Formale Grundlagen der Informatik I
7. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Ulrich Kohlenbach
Alexander Kreuzer
Pavol Safarik

SS 2012

Gruppenübung

Aufgabe G1
Wahr oder falsch? Begründen Sie Ihre Antwort!

(a) Reguläre Sprachen sind entscheidbar.

(b) Es gibt kontextfreie Sprachen, die regulär sind.

(c) Ist L_1 regulär und $L_2 \subseteq L_1$, so ist L_2 auch regulär.

(d) Ist L_1 regulär und L_2 beliebig, dann ist

$$L = \{x \in \Sigma^* \; : \; \text{es existiert ein } y \in L_2, \text{ so dass } xy \in L_1\}$$

regulär.

(e) Das Komplement einer Sprache, die von einer Grammatik erzeugt wird, wird auch wieder von einer Grammatik erzeugt.

(f) Sind L_1 und L_2 kontextfrei, so ist $L_1 \backslash L_2$ entscheidbar.

(g) Die Sprache aller regulären Ausdrücke ist regulär.

Lösungsskizze:

(a) Richtig. (Jede kontextsensitive Sprache ist entscheidbar und jede reguäre Sprache ist auch eine kontextsensitive Sprache.)

(b) Richtig. (Jede reguläre Sprache ist auch kontextfrei.)

(c) Falsch. (Sei $L_1 = \Sigma^*$ und L_2 eine nicht-reguläre Σ-Sprache.)

(d) Richtig. (Sei $\mathcal{A}$ ein DFA für die Sprache L_1. Wir können $\mathcal{A}$ in einem DFA $\mathcal{A}'$ für die Sprache L umwandeln, indem wir die Menge der akzeptierenden Zustände ändern: wir erklären nämlich diejenigen Zustände in $\mathcal{A}'$ zu akzeptierenden, aus denen man mit einem Wort $y \in L_2$ in einen akzeptierenden Zustand von $\mathcal{A}$ gelangt. Diese Definition ist natürlich nicht unbedingt rekursiv, d.h., es gibt Sprachen L_2, so dass der Automat für L sich nicht aus dem Automaten für L_1 berechnen lässt).

(e) Falsch. (Die Sprachen, die von Grammatiken erzeugt werden, sind gerade die rekursiv aufzählbaren Sprachen. Diese sind nicht unter Komplement abgeschlossen. Rekursiv aufzählbare Mengen, deren Komplement auch rekursiv aufzählbar ist, sind gerade die entscheidbaren Mengen: siehe Beobachtung 4.3.3 auf Seite 66 im Skript.)

(f) Richtig. (Kontextfrei Sprachen sind entscheidbar und entscheidbare Sprachen sind unter allen Booleschen Operationen abgeschlossen.)

(g) Falsch. (Folgt aus einer analogen Anwedung des Pumping Lemmas, die wir genutzt haben um zu zeigen, dass die Sprache der Klammerwörter nicht regulär ist.)

Aufgabe G2

Zeigen Sie, dass die Klasse der rekursiv aufzählbaren Σ-Sprachen unter Durchschnitt und Vereinigung abgeschlossen ist.
Folgerns sie daraus, dass auch die Klasse der entscheidbaren Σ-Sprachen unter Durchschnitt und Vereinigung abgeschlossen sind.
Hinweis: Sie können dafür die Tatsache benutzten, dass es für zwei beliebige Turingmaschinen immer eine Turingmaschine gibt, die diese beiden parallel simuliert. Diese Turingmaschine kann man aus der Beschreibung der beiden anderen explizit konstruieren (was aufwendig ist). Unter Annahme der Church-Turing-These ist klar, dass eine solche Konstruktion existieren muss, denn auch ein Computer kann die Ausführung von zwei Programmen parallel simulieren.

Lösungsskizze:
Nach Definition 4.3.2 heißt $L \subseteq \Sigma^*$ aufzählbar, wenn L von einer Turingmaschine akzeptiert wird. Nehmen wir deshalb an, dass $K \subseteq \Sigma^*$ von einer Turingmaschine $\mathcal{M}_K$ und $L \subseteq \Sigma^*$ von einer Turingmaschine $\mathcal{M}_L$ akzeptiert werden.
Die folgende Maschine $\mathcal{M}$ akzeptiert die Sprache $K \cap L$: auf einer beliebigen Eingabe $w \in \Sigma^*$ simuliert die Maschine $\mathcal{M}$ gleichzeitig die Berechnungen von $\mathcal{M}_K$ und $\mathcal{M}_L$ auf dieser Eingabe. Wenn beide Simulationen in akzeptierenden Zuständen terminieren, dann akzeptiert $\mathcal{M}$ die Eingabe. Andernfalls gilt die Eingabe als verworfen. (Dies gilt insbesondere wenn eine der beiden Simulationen nicht terminiert).
Die Sprache $K \cup L$ wird von der folgenden Maschine $\mathcal{M}'$ akzeptiert: Auch diese Maschine simuliert die beiden Maschinen $\mathcal{M}_K$ und $\mathcal{M}_L$, nur die Akzeptanzbedingung ändert sich. Sobald eine der beiden Simulationen in einem akzeptierenden Zustand terminiert, terminiert auch $\mathcal{M}'$ und akzeptiert die Eingabe. Wenn dieser Fall nie eintritt, gilt die Eingabe als verworfen.
Um die entsprechen Abschlüsse für entscheidbare Sprachen zu folgern wenden wir Beobachtung 4.3.3. aus dem Skript an: Sind $K, L \subseteq \Sigma^*$ entscheidbare Sprachen, dann ist $K, \overline{K}, L, \overline{L}$ aufzählbar. Nach dem oben gezeigten gilt dann auch, dass $K \cap L$ und $\overline{K \cap L} = \overline{K} \cup \overline{L}$ aufzählbar sind also folgt mit dem Satz von Post, dass $K \cap L$ entscheidbar ist.
Die Entscheidbarkeit von $K \cup L$ kann man analog zeigen.

