Universität Siegen
Lehrstuhl Theoretische Informatik
Markus Lohrey

Formale Sprachen und Automaten
SS 2025

Übungsblatt 11

Aufgabe 1. Sind die folgenden Aussagen wahr oder falsch?

(a) Wenn L eine Sprache mit $\mathrm{index}(R_L) = \infty$ ist, dann ist L kontextfrei.

(b) Wenn L eine nicht kontextfreie Sprache ist, dann ist $\mathrm{index}(R_L) = \infty$.

(c) Wenn G eine Grammatik in CNF ist, dann ist $L(G)$ kontextfrei und nicht regulär.

(d) Es existieren kontextfreie Sprachen L_1 und L_2 so, dass $L_1 \cap L_2$ auch kontextfrei ist.

(e) Sei L eine Sprache mit $\mathrm{index}(R_L) = \infty$. Dann gilt für jeden Homomorphismus h und jede Sprache L' mit $h(L') = L$, dass L' kontextfrei, aber nicht regulär ist.

(f) Für jedes nicht unäre Alphabet Σ existieren unendlich viele kontextfreie, nicht reguläre Sprachen $L_i \subseteq \Sigma^*$ mit folgender Eigenschaft: Sei $\mathcal{H}$ die Menge aller Homomorphismen $h : \Sigma^* \to \Sigma^*$. Dann ist $\bigcup_{h \in \mathcal{H}} h(L_i)$ regulär.

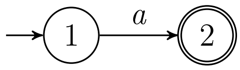
Lösung zu Aufgabe 1.

(a) Falsch, weil für alle nicht-regulären Sprachen L auch $\mathrm{index}(R_L) = \infty$ gilt, aber nicht jede nicht-reguläre Sprache ist eine kontextfreie Sprache, zum Beispiel $L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$.

(b) Wahr, weil jede Sprache, die nicht kontextfrei ist, auch nicht regulär ist und somit $\mathrm{index}(R_L) = \infty$ gilt.

(c) Falsch: Sei z.B. $G = (\{S\}, \{a\}, P, S)$ mit $P = \{S \to a\}$.

Diese Grammatik ist in CNF, aber es gilt $L(G) = \{a\}$ und das ist eine reguläre Sprache. Sie wird beispielsweise vom nachfolgenden NFA erzeugt:

M_G: 

(d) Wahr: Sei z.B. $L_1 = L_2$, dann ist $L_1 \cap L_2 = L_1 = L_2$ und somit genau dann kontextfrei, wenn L_1 und L_2 kontextfrei sind. Die Aussage ist also auch dann wahr, wenn man sie zu „kontextfrei, aber nicht regulär“ verschärft.

(e) Falsch, z.B. $L' = L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ und $h = \mathrm{id}$.

Anmerkung: Selbst, wenn man fordert, dass L kontextfrei (und nicht regulär) ist, ist die Aussage falsch (L' wie oben, $L = \{a^n b^n \mid n \in \mathbb{N}\}$ und h definiert durch $h(a) = a, h(b) = b, h(c) = \varepsilon$). Wenn wiederum dann h auf Isomorphismen eingeschränkt wird, ist die Aussage richtig (h^{-1} bildet von einer kontextfreien auf eine kontextfreie Sprache ab).

(f) Wahr: Sei $a \in \Sigma$. Es gibt unendlich viele nicht reguläre, aber kontextfreie Sprachen mit $a \in L_i$, zum Beispiel $L_i = \{a^n b^n \mid n \geq i\} \cup \{a\}$ für $\Sigma = \{a, b\}$. Die Menge aller Homomorphismen $\mathcal{H}$ enthält insbesondere folgende Abbildungen: $h_w : \Sigma^* \to \Sigma^*$ mit $h_w(a) = w$ und $h_w(s) = \varepsilon$ mit $w \in \Sigma^*$ und $s \in \Sigma \setminus \{a\}$. Also enthält $h_w(L_i)$ das Wort w und $\bigcup_{h \in \mathcal{H}} h(L_i)$ enthält Σ^*. Somit haben wir für jedes $i \in \mathbb{N}$ eine Sprache L_i mit $\bigcup_{h \in \mathcal{H}} h(L_i) = \Sigma^*$ (und Σ^* ist regulär).

Aufgabe 2. Gegeben ist die kontextfreie Grammatik $G = (V, \Sigma, P, S)$ in Chomsky-Normalform über $\Sigma = \{a, b\}$ mit $V = \{S, X, Y, A, B\}$ und den folgenden Produktionen:

$$\begin{aligned} P: \ & S \to a \mid b \mid AA \mid BB \mid XA \mid YB \\ & X \to AS \\ & Y \to BS \\ & A \to a \\ & B \to b \end{aligned}$$

(a) Überprüfen Sie mit dem CYK-Algorithmus, ob $abbbba \in L(G)$ gilt.

(b) Welche Sprache erzeugt G?

