# Tutorium 11

**Aufgabe 1: CYK-Algorithmus**

Gegeben sei eine Menge Nicht-Terminale $V \triangleq \{ A, B, C \}$, ein Alphabet $\Sigma \triangleq \{ a, b \}$, sowie eine CNF-Grammatik $G \triangleq (V, \{ a, b \}, P, S)$ mit

$$\begin{array}{rlll} P: & S & \to & AB \mid BC \\ & A & \to & BA \mid a \\ & B & \to & CC \mid b \\ & C & \to & AB \mid a \end{array}$$

1.a) *Begründe:* Warum ist die Grammatik G eine CNF-Grammatik?

- - - - - - Lösung - - - - - -

Für jede Produktionsregel aus $P$ gilt, dass auf der rechten Seite genau nur zwei Nicht-Terminale oder ein einzelnes Terminalsymbol stehen.

—— /Lösung ——

1.b) *Berechne:* Gegeben sei ein Wort $w_1 \triangleq baaba$. Löse mit dem CYK-Algorithmus das Wortproblem: $w_1 \in L(G)$ oder $w_1 \notin L(G)$?

- - - - - - Lösung - - - - - -

| $CYK_w(i,j)$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1: b | $\{ B \}$ | $\{ S, A \}$ | $\emptyset$ | $\emptyset$ | $\{ S, A, C \}$ |
| 2: a | $\{ A, C \}$ | $\{ B \}$ | $\{ B \}$ | $\{ S, A, C \}$ | |
| 3: a | $\{ A, C \}$ | $\{ S, C \}$ | $\{ B \}$ | | |
| 4: b | $\{ B \}$ | $\{ S, A \}$ | | | |
| 5: a | $\{ A, C \}$ | | | | |

Es gilt also $w_1 \in L(G)$, da $S \in CYK_w(1,5)$.

ERKLÄRUNG: Die Erklärung ist nicht mehr Teil der Lösung, sondern dient nur zum besseren Verständnis der Lösung. Wir betrachten die Berechnung der Einträge an zwei Beispielen entsprechend der Proposition 3.2.1 aus der Formelsammlung:

- $CYK_w(2,1) = \{ X \in V \mid (X \to a) \in P \} = \{ A, C \}$ wegen $(A \to a) \in P$ und $(C \to a) \in P$
- $CYK_w(4,2) = \{ X \in V \mid \exists Y \in CYK_w(4, 1).\exists Z \in CYK_w(5, 1).(X \to YZ) \in P \} = \{ S, A \}$ wegen $(S \to BC) \in P$ und $(A \to BA) \in P$

Algorithmisch betrachtet gehen wir beim Ausfüllen der Tabelle von links nach rechts vor und benutzen dabei bereits berechnete Einträge. Zum Beispiel betrachten wir für die Berechnung des Eintrags $CYK_w(1,3)$ jeweils die folgenden beiden Einträge in Kombination:

- die Einträge $CYK_w(1,1)$ und $CYK_w(2,2)$ (hier finden wir keine Produktionsregel, die $BB$ produziert), sowie
- die Einträge $CYK_w(1,2)$ und $CYK_w(3,1)$ (hier finden wir keine Produktionsregel, die $SA$, $SC$, $AA$ oder $AC$ produziert).

Wir erhalten deswegen $CYK_w(1,3) = \emptyset$.

—— /Lösung ——

1.c) *Berechne:* Gegeben sei ein Wort $w_2 \triangleq bba$. Löse mit dem CYK-Algorithmus das Wortproblem: $w_2 \in L(G)$ oder $w_2 \notin L(G)$?

- - - - - - Lösung - - - - - -

| $CYK_w(i,j)$ | 1 | 2 | 3 |
|---|---|---|---|
| 1: b | $\{ B \}$ | $\emptyset$ | $\{ A \}$ |
| 2: b | $\{ B \}$ | $\{ S, A \}$ | |
| 3: a | $\{ A, C \}$ | | |

Es gilt also $w_2 \notin L(G)$, da $S \notin CYK_w(1,3)$.

—— /Lösung ——

**Aufgabe 2: Die schöne Seite des kontextfreien Pumping-Lemmas**

Sei $\Sigma \triangleq \{ a \}$ und $A \triangleq \{ a^p \mid p \text{ ist eine Primzahl} \}$.

