Blatt 6

Soriano Ortiz, Karime (469451)
Grez Leisinger, Karl Nicolas (471888)
Galytskyy, Danylo (469032)

23. Mai 2025

Aufgabe 1

Aufgabe 1 **4 Punkte**

Sei G die Grammatik $(\{S, A, B\}, \{a, b, c\}, P, S)$, wobei P die folgenden Regeln enthält:

$$\begin{array}{rcl} S & \rightarrow & BAc \mid BcA \\ A & \rightarrow & Ab \mid aa \mid B \\ B & \rightarrow & BAB \mid \varepsilon \end{array}$$

Wandeln Sie G in eine kontextfreie Grammatik in Chomsky-Normalform um, die $L(G)$ erzeugt. Folgen Sie dazu dem Vorgehen aus der Vorlesung und geben Sie alle vier Grammatiken an, die Sie in den Zwischenschritten erhalten. (Beachten Sie auch die ε-Bedingung.)

Schritt 1: Neues Startsymbol

Wir führen ein neues Startsymbol S_0 ein:

$$S_0 \rightarrow S$$

Schritt 2: Entfernen der ε-Produktionen

ε-Produktion: $B \rightarrow \varepsilon$
Nullable Variable: B
Neue Produktionen durch Weglassen von B:

$$S \rightarrow BAc \mid BcA \mid Ac \mid cA$$
$$A \rightarrow Ab \mid aa \mid B \mid a$$
$$B \rightarrow BAB \mid AB \mid BA \mid A$$

Dann entfernen wir $B \rightarrow \varepsilon$.

Schritt 3: Entfernen von Unit-Produktionen

Unit-Produktionen:

$$A \rightarrow B, \quad B \rightarrow A$$

Wir ersetzen sie durch die Produktionen der jeweils anderen Variable:

$$A \rightarrow Ab \mid aa \mid a \mid BAB \mid AB \mid BA$$
$$B \rightarrow BAB \mid AB \mid BA \mid Ab \mid aa \mid a$$

Schritt 4: Terminale in langen Regeln ersetzen

Einführung neuer Variablen für Terminale:

$$\begin{aligned} X_a &\to a \\ X_b &\to b \\ X_c &\to c \end{aligned}$$

Ersetzung in Regeln:

$$\begin{aligned} Ab &\to AX_b \\ aa &\to X_aX_a \\ BAc &\to BAX_c \\ BcA &\to BX_cA \\ a &\to X_a \quad \text{(falls notwendig)} \end{aligned}$$

Schritt 5: Binäre Produktionen (Länge = 2)

Einführung zusätzlicher Hilfsvariablen:

$$\begin{aligned} Y_1 &\to AX_c \quad \Rightarrow S \to BY_1 \\ Y_2 &\to X_cA \quad \Rightarrow S \to BY_2 \\ Y_3 &\to AB \quad \Rightarrow B \to BY_3 \end{aligned}$$

Endgültige Grammatik in Chomsky-Normalform

Nichtterminale: $S_0, S, A, B, X_a, X_b, X_c, Y_1, Y_2, Y_3$
Terminale: a, b, c
Produktionen:

$$\begin{aligned} S_0 &\to S \\ S &\to BY_1 \mid BY_2 \mid AX_c \mid X_cA \\ Y_1 &\to AX_c \\ Y_2 &\to X_cA \\ Y_3 &\to AB \\ A &\to AX_b \mid X_aX_a \mid BY_3 \mid AB \mid BA \\ B &\to BY_3 \mid AB \mid BA \mid AX_b \mid X_aX_a \mid A \\ X_a &\to a \\ X_b &\to b \\ X_c &\to c \end{aligned}$$

Aufgabe 2

Aufgabe 2 **2+1 Punkte**

Wir betrachten die kontextfreie Grammatik $G = (N, \{a, b, c, d\}, P, S)$ mit $N = \{S, A, B, C, D, E\}$. P beinhaltet die folgenden Regeln:

$$\begin{array}{lcl} S & \to & CE \mid EEB \\ A & \to & CEc \mid CaC \mid aACD \mid EeB \\ B & \to & DB \mid BbSD \\ C & \to & CEb \mid aDA \mid cad \\ D & \to & AEC \mid CA \\ E & \to & CEB \mid aCD \end{array}$$

(a) Wenden Sie den Markierungsalgorithmus an, um zu bestimmen, welche Nichtterminalsymbole terminierend sind. Geben Sie diese an.

(b) Ist $L(G)$ leer? Begründen Sie kurz Ihre Antwort.

