Prof. J. Esparza
S. Sickert, J. Krämer

Technische Universität München
KEINE ABGABE

Einführung in die theoretische Informatik

Sommersemester 2017 – Übungsblatt Lösungsskizze 12

Übungsblatt

Wir unterscheiden zwischen Übungs- und Abgabeblättern. Auf diesem *Übungsblatt* finden Sie eine Übersicht über die Kernaspekte, die Sie in Kalenderwoche 29 in den Tutorien diskutieren, üben und vertiefen. Die Aufgaben auf diesem Blatt dienen dem Üben und Verstehen des Vorlesungsstoffes, sowie dem *eigenständigen Erarbeiten* der Kernaspekte. Außerdem sollen Ihnen diese Aufgaben auch helfen, ein Gefühl dafür zu bekommen, was Sie inhaltlich in der Klausur erwartet. Klausuraufgaben können jedoch deutlich von den hier gestellten Aufgaben abweichen. Abschreiben und Auswendiglernen von Lösungen wird Ihnen daher keinen dauerhaften Erfolg in der Vorlesung bringen. Fragen zu den Übungsblättern können Sie montags bis donnerstags von 12 Uhr bis 14 Uhr in der *THEO-Sprechstunde* in Raum 03.11.034 stellen.

Kernaspekte

K12.1 korrektes Wiedergeben der folgenden Definitionen und Algorithmen

- PCP

K12.2 mithilfe des Satzes von Rice zeigen, dass eine Sprache unentscheidbar ist
K12.3 berechnen von Lösungen für PCP-Instanzen
K12.4 beweisen, dass PCP-Instanzen keine Lösung besitzen
K12.5 Aussagen über Reduktionen beweisen oder widerlegen
K12.6 Fehler in inkorrekten Reduktionen erkennen
K12.7 begründet entscheiden, ob gegebene Beispiele neu eingeführte Definitionen erfüllen
K12.8 Aussagen mit neu eingeführten Definitionen beweisen oder widerlegen

AUFGABE 12.1.

Stufe B

Sei $\mathsf{H}_0 \coloneqq \{w \in \{0,1\}^* \mid \mathsf{M}_w[\varepsilon] \downarrow\}$ und sei $A \coloneqq \{w \in \Sigma^* \mid \exists i \in \mathbb{N}_0.\ |w| = 5i + 3\}$ mit $\Sigma = \{\mathsf{a}, \mathsf{b}\}$. Erklären Sie, warum die angegebenen Funktionen keine Reduktionen gemäß Vorlesungsdefinition sind.

(a) Behauptung: $\mathsf{H}_0 \leq A$
Reduktion:

$$f(w) = \begin{cases} \mathsf{aaa} & \text{falls } w \in \mathsf{H}_0 \\ \mathsf{b} & \text{sonst} \end{cases}$$

(b) Behauptung: $A \leq \mathsf{H}_0$
Reduktion: f bildet jedes Element $x \in \Sigma^*$ auf die Kodierung einer TM M_x, die wie folgt definiert ist: Die TM M_x löscht die Eingabe und schreibt x aufs Band, bestimmt dann die Länge von x, zieht 3 ab und prüft anschließend, ob das Ergebnis durch 5 teilbar ist. Dementsprechend gibt die Maschine "Ja" (1) und "Nein" (0) aus.

(c) Behauptung: $\overline{\mathsf{H}_0} \leq \mathsf{H}_0$
Reduktion: f bildet jedes $w \in \{0,1\}^*$ auf die Kodierung $f(w)$ einer TM $\mathsf{M}_{f(w)}$ ab, die $\mathsf{M}_w[\varepsilon]$ simuliert. Falls M_w hält, geht $\mathsf{M}_{f(w)}$ in eine Endlosschleife. Falls $\mathsf{M}_w[\varepsilon]$ nicht hält, hält $\mathsf{M}_{f(w)}$.

(d) Behauptung: $H_{\Sigma^*} \leq \mathsf{H}_0$ mit $H_{\Sigma^*} = \{w \in \{0,1\}^* \mid \forall x \in \Sigma^*.\ \mathsf{M}_w[x] \downarrow\}$.
Reduktion: f bildet jedes $w \in \{0,1\}^*$ auf die Kodierung $f(w)$ einer TM $\mathsf{M}_{f(w)}$ ab, die erst die Eingabe löscht und nicht deterministisch $x \in \Sigma^*$ erzeugt und dann $\mathsf{M}_w[x]$ simuliert.