2.a) *Beweise* mit Hilfe des Pumping-Lemmas für kontextfreie Sprachen, dass $A$ nicht kontextfrei ist.
*Hinweis: Hier kommt es vor allem darauf an* $k$ *aus PUMP-CFL richtig zu wählen.*

- - - - - - - - - - - - - - - Lösung - - - - - - - - - - - - - - -

Wir zeigen, dass PUMP-CFL für $A$ nicht gilt. Sei $n \in \mathbb{N}$ beliebig, aber fest. Wir wählen das Wort $z = a^p$, wobei $p$ eine Primzahl größer oder gleich $n$ ist und somit gilt $z \in A$ und $| z | \geqslant n$. (Da es unendlich viele Primzahlen gibt, gibt es auch immer eine solche Primzahl $p$.) Sei $z = uvwxy$ eine beliebige Zerlegung mit $|vx| \geqslant 1$ und $|vwx| \leqslant n$. Dann gilt $u = a^i, v = a^j, w = a^l, x = a^m$ und $y = a^{p-i-j-l-m}$ mit $i + j + l + m \leqslant p$, $j + l + m \leqslant n$ und $j + m \geqslant 1$. Wähle $k = p + 1$. Dann gilt $uv^{p+1}wx^{p+1}y = a^{i+j(p+1)+l+m(p+1)+p-i-j-l-m} = a^{p+pj+pm} = a^{p(1+j+m)}$. Da $p$ eine Primzahl ist, gilt $p \neq 1$. Aus $j + m \geqslant 1$ folgt weiterhin $1 + j + m \neq 1$ und somit ist $p(1 + j + m)$ keine Primzahl. Also $uv^{p+1}wx^{p+1}y \notin A$ und es folgt, dass $A$ nicht kontextfrei ist.

——————————— /Lösung ———————————

**Aufgabe 3: Die nicht so schöne Seite des kontextfreien Pumping-Lemmas**

Sei $\Sigma \triangleq \{ a, b \}$ und $A \triangleq \{ a^n b^n a^n \mid n \in \mathbb{N}^+ \}$. Gegeben ist eine Abbildung $B : \Sigma^+ \to \mathbb{N}$ mit:

$$B(z) = \begin{cases} 0, & \text{for } z \in \Sigma \\ B(uv) + 1, & \text{for } z = uvw \text{ and } v, w \in \Sigma \text{ and } v \neq w \\ B(uv), & \text{for } z = uvw \text{ and } v, w \in \Sigma \text{ and } v = w. \end{cases}$$

Die Abbildung $B$ ordnet also jedem Wort $z \in \Sigma^+$ die Anzahl der Buchstabenwechsel (oder auch Buchstabengrenzen) von $z$ zu. Wir sprechen folgend auch von Blöcken von Buchstaben.

3.a) *Beweise:* Für alle Wörter $w \in A$ gilt $B(w) = 2$.

- - - - - - - - - - - - - - - Lösung - - - - - - - - - - - - - - -

Sei $w \in A$ beliebig. Dann gibt es ein $n \in \mathbb{N}^+$ mit $w = a^n b^n a^n$. Es gilt also $(w)_n \neq (w)_{n+1}$ und $(w)_{2n} \neq (w)_{2n+1}$. Für alle (anderen) $i \in \{ 1, \ldots, 3n - 1 \} \setminus \{ n, 2n \}$ gilt $(w)_i = (w)_{i+1}$.

——————————— /Lösung ———————————

3.b) *Beweise:* Für alle $v, w \in \Sigma^+$ gilt $B(v) + B(w) \leqslant B(vw)$.

- - - - - - - - - - - - - - - Lösung - - - - - - - - - - - - - - -

Seien $v, w \in \Sigma^+$. Wir stellen zunächst fest, dass jeder Buchstabenwechsel von $v$ auch ein Buchstaben wechsel von $vw$ ist. Genauso ist jeder Buchstabenwechsel von $w$ auch einer von $vw$. Sei weiterhin $n = |v|$. Wir unterscheiden zwei Fälle.

- 1. Fall: $(v)_n = (w)_1$. In diesem Fall kommt am Übergang von $v$ nach $w$ kein weiterer Buchstabenwechsel hinzu. Es gilt also $B(v) + B(w) = B(vw)$.
- 2. Fall: $(v)_n \neq (w)_1$. In diesem Fall zählen wir einen weiteren Buchstabenwechsel am Übergang von $v$ nach $w$. Es gilt also $B(v) + B(w) < B(vw)$.

Da dies alle Fälle sind, gilt die Aussage.

——————————— /Lösung ———————————

3.c) *Vervollständige* mit Hilfe der Abbildung $B$ und des Pumping-Lemmas für kontextfreie Sprachen, den folgenden Beweis, dass $A$ nicht kontextfrei ist.

