Formale Grundlagen der Informatik I 7. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Martin Ziegler
Alexander Kreuzer
Carsten Rösnick

SS 2011
25.05.11

Minitest Lösung

a) Das Wortproblem für kontextfreie Sprachen ist entscheidbar. ☒ Richtig ☐ Falsch

Begründung: Sei $L \subseteq \Sigma^*$ eine kontextfreie Sprache. Falls $\varepsilon \notin L$ hat L eine Grammatik in Chomsky-Normalform (Satz 3.3.2) und das Wortproblem kann mit dem CYK-Algorithmus (3.3.13) entschieden werden. Falls $\varepsilon \in L$, dann kann das Wortproblem für $L \setminus \{\varepsilon\}$, wie oben, mit dem CYK-Algorithmus entschieden werden und das Wortproblem für $\{\varepsilon\}$ ist trivial.

Alternativ kann man auch einen PDA konstruieren, der die Sprache erkennt (Satz 4.1.5).

b) Gegeben sei die Turingmaschine $\mathcal{M}$ über dem Alphabet $\{a, b\}$ mit Startzustand q_0 und folgender Übergangsfunktion:

δ	$\Box$	a	b
q_0	$(\Box, >, q_a)$		
q_a	$(\Box, \circ, q^+)$	$(a, >, q_b)$	$(b, \circ, q^-)$
q_b	$(\Box, \circ, q^-)$	$(a, \circ, q^-)$	$(b, >, q_a)$

Hält $\mathcal{M}$ bei jeder Eingabe? Und wenn ja, welche Sprache erkennt $\mathcal{M}$?

☐ $\mathcal{M}$ hält immer und erkennt $L((a + b)^*)$.
☒ $\mathcal{M}$ hält immer und erkennt $L((ab)^*)$.
☐ $\mathcal{M}$ hält nicht bei jeder Eingabe.

Begründung: Die DTM hält immer: In jedem Schritt, in dem $\mathcal{M}$ noch nicht hält, wird der Kopf nach rechts bewegt. D.h. irgendwann muss $\Box$ eingelesen werden, weil die Eingabe zu Ende ist. Nun gilt, dass $\mathcal{M}$ nur im ersten Schritt in q_0 ist, d.h. irgendwann wird $\Box$ in dem Zustand q_a oder q_b eingelesen und damit geht die DTM in q^+ oder q^- und hält (wenn sie nicht schon vorher gehalten hat).

Die Maschine entscheidet $L((ab)^*)$, denn anhand von δ sieht man, dass für ein a ohne folgendes b die DTM in den nicht-akzeptierenden Zustand q^- übergeht, wobei wegen $\delta(q_a, a) = (a, >, q_b)$ nach Lesen eines a stets in den Zustand q_b gewechselt wird, in dem (wie eben begründet) als nächstes Zeichen nur ein b erlaubt ist.

c) Für das Komplement $\overline{H} = \Sigma^* \setminus H$ des Halteproblems gilt:

☐ $\overline{H}$ ist entscheidbar
☒ $\overline{H}$ ist nicht aufzählbar
☐ es gibt eine Typ 3 Grammatik G mit $L(G) = \overline{H}$

Begründung: Für die ersten beiden Punkte, siehe Skript S. 67. Gäbe es eine solche Typ 3 Grammatik, die $\overline{H}$ akzeptiert, so könnte man sie in eine Typ 3 Grammatik überführen, die H akzeptiert. Damit wäre H entscheidbar, was im Widerspruch zum Beweis aus der Vorlesung steht.

Gruppenübung

Aufgabe G1 (Konstruktion von Turingmaschinen)
Beschreiben Sie formal (d.h. unter Angabe der Komponenten Σ, Q, δ) jeweils eine 1-Band Turingmaschine für die folgenden Funktionen:

(a) Bei Eingabe der Binärdarstellung einer Zahl $n \in \mathbb{N}$ soll die Binärdarstellung der Zahl $2 \cdot n$ berechnet werden.