Stufe B

Lösungsskizze

(a) f ist unberechenbar, da H_0 unentscheidbar ist und somit χ_{H_0} unberechenbar ist.
(b) f bildet auf Kodierungen von Turing-Maschinen ab, die immer terminieren. Da $\mathsf{a} \notin A$, aber $f(\mathsf{a}) \in \mathsf{H}_0$, erfüllt die Funktion f nicht die Definition einer Reduktion.
(c) f ist nicht wohldefiniert. Wenn $\mathsf{M}_{f(w)}$ die Berechnung von $\mathsf{M}_w[\varepsilon]$ simuliert und $\mathsf{M}_w[\varepsilon]$ nicht hält, dann hält definitiv $\mathsf{M}_{f(w)}$ auch nicht.
(d) Sei M irgendeine Turing-Maschine mit $\mathsf{M}[\varepsilon]\downarrow$ und $\neg\mathsf{M}[0]\downarrow$ und sei $w \in \Sigma^*$ mit $\mathsf{M}_w = \mathsf{M}$. Dann gilt $w \notin H_{\Sigma^*}$ und $f(w) \in \mathsf{H}_0$.

AUFGABE 12.2. (*Reduktionen*)
Entscheiden Sie, ob folgende Aussagen korrekt oder inkorrekt sind, und begründen Sie Ihre Antwort, indem Sie einen Beweis bzw. ein passendes Gegenbeispiel angeben. Sei $\Sigma = \{0, 1\}$.
(a) $\forall A \subseteq \Sigma^*.\ A \leq \Sigma^*$
(b) $\forall A, B \subseteq \Sigma^*.\ A \leq B \iff \overline{A} \leq \overline{B}$
(c) $\forall A \subseteq \Sigma^*.\ A \neq \emptyset \wedge A \neq \Sigma^* \implies A \leq \overline{A}$
(d) $\forall A, B, C \subseteq \Sigma^*.\ A \leq B \wedge B \leq C \implies A \leq C$

Lösungsskizze

(a) Falsch. Sei $A = \emptyset$. Damit $\overline{A} = \Sigma^*$. Dann muss für eine Reduktionsfunktion f gelten: $\forall x \in \overline{A}.\ f(x) \notin \Sigma^*$. Eine solche Funktion f existiert aber nicht.
(b) Wahr. Gelte $A \leq B$. Dann existiert ein totales und berechenbares f mit: $\forall x \in \Sigma^*.x \in A \Leftrightarrow f(x) \in B$. Somit gilt auch: $\forall x \in \Sigma^*.x \in \overline{A} \Leftrightarrow f(x) \in \overline{B}$. Daraus folgt dann $\overline{A} \leq \overline{B}$. Die Rückrichtung geht analog.
(c) Falsch. Sei $A = \mathsf{H}_0$. Nach Vorlesung wissen wir H_0 ist semi-entscheidbar und $\overline{\mathsf{H}_0}$ ist nicht semi-entscheidbar. Insbesondere sind damit die Mengen nicht trivial. Angenommen (c) gilt, dann haben wir $\mathsf{H}_0 \leq \overline{\mathsf{H}_0}$ und mit (b) auch $\overline{\mathsf{H}_0} \leq \mathsf{H}_0$. Dann sind aber beide Menge semi-entscheidbar und somit H_0 auch entscheidbar. Widerspruch!
(d) Wahr. Seien f und g Reduktionen von $A \leq B$ und $B \leq C$. Sei nun $h(x) = g(f(x))$. h ist total und berechenbar, da f und g total und berechenbar sind. Sei $x \in A$ beliebig. Dann gilt $f(x) \in B$ und $h(x) = g(f(x)) \in C$. Sei nun $x \notin A$ beliebig. Dann gilt $f(x) \notin B$ und damit $h(x) = g(f(x)) \notin C$. Somit ist h eine geeignete Reduktion.

