

Übungsblatt 8

Vorlesung Theoretische Grundlagen der Informatik im WS 16/17

Ausgabe 31. Januar 2017

Abgabe 09. Februar 2017, 11:00 Uhr (im Kasten im UG von Gebäude 50.34)

Bitte nutzen Sie den WebInScribe Deckblattgenerator
und heften Sie das Deckblatt an Ihr Übungsblatt.
<https://webinscribe.ira.uka.de/deckblatt/index.php?course=10588>.

Aufgabe 1

(2 + 2 + 1 Punkte)

Es ist folgende Verteilung für die Zeichen $Z = \{a, b, c, d, e, f\}$ gegeben.

Zeichen	Häufigkeit
a	3
b	97
c	33
d	33
e	10
f	23

- Kodieren Sie die Tabelle mittels Shannon-Fano. Geben Sie dazu in jedem Schritt den Kodierungsbaum an. Notieren Sie an den Blättern die repräsentierte Mengen. Geben Sie die Kodierung der einzelnen Zeichen an.
- Kodieren Sie Tabelle mittels dem Verfahren von Huffman. Stellen die Kodierungsschritt zu jedem Iterationsschritt dar. Geben Sie die Kodierung der einzelnen Zeichen an.
- Bestimmen Sie jeweils die mittlere Wortlänge.

Lösung:

-

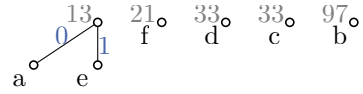
I)		Zeichen b c d f e a	Häufigkeit 97 33 33 21 10 3	Kodierung
II)		Zeichen b c d f e a	Häufigkeit 97 33 33 23 10 3	Kodierung 0 1 1 1 1 1
III)		Zeichen b c d f e a	Häufigkeit 97 33 33 23 10 3	Kodierung 0 1 0 1 0 1 1 1 1 1 1
IV)		Zeichen b c g f e a	Häufigkeit 97 33 33 23 10 3	Kodierung 0 1 0 1 1 0 1 1 1 1 1 1 1
V)		Zeichen b c d f e a	Häufigkeit 97 33 33 23 10 3	Kodierung 0 1 0 1 1 0 1 1 1 0 1 1 1 1 1 1
VI)		Zeichen b c d f e a	Häufigkeit 97 33 33 23 10 3	Kodierung 0 1 0 1 1 0 1 1 1 0 1 1 1 0 1 1 1 1

(b)

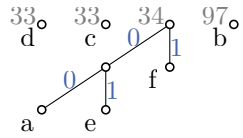
I)

3_a 10_e 21_f 33_d 33_c 97_b

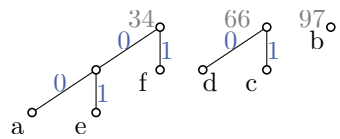
II)



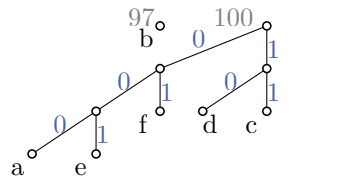
III)



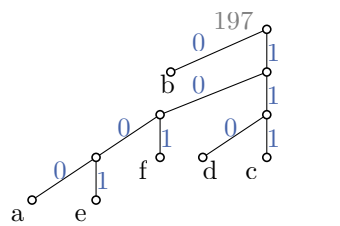
III)



IV)



V)



Zeichen	Häufigkeit	Shannon-Fano	Huffmann
b	97	0	0
c	33	1 0 1	1 1 1
d	33	1 0 1	1 1 0
f	23	1 1 0	1 0 1
e	10	1 1 1 0	1 0 0 1
a	3	1 1 1 1	1 0 0 0

(c)

(d)

$$\frac{97}{199} \cdot 1 + \frac{33}{199} \cdot 3 + \frac{33}{199} \cdot 3 + \frac{23}{199} \cdot 3 + \frac{10}{199} \cdot 4 + \frac{3}{199} \cdot 4 = 2.09$$

Aufgabe 2

(Punkte)

Sie und Ihre Kommilitonen Sina und Sascha haben soeben jeweils einen nichtdeterministischen PDA im Markt für Medien erworben. Sina kauft nun den Stack von Saschas PDA und baut ihn in ihrem eigenen PDA als Zweitstack ein, den sie unabhängig vom Erststack benutzen kann.

- (a) Sascha fragt sich, ob sein PDA jetzt weniger Sprachen erkennen kann als Ihrer. Beantworten Sie diese Frage, indem Sie begründen, welche Klasse von Sprachen von Saschas PDA erkannt

werden können.

Sina hofft, dass sie mit ihrem PDA mehr Sprachen erkennen kann als Ihrer.

- (b) Zeigen Sie, dass Sina recht hat, indem Sie eine Sprache angeben, die von Sinas PDA erkannt wird, von Ihrem aber nicht erkannt werden kann. Begründen Sie beides!
- (c) Welche Klasse von Sprachen kann Sinas PDA erkennen?
Hinweis: Kann Sina mit ihrem PDA ein mächtigeres Berechnungsmodell simulieren?