Aufgabe G3

Bestimmen Sie eine kontextfreie Grammatik für die arithmetischen Terme über $(\mathbb{N}, +, \cdot, 0, 1)$, die minimal geklammert sind (d.h., wir berücksichtigen die Assoziativität der Addition und Multiplikation und geben der Multiplikation Priorität über die Addition).

Lösungsskizze: Wir setzen $\Sigma = \{0, 1, +, \cdot, (,)\}$ und benutzen die Nichtterminale $\{S, B, A, M\}$. S ist das Startsymbol, B steht für Terme die nur aus einer der Konstanten 0 oder 1 bestehen, A steht für alle Terme, die Summen kleinerer Terme sind, und M steht für alle Terme, die Produkte kleinerer Terme sind. Wir benutzen die folgenden Produktionen:

$$\begin{array}{lcl} S & \to & B \,|\, A \,|\, M \\ B & \to & 0 \,|\, 1 \\ A & \to & A + A \,|\, A + B \,|\, B + A \,|\, B + B \,|\, M + A \,|\, A + M \,|\, M + M \,|\, M + B \,|\, B + M \\ M & \to & M \cdot M \,|\, M \cdot B \,|\, B \cdot M \,|\, B \cdot B \,|\, M \cdot (A) \,|\, (A) \cdot M \,|\, (A) \cdot (A) \,|\, (A) \cdot B \,|\, B \cdot (A) \end{array}$$

Aufgabe G4

Sei $L \subseteq \{a\}^*$ eine kontextfreie Sprache über einem einelementigen Alphabet.

(a) Zeigen Sie, dass man L darstellen kann als endliche Vereinigung

$$L = L_0 \cup L_{k_1,p_1} \cup \cdots \cup L_{k_m,p_m} \,,$$

wobei L_0 eine endliche Sprache ist und

$$L_{k,p} := \{ a^{k+ip} : i \in \mathbb{N} \} .$$

(b) Folgern Sie hieraus, dass jede kontextfreie Sprache über einem einelementigen Alphabet regulär ist.

Hinweis zu (a): Benutzen Sie das Pumping Lemma um zu zeigen, dass es eine Zahl $n \in \mathbb{N}$ gibt, so dass

$$L = L_0 \cup L_{k_1,p_1} \cup L_{k_2,p_2} \cup \ldots \quad \text{mit } p_1, p_2, \ldots \leq n ,$$

wobei Sie zunächst unendlich viele Sprachen der Form $L_{k,p}$ zulassen. Argumentieren Sie im zweiten Schritt, wieso Sie mit endlich vielen solchen Sprachen auskommen.

Lösungsskizze:

(a) Sei n die Konstante aus dem Pumping Lemma. Wir setzen

$$\begin{aligned} L_0 &:= \{ w \in L : |w| < n \} \text{ und} \\ S &:= \{k \in \mathbb{N} \,:\, a^k \in L \text{ und } k \geq n\}. \end{aligned}$$

Für jedes $k \in S$ gibt es eine Zerlegung $a^k = yuvwz$ mit $|uvw| \leq n$, $uw \neq \varepsilon$ und $yu^ivw^iz \in L$ für alle $i \in \mathbb{N}$. Setze $p_k := |uw|$. Dann folgt, dass

$$yu^ivw^iz = a^{k+ip_k} \in L .$$

Also ist $L_{k,p_k} \subseteq L$ und

$$L = L_0 \cup \bigcup_{k \in S} L_{k,p_k} .$$

Sind $k, l \in S$ zwei Zahlen mit

$$p_k = p_l , \quad k \equiv_{p_k} l \quad \text{und} \quad k < l ,$$

dann gilt $L_{l,p_l} \subseteq L_{k,p_k}$. Wir können also L_{l,p_l} aus der Vereinigung entfernen. Deswegen gilt

$$L = L_0 \cup \bigcup_{k \in S_0} L_{k,p_k} ,$$

wobei

$$S_0 := \left\{ k \in S \,:\, \text{es gibt kein } l \in S \text{ mit } l < k ,\ p_l = p_k \text{ und } k \equiv_{p_k} l \right\} .$$

Da $p_k \leq n$ ist, folgt, dass S_0 weniger als n^2 Elemente enthält. Insbesondere ist die obige Vereinigung also endlich.