AUFGABE 12.3. (*PCP*) Stufe B - D
Wir betrachten in dieser Aufgabe das Post'sche Korrespondenzproblem (PCP).
(a) Bestimmen Sie *alle* Lösungen für das folgende PCP: $P_1 = ((d, cd), (d, d), (abc, ab))$.
(b) Zeigen Sie, dass die folgende Instanz des PCPs keine Lösung hat: $P_2 = ((ab, aba), (baa, aa), (aba, baa))$.
(c) Zeigen Sie, dass das Post'sche Korrespondenzproblem über einem Alphabet mit nur einem Symbol entscheidbar ist, indem Sie einen Algorithmus angeben. Begründen Sie auch dessen Korrektheit.
(d) Sei $P = (c_1, c_2)$ ein PCP über einem beliebigem Alphabet Σ mit $c_i = (x_i, y_i)$ und $||x_i| - |y_i|| = 1$ für $i \in \{1, 2\}$. Zeigen Sie die Entscheidbarkeit für diese Variante des PCPs. Geben Sie hierzu einen Algorithmus an und begründen Sie dessen Korrektheit.

Lösungsskizze

(a) Die Menge aller Lösungen ist: $\mathsf{L}((2 \mid (31))^*) \setminus \{\varepsilon\}$
(b) Zu Beginn kann nur die Karte (ab, aba) verwendet werden, da 2 und 3 als Startkarte ungeeignet sind und somit alle Lösungen mit 1 beginnen müssen. Wir müssen deshalb mit dem Überhang (ε, a) fortfahren. Offensichtlich kann nun weder (ab, aba) (Karte 1) noch (baa, aa) (Karte 2) angewendet werden. Durch (aba, baa) (Karte 3) erhalten wir aber $(aba, abaa)$, was wiederum zu dem Überhang (ε, a) führt. Wir können somit keinen Abschluss finden und damit kann diese Instanz des Post'schen Korrespondenzproblems keine Lösung besitzen.
(c) In diesem Fall haben alle Karten c die Form $c = (a^i, a^j)$ für $\Sigma = \{a\}$ und $i, j \geq 0$.
- Falls $i = j$, ist die Karte c alleine eine Lösung.
- Wenn alle Karten oben länger als unten $(i > j)$ sind, gibt es keine Lösung.
- Analog für den umgekehrten Fall $(i < j)$.
- Wenn es zwei Karten c_1, c_2 mit $c_1 = (a^{j_1}, a^{k_1})$, $c_2 = (a^{j_2}, a^{k_2})$, $j_1 > k_1$ und $j_2 < k_2$ gibt, sei i_1 der Index von c_1 und sei i_2 der von c_2. Dann ist $i_1^{k_2 - j_2} i_2^{j_1 - k_1}$ eine Lösung des PCP.

Eine TM kann diese Vorbedingungen prüfen und somit bestimmen, ob es eine Lösung gibt.
(d) Entscheidbar, da es genau dann eine Lösung gibt, wenn 12 oder 21 eine Lösung ist, und diese beiden Fälle von einer TM geprüft werden können.
Beweis:
Sei $i_1 i_2 \dots i_k$ eine kürzeste Lösung von P. Aus der Längenbedingung der Karten folgt sofort, dass $k \geq 2$ gilt. Falls $i_1 \neq i_2$, dann gilt $|x_{i_1} x_{i_2}| = |y_{i_1} y_{i_2}|$. Somit ist bereits $i_1 i_2$ (12 oder 21) eine Lösung.
Sei nun $i_1 = i_2$ und o.B.d.A. $i_1 = i_2 = 1$. Wir nehmen ebenfalls o.B.d.A. an, dass $|x_1| > |y_1|$ gilt. Somit $x_1 = u_1 \dots u_n u_{n+1}$ und $y_1 = u_1 \dots u_n$ mit $u_i \in \Sigma$ gilt. Da $i_1 i_2$ Teil einer Lösung ist, gilt für $(x_1 x_1, y_1 y_1)$:

$$x_1 x_1 = y_1 y_1 u_n u_{n+1}$$

und somit

$$u_{n+1} = u_1 = u_2 = \cdots = u_{n-1} = u_n$$

Die Karte c_1 hat somit die Gestalt $c_1 = (a^{n+1}, a^n)$ für irgendeinen Buchstaben $a \in \Sigma$.
In der Lösung muss es auch eine solche Teilsequenz (22) für c_2 geben, damit der Überhang von c_1 ausgeglichen wird. Analog folgt dann $c_2 = (b^m, b^{m+1})$ für ein $b \in \Sigma$. Da c_1 und c_2 sich überschneiden, gilt $a = b$. Somit ist auch 12 eine Lösung ist.