(b) Bei Eingabe der Binärdarstellung einer Zahl $n \in \mathbb{N}$ soll die Binärdarstellung der Zahl $n+3$ berechnet werden. Ihre Turingmaschine soll dabei, neben den Zuständen q_0 und q^+ nur höchstens drei weitere Zustände enthalten.

Tipp: Nehmen Sie zur Vereinfachung an, dass Sie die Zahl n in der Form $b_0 b_1 \dots b_k$ (von links nach rechts geschrieben) auf dem Band vorliegen haben, wobei $n = \sum_{i=0}^{k} b_i \cdot 2^i$, $b_i \in \{0, 1\}$.

Lösungsskizze:

(a) Definiere $\mathcal{M} = (\Sigma, Q, q_0, \delta, q^+, q^-)$ durch

$$\begin{aligned} \Sigma &= \{0, 1\}, \\ Q &= \{q_0, q^+, q^-, q_a, q_b\} \end{aligned}$$

mit folgender Übergangsfunktion:

δ	$\Box$	0	1
q_0	$(\Box, >, q_0)$	$(0, >, q_a)$	$(0, >, q_b)$
q_a	$(\Box, \circ, q^+)$	$(0, >, q_a)$	$(0, >, q_b)$
q_b	$(1, \circ, q^+)$	$(1, >, q_a)$	$(1, >, q_b)$

(b) Definiere $\mathcal{M} = (\Sigma, Q, q_0, \delta, q^+, q^-)$ durch

$$\begin{aligned} \Sigma &= \{0, 1\}, \\ Q &= \{q_0, q^+, q^-, q_{bc}, q_b, q_c\} \end{aligned}$$

mit folgender Übergangsfunktion:

δ	$\Box$	0	1
q_0	$(\Box, >, q_b)$		
q_b	$(1, >, q_c)$	$(1, >, q_c)$	$(0, >, q_{bc})$
q_c	$(1, \circ, q^+)$	$(1, \circ, q^+)$	$(0, >, q_c)$
q_{bc}	$(0, >, q_c)$	$(0, >, q_c)$	$(1, >, q_c)$

Dies beschreibt die normale Binäraddition mit folgender Bedeutung der Zustände: Der Zustand q_b beschreibt, dass der Binärzahl auf dem Band noch ein Bit des ursprünglichen Summanden $11_2 = 3_{10}$ hinzuzuaddieren sind; Zustand q_c beschreibt den Fall eines Übertrags aus voriger Addition (carry-Bit); Zustand q_{bc} beschreibt die Kombination der Fälle von q_b und q_c.

Aufgabe G2 (Simulation von Turingmaschinen)
Beschreiben Sie informell wie eine deterministische Turingmaschine $\mathcal{M}_0$ für jeden der nachfolgenden Punkte konstruiert werden kann. Geben Sie die dazu benötigten Zustände von $\mathcal{M}_0$ an und beschreiben Sie deren Funktion.

(a) Gegeben eine feste Zeichenkette $c = c_1 \dots c_k \in \Sigma^*$. Die Maschine $\mathcal{M}_0$ lösche die Eingabe, schreibe c auf das Eingabeband, fahre zurück an die Startposition und halte.

(b) Gegeben eine feste Turingmaschine $\mathcal{M}$ und eine feste Zeichenkette $c \in \Sigma^*$. Die Maschine $\mathcal{M}_0$ lösche ihre Eingabe und rechne anschliessend weiter wie die Maschine $\mathcal{M}$ bei Eingabe der Konstanten c.

(c) Gegeben eine feste Turingmaschine $\mathcal{M}$. Die Maschine $\mathcal{M}_0$ simuliere $\mathcal{M}$ auf Eingabe $\langle\mathcal{M}\rangle$ so dass

$$\begin{aligned} w \xrightarrow{\mathcal{M}_0} \infty \quad &\Leftrightarrow \quad \langle\mathcal{M}\rangle \xrightarrow{\mathcal{M}} q^+, \\ w \xrightarrow{\mathcal{M}_0} q^- \quad &\Leftrightarrow \quad \langle\mathcal{M}\rangle \xrightarrow{\mathcal{M}} q^- \end{aligned}$$

für alle Eingaben $w \in \Sigma^*$ der Maschine $\mathcal{M}_0$. Sie dürfen hier verwenden, dass für alle Turingmaschinen $\mathcal{M}$ solch eine Kodierung $\langle\mathcal{M}\rangle$ existiert[1].