Sei $n \in \mathbb{N}$. Wir wählen das Wort $z = a^{n+1}b^{n+1}a^{n+1}$ mit $z \in A$, denn $n + 1 \in \mathbb{N}^+$. Sei $z = uvwxy$ eine beliebige Zerlegung mit $|vx| \geqslant 1$ und $|vwx| \leqslant n$. Da $z \in A$ gilt $B(z) = 2$. Es gibt also drei Blöcke an Buchstaben, die jeweils $n + 1$ lang sind. Wir betrachten in den folgenden Fälle, wie sich die Zerlegung auf diese drei Blöcke verteilen kann:

- - - - - - - - - - - - - - - Lösung - - - - - - - - - - - - - - -

*Hinweis: Im folgenden geben wir zwei Lösungen. Die Erste soll dabei helfen eine Intuition für die Beweisstruktur zu erlangen. Die Zweite ist dann die formal korrekte Version des Beweises an.*

1. Fall: Sowohl $v$ als auch $x$ liegen in einem Block: $B(vwx) = 0$. Da $|vwx| \leqslant n$ ist es möglich, dass $v$ und $x$ in nur einem Block liegen. Somit gibt es zwei Blöcke, in denen $v$ und $x$ jeweils nicht enthalten sind. Wähle $k = 2$. Für $uv^2wx^2y$ gilt, dass die beiden Blöcke, die $v$ und $x$ nicht enthalten, weiterhin jeweils $n+1$ lang sind. Gleichzeitig gilt für den Block, der $v$ und $x$ enthält, dass er in $uv^2wx^2y$ mindestens einen weiteren Buchstaben enthält (da $|vx| \geqslant 1$) und somit ist er echt größer als $n+1$. Es folgt also $uv^2wx^2y \notin A$.
2. Fall: Die Worte $v$ und $x$ liegen in zwei verschiedenen Blöcken und dabei liegen $v$ und $x$ jeweils vollständig in einem Block: $B(vwx) = 1, B(v) = 0$ und $B(x) = 0$. Somit gibt es einen Block, der weder $v$ noch $x$ enthält. Wähle $k = 2$. Für $uv^2wx^2y$ gilt, dass der Block, der weder $v$ noch $x$ enthält weiterhin $n+1$ lang ist. Gleichzeitig gilt für mindestens entweder für den Block, der $v$ enthält oder für den, der $x$ enthält, dass er für $uv^2wx^2y$ mindestens einen weiteren Buchstaben enthält (da $|vx| \geqslant 1$) und somit ist er echt größer als $n+1$. Es folgt $uv^2wx^2y \notin A$.
3. Fall: Entweder $v$ oder $x$ erstreckt sich über zwei Blöcke und es gilt: Entweder $B(v) = 1$ oder $B(x) = 1$. Da $|vwx| \leqslant n$ gilt, können sich nicht $v$ und $x$ jeweils über zwei Blöcke erstrecken. Wir betrachten o.B.d.A. den Fall $B(v) \geqslant 1$. Wähle $k = 2$ und betrachte das Wort $uv^2wx^2y$. Damit $uv^2wx^2y \in A$ müsste nach 3a) gelten $B(uv^2wx^2y) = 2$. Es gibt eine Grenze zwischen zwei Blöcken außerhalb von $v$. Diese liegt entweder links oder rechts von $v$. Wir betrachten die Anzahl der Blöcke für beide Fälle:
   1. Fall: Die weitere Grenze liegt rechts von $v$. $2 \neq 3 = 0+2+1 \leqslant B(u) + 2B(v) + B(wx^2y) \leqslant B(uv^2wx^2y)$. Und somit $uv^2wx^2y \notin A$ nach 3a).
   2. Fall: Die weitere Grenze liegt links von $v$. $2 \neq 3 = 1+2+0 \leqslant B(u) + 2B(v) + B(wx^2y) \leqslant B(uv^2wx^2y)$. Und somit $uv^2wx^2y \notin A$ nach 3a).

Da $|vwx| \leqslant n$ sind das alle möglichen Fälle. Somit gilt die Aussage.

------------------------- alternative Lösungen -------------------------

- Wir betrachten zunächst den Fall, dass $B(v) \geqslant 1$. Dann gilt einer der folgenden Fälle:
  - $u = a^i$, $v = a^{n+1-i}b^j$ und $wxy = b^{n+1-j}a^{n+1}$ (mit Einschränkungen entsprechend 2, 3 PUMP-CFL)
  - $u = a^{n+1}b^i$, $v = b^{n+1-i}a^j$ und $wxy = a^{n+1-j}$ (mit Einschränkungen entsprechend 2, 3 PUMP-CFL)