AUFGABE 12.4. Stufe C

Entscheiden Sie, ob die folgenden Mengen unentscheidbar für $\Sigma = \{0, 1\}$ sind, und begründen Sie Ihre Antworten mit dem Satz von Rice (falls anwendbar). Geben Sie dabei die Menge $\mathcal{F}$ genau an und argumentieren Sie, warum die Menge nicht trivial ist.

- $\mathsf{L}_1 = \{w \in \Sigma^* \mid \{u \in \Sigma^* \mid \varphi_w(u) = 1\} \text{ ist regulär}\}$
- $\mathsf{L}_2 = \{w \in \Sigma^* \mid \forall n \in \mathbb{N}_0.\ \varphi_w(n) = n * (n - 23) + 42\}$
- $\mathsf{L}_3 = \{w \in \Sigma^* \mid \forall p \in \mathbb{N}_0.\ (|w| > p \wedge p \text{ ist prim}) \to w_p = 0\}$

Hinweis: $w_p \in \Sigma$ bezeichnet den Buchstaben an der p-ten Stelle im Wort w.

Lösungsskizze

- Sei $\mathcal{F} = \{f \mid f \text{ ist berechenbar } \wedge f^{-1}(1) \text{ ist regulär}\}$. Sei nun $g(w) = 1$ und

$$h(w) = \begin{cases} 1 & \text{falls } \exists i \geq 0.\ w = 0^i 1^i \\ 0 & \text{sonst} \end{cases}$$

  zwei berechenbare Funktionen. Dann gilt $g \in \mathcal{F}$ und $h \notin \mathcal{F}$. Somit ist $\mathcal{F}$ nicht die Menge aller berechenbarer Funktionen. Damit folgt aus dem Satz von Rice, dass L_1 unentscheidbar ist.
- Sei $\mathcal{F} = \{f \mid f \text{ ist berechenbar } \wedge \forall n \in \mathbb{N}_0.\ f(n) = n * (n - 23) + 42\}$. Dann gilt für $g(n) = 0$: $g \notin \mathcal{F}$ und somit ist $\mathcal{F}$ nicht die Menge aller berechenbarer Funktionen. Weiterhin ist $\mathcal{F}$ auch nicht leer, da das Polynom in der Definition berechenbar ist. Somit ist nach Satz von Rice L_2 unentscheidbar.
- L_3 ist entscheidbar, da w nur syntaktischen Kriterien erfüllen muss. Eine TM kann alle Primzahlen kleiner gleich $|w|$ berechnen und an diesen Stellen in w prüfen, ob $w_p = 0$ gilt.

AUFGABE 12.5. (*Semi-Entscheidbarkeit*) Stufe D

Sei $A \subseteq \Sigma^*$. Zeigen Sie die folgende Behauptung:

A ist semi-entscheidbar gdw. A ist Wertebereich einer berechenbaren Funktion

Lösungsskizze
A semi-entscheidbar $\rightsquigarrow$ A rekursiv aufzählbar $\rightsquigarrow$ A Wertebereich eine berechenbaren Funktion (genau die Funktion aus Definition 5.44, die A aufzählt)
A Wertebereich einer berechenbaren Funktion f $\rightsquigarrow$ sei T TM zu f, simuliere T für ansteigende Grenze N auf allen Eingaben der Länge $\leq N$ für N Schritte, falls Simulation terminiert, gib berechneten Wert aus $\rightsquigarrow$ A aufzählbar, da schließlich jede Eingabe und damit jeder Ausgabe erzeugt wird $\rightsquigarrow$ A semi-entscheidbar.