Lösungsskizze:

(a) Wir benötigen die Zustände $q_0, q^+, q_{c_1}, \ldots, q_{c_k}, q^{c\rightarrow}, q^{\leftarrow}$ mit folgender Bedeutung:

- q_0 Startzustand
- q^+ Akzeptierender Endzustand
- q_{c_i} Zustand, in dem der aktuelle Zelleninhalt mit c_i überschrieben und zum Zustand $q_{c_{i+1}}$ übergegangen wird, falls $i < k$, sonst wird zu $q^{c\rightarrow}$ übergegangen.
- $q^{c\rightarrow}$ Lösche den Rest des Eingabewortes vom Band indem so lange nach rechts gefahren und $\Box$ geschrieben wird, bis das erste $\Box$ erreicht ist. Gehe dann in den Zustand $q^{\leftarrow}$ über.
- $q^{\leftarrow}$ Fahre zur Anfangsposition des Bandes (d.h. in die durch die Konfiguration $C_{\text{start}} := (\Box, q^{\leftarrow}, \Box, c)$ beschriebene Position) und gehe in den Zustand q^+ über.

(b) Sei $\mathcal{M} = (\Sigma, Q, q'_0, \delta', q'^+, q'^-)$. Wir verwenden die (disjunkte) Vereinigung von Q mit den Zuständen aus (a). Die Zustände aus (a) sind nun mit folgender geänderter Bedeutung belegt:

- $q^{\leftarrow}$ Fahre zur Anfangsposition des Bandes und (im Unterschied zur Maschine in (a)) gehe in den Startzustand q'_0 der Maschine $\mathcal{M}$ über.
- q'_0 Da der Lesekopf der Maschine $\mathcal{M}_0$ auf der Startposition, sowie c auf dem Band steht, verfahre $\mathcal{M}_0$ nun wie $\mathcal{M}$, indem es die Übergangsfunktion von $\mathcal{M}$ ausführt.
- q'^+ Gehe in den Zustand q^+ über.
- q'^- Gehe in den Zustand q^- über.

(c) Wie in (b) mit $c = \langle\mathcal{M}\rangle$, allerdings mit dem zusätzlichen Zustand q^∞, so dass q'^+ in den Zustand q^∞ übergehe. Im Zustand q^∞ gelte $\delta_{\mathcal{M}_0}(q^\infty, z) = (z, \circ, q^\infty)$ für alle $z \in (\Sigma_{\mathcal{M}_0} \cup \{\Box\})$.

Aufgabe G3 (Chomsky-Hierarchie)
Sei $\Sigma = \{a, b, c\}$. Zu welchem Niveau der Chomsky-Hierarchie gehören die folgenden Sprachen?

$L_1 = \{w \in \Sigma^* :$ zu jedem a kann man eine spätere Stelle mit einem b finden derart, dass jedes b zu höchstens einem a gehört $\}$

$L_2 = \{w \in \Sigma^* :$ wenn in w ein a vorkommt, dann gibt es eine spätere Stelle, an der ein b steht, wobei dieses b zu mehreren a's gehören kann $\}$

Lösungsskizze:

(i) L_1 ist kontextfrei aber nicht regulär.

Angenommen L_1 wäre regulär. Sei n die Konstante aus dem Pumping Lemma und setze $x = a^n b^n$. Nach dem Lemma gibt es eine Zerlegung $x = uvw$ mit $|uv| \leq n, v \neq \varepsilon$ und $uv^m w \in L_1$ für alle $m \in \mathbb{N}$. Da $|uv| \leq n$, ist $u = a^i$ und $v = a^k$ für geeignete $i, k \leq n$. Also ist $uv^2 w = a^{n+k} b^n \notin L_1$. Widerspruch!