Lösung zu Aufgabe 2. (a)

	a	b	b	b	b	a
j = 1	S,A	S,B	S,B	S,B	S,B	S, A
j = 2	X	Y,S	Y,S	Y,S	Y	
j = 3	X	Y,S	Y,S	∅		
j = 4	X	Y,S	∅			
j = 5	X	∅				
j = 6	S					

Da S im letzten Feld steht, gilt $abbbba \in L(G)$.

Wir verwenden die Anleitung aus dem FSA-Skript (ab Folie 236) um die obige Tabelle auszufüllen.

(b) Wir erhalten einen besseren Überblick über die von G erzeugte Sprache, indem wir einige Nicht-Terminale eliminieren.

Zunächst streichen wir die Produktionen $A \to a$ und $B \to b$ und setzen den jeweiligen Buchstaben dort ein wo zuvor A oder B stand:

$$\begin{aligned} P' : S &\to a|b|aa|bb|Xa|Yb \\ X &\to aS \\ Y &\to bS \end{aligned}$$

Dann können wir X auf rechten Seiten durch aS und Y durch bS ersetzen und die entsprechenden Regeln für X und Y löschen:

$$P'' : S \to a|b|aa|bb|aSa|bSb$$

Die Grammatik $G' = (V, \Sigma, P'', S)$ ist eine zu G äquivalente Grammatik, also gilt $L(G) = L(G')$. In dieser vereinfachten Form sehen wir, dass G' alle nicht-leeren Palindrome über dem Alphabet $\Sigma = \{a, b\}$ erzeugt.

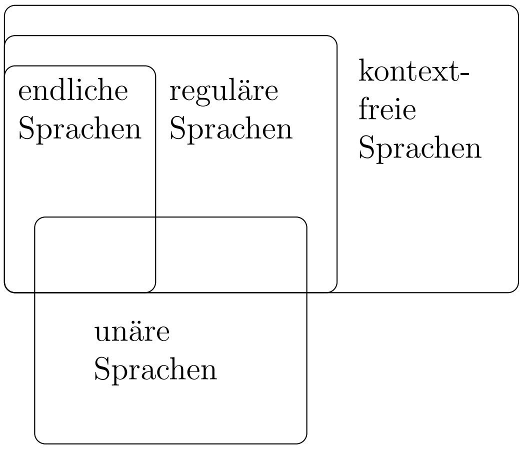
In Mengenschreibweise:

$L(G') = L(G) = \{w \in \Sigma^+ \mid w = w^r\}$.

Aufgabe 3. Stellen Sie die folgenden Sprachklassen über dem Alphabet $\{a, b\}$ in einem Venn-Diagramm dar. Geben Sie zusätzlich Beispielsprachen für alle Teilbereiche in ihrem Diagramm an.

- Reguläre Sprachen
- Kontextfreie Sprachen
- Endliche Sprachen
- Unäre Sprachen (alle Sprachen L mit $L \subseteq \{a\}^*$)

Lösung zu Aufgabe 3.



- endlich und unär: $\{a\}$
- endlich und nicht unär: $\{ab\}$
- regulär, unär und nicht endich: $\{a^n \mid n \in \mathbb{N}\}$
- regulär, nicht unär und nicht endlich: $\{ab^n \mid n \in \mathbb{N}\}$
- unär, nicht kontextfrei: $\{a^p \mid p \text{ ist Primzahl}\}$
- nicht unär, nicht regulär, kontextfrei: $\{a^n b^n \mid n \in \mathbb{N}\}$
- nicht unär, nicht kontextfrei: $\{ww \mid w \in \{a, b\}^*\}$

Hinweis: Jede kontextfreie, unäre Sprache ist auch regulär (Satz von Parikh, siehe Übungsblatt 8, Aufgabe 3).

Aufgabe 4. Sei $L \subseteq \Sigma^*$ eine beliebige Sprache und h ein Homomorphismus. Der *inverse Homomorphismus* von L ist definiert als

$$h^{-1}(L) = \{w \mid h(w) \in L\}.$$

Seien nun $L = \{w \in \{a, b\}^* \mid \#_a(w) = \#_b(w)\}$ und h definiert via $h(0) = ab$, $h(1) = ba$. Beschreiben Sie $L' = h^{-1}(L) \subseteq \{0, 1\}^*$ und geben Sie eine kontextfreie Grammatik an, die L' erzeugt.

Lösung zu Aufgabe 4. Wichtig: Nicht für jedes $v \in L$ muss ein $w \in L'$ existieren mit $h(w) = v$! Wir stellen fest, dass das Bild jedes Wortes in $\{0, 1\}^*$ genau gleich viele a's und b's enthält. Somit ist $h^{-1}(L) = \{0, 1\}^*$. Die Sprache ist sogar regulär. Eine Grammatik ist z.B.

$$G = (\{S\}, \{0, 1\}, \{S \to 0S \mid 1S \mid \varepsilon\}, S).$$