  Wähle $k = 2$ und betrachte das Wort $uv^2wx^2y$. Entsprechend der beiden möglichen Zerlegungen gilt:
  - $2 \neq 3 = 0+2+1 \leqslant B(u) + 2B(v) + B(wx^2y) \leqslant B(uv^2wx^2y)$. Und somit $uv^2wx^2y \notin A$ nach 3a).
  - $2 \neq 3 = 1+2+0 \leqslant B(u) + 2B(v) + B(wx^2y) \leqslant B(uv^2wx^2y)$. Und somit $uv^2wx^2y \notin A$ nach 3a).
- Betrachte nun den Fall, dass $B(x) \geqslant 1$. Dann gilt einer der folgenden Fälle:
  - $uvw = a^i$, $x = a^{n+1-i}b^j$ und $y = b^{n+1-j}a^{n+1}$ (mit Einschränkungen entsprechend 2, 3 PUMP-CFL)
  - $uvw = a^{n+1}b^i$, $x = b^{n+1-i}a^j$ und $y = a^{n+1-j}$ (mit Einschränkungen entsprechend 2, 3 PUMP-CFL)

  Wähle $k = 2$ und betrachte das Wort $uv^2wx^2y$. Entsprechend der beiden möglichen Zerlegungen gilt:
  - $2 \neq 3 = 0+2+1 \leqslant B(uvw) + 2B(x) + B(y) \leqslant B(uv^2wx^2y)$. Und somit $uv^2wx^2y \notin A$ nach 3a).
  - $2 \neq 3 = 1+2+0 \leqslant B(uvw) + 2B(v) + B(y) \leqslant B(uv^2wx^2y)$. Und somit $uv^2wx^2y \notin A$ nach 3a).
- Für den Fall, dass $B(v) = 0$ und $B(x) = 0$ gilt, ergeben sich die folgenden Fälle:
  - ($v$ und $x$ liegen beide im ersten a-Block) Es gilt $u = a^i, v = a^j, w = a^l, x = a^m$ und $y = a^{n+1-i-j-l-m}b^{n+1}a^{n+1}$ mit $j+l+m \leqslant n$, $i+j+l+m \leqslant n+1$ und $j+m \geqslant 1$. Wähle k = 2. Dann ist $uv^2wx^2y = a^{n+1+j+m}b^{n+1}a^{n+1}$ und da $j+m \geqslant 1$ gilt $n+1+j+m \neq n+1$ und somit $uv^2wx^2y \notin A$.
  - ($v$ und $x$ liegen beide im b-Block) Es gilt $u = a^{n+1}b^i, v = b^j, w = b^l, x = b^m$ und $y = b^{n+1-i-j-l-m}a^{n+1}$ mit $j+l+m \leqslant n$, $i+j+l+m \leqslant n+1$ und $j+m \geqslant 1$. Wähle k = 2. Dann ist $uv^2wx^2y = a^{n+1}b^{n+1+j+m}a^{n+1}$ und da $j+m \geqslant 1$ gilt $n+1+j+m \neq n+1$ und somit $uv^2wx^2y \notin A$.
  - ($v$ und $x$ liegen beide im zweiten a-Block) Es gilt $u = a^{n+1}b^{n+1}a^i, v = a^j, w = a^l, x = a^m$ und $y = a^{n+1-i-j-l-m}$ mit $j+l+m \leqslant n$, $i+j+l+m \leqslant n+1$ und $j+m \geqslant 1$. Wähle k = 2. Dann ist $uv^2wx^2y = a^{n+1}b^{n+1}a^{n+1+j+m}$ und da

$j+m \geqslant 1$ gilt $n+1+j+m \neq n+1$ und somit $uv^2wx^2y \notin A$.

- (ν liegt im ersten a-Block und x im b-Block) Es gilt $u = a^i, v = a^j, w = a^{n+1-i-j}b^l, x = b^m$ und $y = b^{n+1-l-m}a^{n+1}$ mit $i+j \leqslant n+1$, $j+l+m \leqslant n$ und $j+m \geqslant 1$. Wähle k = 2. Dann ist $uv^2wx^2y = a^{n+1+j}b^{n+1+m}a^{n+1}$. Mit $j+m \geqslant 1$ folgt $n+1+j \neq n+1$ oder $n+1+m \neq n+1$. Und somit $uv^2wx^2y \notin A$.
- (ν liegt im b-Block und x im zweiten a-Block) Es gilt $u = a^{n+1}b^i, v = b^j, w = b^{n+1-i-j}a^l, x = a^m$ und $y = a^{n+1-l-m}$ mit $i+j \leqslant n+1$, $j+l+m \leqslant n$ und $j+m \geqslant 1$. Wähle k = 2. Dann ist $uv^2wx^2y = a^{n+1}b^{n+1+j}a^{n+1+m}$. Mit $j+m \geqslant 1$ folgt $n+1+j \neq n+1$ oder $n+1+m \neq n+1$. Und somit $uv^2wx^2y \notin A$.

Da $|vwx| \leqslant n$ sind das alle Fälle. Und somit gilt die Aussage.

/alternative Lösungen

/Lösung