[1] siehe auch: Gödelnummer (`http://de.wikipedia.org/wiki/Gödelnummer`)

Ein PDA für L_1 ist $\mathcal{P} = (\Sigma, Q, q, \Delta, A, \Gamma, \#)$ mit $Q = \{q\}$, $\Gamma = \{\#, |\}$, $A = \{q\}$ und Transitionen

$$\begin{array}{llllll} \{ & (q, & \#, & \varepsilon, & \varepsilon, & q) \\ & (q, & \#, & b, & \#, & q) \\ & (q, & \#, & a, & |\#, & q) \\ & (q, & |, & a, & ||, & q) \\ & (q, & |, & b, & \varepsilon, & q) \\ & (q, & \#, & c, & \#, & q) \\ & (q, & |, & c, & |, & q) \quad \}. \end{array}$$

(ii) $L_2 = L(c^* + (a+b+c)^* bc^*)$ ist regulär.

Hausübung

Aufgabe H1 (Chomsky-Hierarchie) (8 Punkte)
Welche der folgenden Sprachen über dem Alphabet $\Sigma = \{a, b, c\}$ sind (i) regulär, (ii) kontexfrei, aber nicht regulär, oder (iii) nicht kontextfrei? Begründen Sie Ihre Antwort!

$$\begin{array}{lcl} L_1 & = & \{x \in \Sigma^* \,:\, |x|_a > 2 \cdot |x|_b\} \\ L_2 & = & \{x \in \Sigma^* \,:\, |x|_a > 2 \cdot |x|_b > 2 \cdot |x|_c\} \\ L_3 & = & \{x \in \Sigma^* \,:\, |x|_a > 2 \cdot |x|_b \text{ und } |x|_b \leq 1337\} \\ L_4 & = & \{x \in \Sigma^* \,:\, |x|_a > 2 \cdot |x|_b \text{ und } |x|_b \geq 1337\} \end{array}$$

Lösungsskizze:

L_1: L_1 ist kontextfrei, aber nicht regulär. Eine Grammatik für L_1 wäre

$$\begin{array}{llll} P: & X_0 & \to & XaX \\ & X & \to & \varepsilon \mid XX \mid aXaXb \mid aXbXa \mid bXaXa \mid aX \mid Xa \mid cX \mid Xc \end{array}$$

L_1 ist nicht regulär. Angenommen L_1 wäre regulär. Sei n die Konstante aus dem Pumping Lemma und setze $x = a^{2n+1}b^n$. Nach dem Lemma gibt es eine Zerlegung $x = uvw$ mit $|uv| \leq n, v \neq \varepsilon$ und $uv^m w \in L_1$ für alle $m \in \mathbb{N}$. Da $|uv| \leq n$ und $v \neq \varepsilon$, ist $v = a^k$ für $k \geq 1$. Also ist $uv^0 w = a^{2n+1-k}b^n \notin L_1$. Widerspruch!

Alternativ (wenn auch nicht notwendig) kann der Beweis auch durch Ausnutzung der Abgeschlossenheit regulärer Sprachen unter Homomormphismen mit $\varphi : \Sigma \to \{a, b\}^*$, $\varphi(a) = a$, $\varphi(b) = b$, $\varphi(c) = \varepsilon$ und $\varphi(L_1) = \{x \in \{a, b\}^* \,:\, |x|_a > 2 \cdot |x|_b\}$ gezeigt werden.

L_2: L_2 ist nicht kontextfrei. Angenommen L_2 wäre kontextfrei. Sei n die Konstante aus dem Pumping Lemma und setze $x = a^{2n+3}b^{n+1}c^n$ ($x \in L_2$, da $|x|_a = 2n + 3 > 2(n + 1) = 2 \cdot |x|_b$ und $2 \cdot |x|_b > 2 \cdot |x|_c \Leftrightarrow |x|_b > |x|_c$). Nach dem Lemma gibt es eine Zerlegung $x = yuvwz$ mit $|uvw| \leq n, uw \neq \varepsilon$ und $yu^m vw^m z \in L_2$ für alle $m \in \mathbb{N}$. Da $|uvw| \leq n$ kann uvw nicht sowohl a's als auch c's enthalten. Deshalb gibt es nur die folgenden Möglichkeiten:

