <i>b</i>      | predict 8.      |
| 10. | <i>bb</i>     | <i>bB</i>     | predict 8.      |
| 11. | <i>b</i>      | $\varepsilon$ | scan 9.         |
| 12. | <i>b</i>      | <i>B</i>      | scan 10.        |
| 13. | <i>b</i>      | <i>b</i>      | predict 12.     |
| 14. | $\varepsilon$ | $\varepsilon$ | scan 13.        |

2. *adbb* ist in der Sprache, da der Parser eine Konfiguration mit leerer Resteingabe und leerem Stack (Konfiguration 14) herleiten konnte.

**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>aS</i>   | <i>bb</i>     | reduce 3  |
|    | 5  | <i>adb</i>  | <i>b</i>      | shift 3   |
|    | 6  | <i>aSb</i>  | <i>b</i>      | shift 4   |
|    | 7  | <i>adB</i>  | <i>b</i>      | reduce 5  |
|    | 8  | <i>aSB</i>  | <i>b</i>      | reduce 6  |
|    | 9  | <i>S</i>    | <i>b</i>      | reduce 8  |
| 1. | 10 | <i>adbb</i> | $\varepsilon$ | shift 5   |
|    | 11 | <i>aSbb</i> | $\varepsilon$ | shift 6   |
|    | 12 | <i>adBb</i> | $\varepsilon$ | shift 7   |
|    | 13 | <i>aSBb</i> | $\varepsilon$ | shift 8   |
|    | 14 | <i>Sb</i>   | $\varepsilon$ | shift 9   |
|    | 15 | <i>adbB</i> | $\varepsilon$ | reduce 10 |
|    | 16 | <i>aSbB</i> | $\varepsilon$ | reduce 11 |
|    | 17 | <i>adBB</i> | $\varepsilon$ | reduce 12 |
|    | 18 | <i>aSBB</i> | $\varepsilon$ | reduce 13 |
|    | 19 | <i>SB</i>   | $\varepsilon$ | reduce 14 |
|    | 20 | <i>adB</i>  | $\varepsilon$ | reduce 15 |
|    | 21 | <i>aSB</i>  | $\varepsilon$ | reduce 16 |
|    | 22 | <i>S</i>    | $\varepsilon$ | reduce 21 |

2.  $S \Rightarrow aSB \Rightarrow aSbB \Rightarrow aSbb \Rightarrow adbb$

**Aufgabe 3** Betrachten Sie nun folgende CFG:

$G = \langle \{S, A, X, B\}, \{a, b\}, \{S \rightarrow AX \mid BX \mid SA \mid a \mid b, X \rightarrow SA \mid SB, A \rightarrow a, B \rightarrow b\}, S \rangle$

Wie sieht die Chart aus, die sich bei einem CYK-Parsing der Eingabe *aabba* mit dieser CFG ergibt?  
(Nur Recognition.)

Lösung:

|          |             |             |             |             |             |          |
|----------|-------------|-------------|-------------|-------------|-------------|----------|
| <i>l</i> |             |             |             |             |             |          |
| 5        | <i>S</i>    |             |             |             |             |          |
| 4        | <i>X</i>    | <i>X, S</i> |             |             |             |          |
| 3        | <i>X, S</i> | <i>S</i>    | <i>S</i>    |             |             |          |
| 2        | <i>X, S</i> | <i>X</i>    | <i>X</i>    | <i>X, S</i> |             |          |
| 1        | <i>S, A</i> | <i>S, A</i> | <i>B, S</i> | <i>B, S</i> | <i>S, A</i> |          |
|          | 1           | 2           | 3           | 4           | 5           | <i>i</i> |