AUFGABE 12.6. (*Entscheidbarkeit und kontextfreie Grammatiken*) Stufe D

Seien $\mathsf{G}_1, \mathsf{G}_2$ CFGs. Beweisen Sie die folgenden beiden Aussagen:

(a) $\mathsf{L}(\mathsf{G}_1) \not\subseteq \mathsf{L}(\mathsf{G}_2)$ ist semi-entscheidbar.
(b) $\mathsf{L}(\mathsf{G}_1) \subseteq \mathsf{L}(\mathsf{G}_2)$ ist unentscheidbar.
Hinweis: In der Vorlesung wurde das Resultat nur erwähnt, zeigen Sie das Resultat jetzt formal. Sie dürfen verwenden, dass für CFGs $\mathsf{G}_1, \mathsf{G}_2$ das Problem $\mathsf{L}(\mathsf{G}_1) \cap \mathsf{L}(\mathsf{G}_2) = \emptyset$ unentscheidbar ist. Schauen Sie sich den entsprechenden Beweis in den Folien an. Charakterisieren Sie die dort verwendeten CFGs möglichst genau nach linear, rechtslinear, linkslinear und deterministisch. Denken Sie außerdem daran, dass $\mathsf{L}(\mathsf{G}_1) \subseteq \mathsf{L}(\mathsf{G}_2) \Leftrightarrow \mathsf{L}(\mathsf{G}_1) \cap \overline{\mathsf{L}(\mathsf{G}_2)} = \emptyset$.

Lösungsskizze

(a) Es gilt $\mathsf{L}(\mathsf{G}_1) \not\subseteq \mathsf{L}(\mathsf{G}_2)$ genau dann, wenn es ein $w \in \mathsf{L}(\mathsf{G}_1) \setminus \mathsf{L}(\mathsf{G}_2)$ gibt.
Für $w \in \Sigma^*$: Für jedes w testet man mittels CYK (o.B.d.A. sind G_1 und G_2 in CNF), ob $w \in \mathsf{L}(\mathsf{G}_1)$ und ob $w \in \mathsf{L}(\mathsf{G}_2)$ gilt. Sobald man das erste $w \in \mathsf{L}(\mathsf{G}_1) \setminus \mathsf{L}(\mathsf{G}_2)$ gefunden hat, stoppt man und gibt 1 aus. Offensichtlich stoppt der Algorithmus im Fall $\mathsf{L}(\mathsf{G}_1) \not\subseteq \mathsf{L}(\mathsf{G}_2)$ stets, im Fall $\mathsf{L}(\mathsf{G}_1) \subseteq \mathsf{L}(\mathsf{G}_2)$ terminiert der Algorithmus allerdings nie. Damit ist das Problem semi-entscheidbar.

(b) Es gilt $\mathsf{L}(\mathsf{G}_1) \subseteq \mathsf{L}(\mathsf{G}_2)$ **gdw** $\mathsf{L}(\mathsf{G}_1) \setminus \mathsf{L}(\mathsf{G}_2) = \emptyset$ **gdw** $\mathsf{L}(\mathsf{G}_1) \cap \overline{\mathsf{L}(\mathsf{G}_2)} = \emptyset$. (O.B.d.A. verwenden G_1 und G_2 dasselbe Alphabet Σ.)
Sei $(x_1, y_1), \ldots, (x_l, y_l) \in \Gamma^* \times \Gamma^*$ eine PCP-Instanz. Sei $\Sigma = \{a_1, \ldots, a_l\}$, o.B.d.A. $\Gamma \cap \Sigma = \emptyset$ und $\mathsf{A} = \Sigma \cup \Gamma$, da wir Σ frei wählen können. Dann sind die in der Vorlesung verwendeten Grammatik $\mathsf{G}_1, \mathsf{G}_2$ linear und die erzeugten Sprachen sogar deterministisch: Da DCFL unter Komplement nach Vorlesung abgeschlossen sind, kann man aus G_2 bzw. dem entsprechenden DPDA eine CFG G_2' mit $\mathsf{L}(\mathsf{G}_2') = \overline{\mathsf{L}(\mathsf{G}_2)}$ konstruieren.
Damit gilt: $\mathsf{L}(\mathsf{G}_1) \subseteq \mathsf{L}(\mathsf{G}_2')$ **gdw** $\mathsf{L}(\mathsf{G}_1) \cap \overline{\mathsf{L}(\mathsf{G}_2')} = \emptyset$ **gdw** $\mathsf{L}(\mathsf{G}_1) \cap \mathsf{L}(\mathsf{G}_2) = \emptyset$ **gdw** die gegebene PCP-Instanz hat keine Lösung. Somit ist die Teilmengenrelation über CFGs unentscheidbar.