- u und w enthalten nur a. Dann enthält $yu^0 vw^0 z$ maximal doppelt so viele a's wie b's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten a und b. Dann enthält $yu^0 vw^0 z$ nicht mehr b's als c's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten nur b. Dann enthält $yu^0 vw^0 z$ nicht mehr b's als c's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten b und c. Dann enthält $yu^2 vw^2 z$ mindestens halb so viele b's wie a's und ist deshalb nicht in L_2 enthalten. Widerspruch!

- u und w enthalten nur c. Dann enthalt yu^2vw^2z nicht mehr b's als c's und ist deshalb nicht in L_2 enthalten. Widerspruch!

L_3: L_3 ist regulär. Für jedes $n \in \mathbb{N}$ ist

$$L_3^n = \{x \in \Sigma^* : |x|_a > 2 \cdot n \text{ und } |x|_b = n\} = \{x \in \Sigma^* : |x|_a > 2 \cdot n\} \cap \{x \in \Sigma^* : |x|_b = n\}$$

ein Durchschnitt von zwei regulären Sprachen und deshalb regulär. Es folgt, dass $L_3 = \bigcup_{n \leq 1337} L_3^n$ auch regulär ist.

L_4: L_4 ist kontextfrei, aber nicht regulär. $L_4 = L_1 \cap \{x \in \Sigma^* : |x|_b \geq 1337\}$ ist Durchschnitt einer kontextfreien Sprache mit einer regulären Sprache und deshalb kontextfrei. Wäre L_4 auch noch regulär, dann wäre $L_1 = L_3 \cup L_4$ das auch. Wir haben aber oben gezeigt, dass L_1 nicht regulär ist.

Aufgabe H2 (Postsches Korrespondenzproblem)
Im Postschen Korrespondenzproblem ist 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, heißt diese eine Lösung des Korrespondenzproblems (x_1, y_1), (x_2, y_2), …, (x_k, y_k). Man kann zeigen, dass man nicht rekursiv entscheiden kann, ob eine Lösung existiert. Zeigen Sie, dass das folgende Korrespondenzproblem keine Lösungen hat.

i	x_i	y_i
1	10	101
2	011	11
3	101	011

Lösungsskizze: Nehmen wir an, es gäbe ein Lösung $i_1, i_2, \ldots, i_n$, und sei $\mathbf{x} = x_{i_1} x_{i_2} \ldots x_{i_n}$ und $\mathbf{y} = y_{i_1} y_{i_2} \ldots y_{i_n}$. Da die Wörter $\mathbf{x}$ und $\mathbf{y}$ mit demselben Buchstabe anfangen müssen, ist $i_1 = 1$. Wir haben also

$$\begin{aligned} \mathbf{x} &= 10 \cdots \\ \mathbf{y} &= 101 \cdots \end{aligned}$$

Es folgt, dass x_{i_2} mit einer 1 anfangen muss, also ist i_2 entweder 1 oder 3. $i_2 = 1$ ist unmöglich, da andererseits der vierte Buchstabe in $\mathbf{x}$ die 0 ist und in $\mathbf{y}$ die 1. Also ist $i_2 = 3$.
Wir können jetzt (induktiv) zeigen, dass wir für alle $n > 2$ auch $i_n = 3$ haben müssen. Wir sind nämlich immer in der Lage, dass das Wort $\mathbf{y}$ einen Buchstabe länger ist als $\mathbf{x}$, und dass der fehlende Buchstabe in $\mathbf{x}$ die 1 ist. Also muss man die Folge $i_1, i_2, \ldots, i_n$ immer erweitern, und wegen der Gründe die wir bereits gesehen haben, muss das immer mit $i_{n+1} = 3$ geschehen. Da unendliche Lösungen nicht erlaubt sind, gibt es für dieses Korrespondenzproblem also keine Lösung.