(b) Endliche Sprachen und jede Sprache der Form $L_{k,p}$ sind regulär:

$$L_{k,p} = L(a^k(a^p)^*) .$$

Da reguläre Sprachen unter Vereinigung abgeschlossen sind, folgt, dass

$$L = L_0 \cup L_{k_1,p_1} \cup \cdots \cup L_{k_n,p_n}$$

ebenfalls regulär ist.

Hausübung

Aufgabe H1 (4+3+2+2 Punkte)

Sei $\Sigma = \{a, b\}$. Zur Erinnerung: Im Post'schen Korrespondenzproblem (PKP) ist eine *Instanz*, d.h. eine endliche Folge von Wortpaaren $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$ mit $x_i, y_i \in \Sigma^+$, gegeben. Gefragt ist, ob es eine Folge von Indizes $i_1, i_2, \ldots, i_n \in \{1, 2, \ldots, k\}$ $(n \geq 1)$ gibt, mit

$$x_{i_1} x_{i_2} \ldots x_{i_n} = y_{i_1} y_{i_2} \ldots y_{i_n}.$$

Wenn eine solche Folge existiert, so heißt diese eine *Lösung* der PKP-Instanz $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$. Insbesondere ist die Frage, ob eine beliebige Instanz des PKP eine Lösung hat, semi-entscheidbar, aber *nicht* entscheidbar.

(a) Entscheiden Sie für die folgenden Instanzen der PKP, ob es eine Lösung gibt. Falls es eine Lösung gibt, geben Sie diese an. Falls nicht, beweisen Sie, dass es keine Lösung gibt.
   i. $(abba, a)$, $(aa, bbaa)$, $(bbbb, abb)$
   ii. (aa, a), (ba, ab), (b, abb)

(b) Geben Sie eine Grammatik an, die genau alle endlichen Folgen von Wortpaaren $(x_1, y_1), \ldots, (x_k, y_k)$ mit $x_i, y_i \in \Sigma^+$ (d.h. die syntaktisch korrekten Instanzen des PKP) erzeugt.

(c) Gibt es eine Grammatik, die genau alle solcher Instanzen des PKP erzeugt, die eine Lösung haben?

(d) Gibt es eine Grammatik, die genau alle solcher Instanzen des PKP erzeugt, die *keine* Lösung haben?

Lösungsskizze:

(a) i. Eine Lösung der Instanz ist $(1, 2, 3, 2)$.

   ii. Angenommen es gibt eine Lösung $(i_1, \ldots, i_n)$ mit $\mathbf{x} = \mathbf{y}$, wobei $\mathbf{x} = x_{i_1} \ldots x_{i_n}$ und $\mathbf{y} = y_{i_1} \ldots y_{i_n}$. Da beide Worte mit dem selben Buchstaben beginnen müssen, ist $i_1 = 1$. Nun muss y_{i_2} mit a beginnen, sprich $i_2 \in \{2, 3\}$. Ist $i_2 = 2$, so muss y_{i_3} ebenfalls mit a beginnen und es folgt wiederum $i_3 \in \{2, 3\}$. Stets $i_k = 2$ für alle $k \geq 2$ zu wählen führt zu *keiner* endlichen Lösung der Instanz. Entsprechend (da die Existenz einer Lösung vorausgesetzt wurde) muss es ein i_k mit $k \geq 2$ geben, so dass $i_k = 3$ gewählt wurde. Im Falle $k = 2$ war $i_1 = 1$, womit nun x_{i_3} mit b beginnen muss. Wird nun $i_3 = 3$ gewählt, so gibt es fortan keine Fortsetzung mehr, so dass die Instanz eine Lösung besitzt, denn es müsste x_{i_4} mit abb beginnen, was nicht möglich ist.
   Im Falle $k > 2$ muss $x_{i_{k+1}}$ mit b beginnen. Wird $i_{k+1} = 2$ gewählt, so kann wieder unendlich fortgesetzt werden, ohne eine endliche Lösung zu erhalten. Wird hingegen $i_{k+1} = 3$ gewählt, so müsste wiederum $x_{i_{k+2}}$ mit abb beginnen, was nicht möglich ist. Entsprechend hat diese Instanz *keine* Lösung.

(b) Eine mögliche Grammatik ist $G = (\Sigma, V, P, S)$ mit $V = \{S, T, U\}$ und den Produktionen

$$\begin{aligned} S &\to T\#S \mid T \\ T &\to U\$U \\ U &\to a_i U \mid a_i \quad (\text{für alle } a_i \in \Sigma) \end{aligned}$$

(c) Ja, denn das PKP ist semi-entscheidbar, womit es eine Typ-0 Grammatik gibt, die alle Instanzen des PKP erzeugt, die eine Lösung haben.

(d) Nein, sonst wäre das PKP entscheidbar – und damit insbesondere auch das wohlbekannte (unentscheidbare) Halteproblem.