(a) Markierungsalgorithmus: terminierende Nichtterminalsymbole

Initialisierung: Markiere alle Regeln mit nur Terminalsymbolen.

$$C \to cad \quad \Rightarrow \quad \text{markiere } C$$

Iteration 1: Regeln, deren rechte Seite nur markierte Nichtterminale oder Terminale enthalten:

- $A \to CaC$ (C markiert) $\Rightarrow$ markiere A
- $D \to CA$ (C, A markiert) $\Rightarrow$ markiere D
- $A \to aACD$ (A, C, D markiert) $\to$ gültig, A bleibt markiert
- $E \to aCD$ (C, D markiert) $\Rightarrow$ markiere E

Iteration 2:

- $S \to CE$ (C, E markiert) $\Rightarrow$ markiere S

Ergebnis: Die Menge der terminierenden Nichtterminale ist:

$$\boxed{\{A, C, D, E, S\}}$$

Nichtterminierend: B

(b) Ist $L(G)$ leer?

Nein. Da das Startsymbol S terminierend ist, existiert mindestens eine Ableitung, die nur Terminale enthält.

$$\boxed{L(G) \neq \emptyset}$$

Aufgabe 3

(a)

Es reicht *nicht*, zunächst Schritt II und dann Schritt I durchzuführen, um eine Grammatik mit nur noch produktiven Nichtterminalen zu erhalten.
Beispielgrammatik:

$$S \to A \qquad A \to B \qquad B \to B$$

Analyse:

- Alle Nichtterminale sind von S aus erreichbar. Also ändert Schritt II nichts.
- B ist jedoch nicht terminierend, da es nur auf sich selbst verweist.
- Führt man nun Schritt I aus, so wird B entfernt, und damit auch A, da $A \to B$ nun ungültig ist.
- Schließlich bleibt $S \to A$, aber A wurde entfernt. Also ist S nun auch nutzlos.

Fazit: Wenn man *zuerst* Schritt II (Entfernen unerreichbarer Nichtterminale) ausführt, bleiben nutzlose, nicht-produktive Nichtterminale möglicherweise erhalten. Daher reicht diese Reihenfolge nicht aus.

(b)

Wird zunächst Schritt I und dann Schritt II durchgeführt, enthält die resultierende Grammatik nur noch produktive Nichtterminale.
Beweisidee:

- Durch Schritt I entfernen wir alle Nichtterminale, die nicht terminierend (also nicht produktiv) sind. Dies schließt auch Produktionen aus, die auf nicht-produktive Nichtterminale verweisen.
- Danach entfernt Schritt II alle Nichtterminale, die nicht erreichbar sind, d. h. die vom Startsymbol aus nicht benutzt werden können.

Fazit: Nach beiden Schritten enthält die Grammatik nur solche Nichtterminale, die sowohl produktiv als auch erreichbar sind. Somit ist jede Produktion in der resultierenden Grammatik gültig, und jedes verwendete Nichtterminal ist produktiv.

Aufgabe 4

Sei G die Grammatik $(\{S, A, B, C\}, \{a, b, c\}, P, S)$ in Chomsky-Normalform, wobei P aus folgenden Regeln besteht:

$$\begin{array}{lcl} S & \rightarrow & CA \mid BC \\ A & \rightarrow & AB \mid a \\ B & \rightarrow & CB \mid CA \mid b \\ C & \rightarrow & AA \mid c \end{array}$$

Stellen Sie mit Hilfe des CYK-Algorithmus durch Ausfüllen der entsprechenden Tabelle fest, ob die Wörter *cabca* und *aabcb* zu $L(G)$ gehören. Nutzen Sie bei einer positiven Antwort die Tabelleneinträge, um eine Ableitung des Wortes in G anzugeben.

Wort: `cabca`

	1	2	3	4	5
1	A	{S,B}	A	S	{S,B}
2		A			
3			A		
4					
5					

Tabelle 1: CYK-Tabelle für das Wort `cabca`

Analyse: In der Zelle (1,5) taucht das Startsymbol S nicht auf.
Ergebnis: `cabca` gehört **nicht** zur Sprache $L(G)$.

Wort: `aabcb`

Analyse: In der Zelle (1,5) befindet sich das Startsymbol S, daher gehört das Wort zur Sprache.
Ableitung von „aabcb“:

$$S \Rightarrow CA$$
$$C \Rightarrow AA \Rightarrow aA \Rightarrow aAB \Rightarrow aaB$$
$$A \Rightarrow AB \Rightarrow aB \Rightarrow ab$$
$$B \Rightarrow CB \Rightarrow cB \Rightarrow cb$$
$$\Rightarrow aabcb$$

	1	2	3	4	5
1	A	{C}	{A}	{S}	{B}
2					
3					
4				{A}	
5	{S}				

Tabelle 2: CYK-Tabelle für das Wort `aabcb`

Ergebnis: $\boxed{\texttt{aabcb} \text{ gehört zur Sprache } L(G).}$