Der ebenso geniale wie auch gewiefte Doktor Meta ist bestürzt. Der pfiffige Informatik-Student Marvin Faulson ist in sein Hauptquartier eingebrochen (Vielleicht hätte man das Schloss doch ersetzen sollen? ... Nein, die Hochglanzpolitur für die Haifischbecken waren wichtiger) und hat den gefangenen internationalen Spitzenagent Sven van Hagen befreit. Ein episches Gefecht zwischen Marvin Faulson mit Sven van Hagen gegen Elsa und Dr. Meta ist unausweichlich. Da beide Parteien Gewalt verabscheuen einigen sie sich auf einen Wettstreit: Wer den besseren PDA bauen kann gewinnt! Während unzähliger Stunden voll rauchender Köpfe und Schweiß, gelingt es Elsa, den alleinstehenden Marvin Faulson mit einem Augenaufschlag abzulenken und ihm seinen soeben fertiggestellten PDA-Stack zu entwenden.

- (a) Um der Arbeit einen neuen Stack zu bauen zu entgehen, überlegt Marvin, ob sein PDA nicht auch ohne den Stack die gleiche Menge von Sprachen erkennen kann. Beantworten Sie diese Frage indem sie die Sprachen, die sein PDA nun noch erkennen kann klassifizieren.

In der Zwischenzeit sondiert Doktor Meta die neue Lage. Er hofft durch den zweiten Stack (der unabhängig vom schon vorhandenen Stack benutzbar ist) nun einen Vorteil zu haben, der sich auch durch Anbau eines neuen Stacks am PDA der Gegnerpartei nicht kompensieren lässt.

- (b) Zeigen Sie, dass seine Hoffnung begründet ist, indem Sie eine Sprache angeben, die von Metas PDA erkannt wird, von einem gewöhnlichen PDA aber nicht erkannt werden kann. Begründen sie beides!
- (c) Welche Klasse von Sprachen kann der Elsa/Meta-Automat nun erkennen? Kann er ein mächtigeres Berechnungsmodell simulieren?
- (d) Angenommen Elsa könnte einen weiteren Stack entwenden (diesmal natürlich vom ebenfalls alleinstehenden Sven van Hagen). Könnte Doktor Meta damit ein noch mächtigeres Berechnungsmodell bauen?

Lösung:

- (a) Aus der Vorlesung ist bekannt, dass Ihr PDA genau die kontextfreien Sprachen erkennen kann. Ohne einen Stack ist Marvin Faulsons PDA ein nichtdeterministischer endlicher Automat, der bekannterweise genau die regulären Sprachen erkennen kann. Da die Menge der regulären Sprachen eine echte Teilmenge der kontextfreien Sprachen ist, kann Faulsons PDA tatsächlich weniger Sprachen erkennen als ein PDA mit einem Stack.
- (b) In Aufgabe 1 von Übungsblatt 7 wurde gezeigt, dass die Sprache

$$L = \{w \in \{a, b, c\}^* : |w|_a = |w|_b = |w|_c\}$$

nicht kontextfrei ist. Da ein ganz normaler PDA genau die kontextfreien Sprachen erkennen kann, kann er insbesondere L nicht erkennen. Elsas PDA mit zwei Stacks kann L allerdings erkennen. Sie kann z.B. so vorgehen, dass sie für jedes a ein Zeichen auf den ersten Stack legt, für jedes b ein Zeichen auf den zweiten Stack legt, und für jedes c jeweils ein Zeichen von den beiden Stacks löscht.

- (c) Elsas PDA mit zwei Stacks L, R ist genauso mächtig wie eine Turingmaschinen und erkennt also die Menge der entscheidbaren Sprachen. Einerseits kann eine Turingmaschine offensichtlich einen PDA simulieren. Andererseits kann auch ein PDA \mathcal{P} mit zwei Stacks eine Turingmaschine \mathcal{M} wie folgt simulieren. Zu erst schreibt \mathcal{P} die Eingabe w auf den Stack R . Anschließend werden alle Symbole von R nach L verschoben, somit steht das erste Zeichen von w an erster Stelle von L . Diese Position stellt nun den Kopf der Turingmaschine dar und Der Stack L kodiert somit das Wort links vom Kopf und das Zeichen auf dem Lesekopf und R das Wort rechts vom Lesekopf. Eine Schreib Operation ist also eine Folge von einer **pop** und einer **push** Operation auf L . Ein Schritt nach Links ist eine **pop** Operation auf L und eine **push** Operation auf R , entsprechend Umgekehrt wird eine Bewegung nach rechts simuliert.
- (d) Nein, da ein PDA mit 3 Stacks lässt sich ebenfalls von einer Turing Maschine simulieren.

Aufgabe 3

(Punkte)

Sei

(a)

$$L_1 = \{p^R\#w \mid p, w \in \{0, 1\}^*, w \text{ enthält } p \text{ als Teilwort}\},$$

(b)

$$L_2 = \{x_1\#x_2\#\dots\#x_k \mid k \geq 1, x_i \in \{0, 1\}^*, \text{ für ein } i, j \text{ gilt } x_i = x_j^R\}.$$

wobei p^R das Spiegelwort von p ist. Können L_1 und L_2 jeweils von einem nichtdeterministischen PDA erkannt werden? Beweisen Sie!

Lösung:

Ja, L_i wird von folgenden kontextfreien Grammatik erzeugt:

(a)

$$\begin{aligned} S &\rightarrow TX \\ T &\rightarrow 0T0 \mid 1T1 \mid \#X \\ X &\rightarrow 0X \mid 1X \mid \varepsilon \end{aligned}$$

(b)

$$\begin{aligned} S &\rightarrow T \mid Y\#T \mid T\#Y \mid Y\#T\#Y \\ T &\rightarrow 0T0 \mid 1T1 \mid \# \mid \#Y\# \mid 1 \mid 0 \\ Y &\rightarrow X \mid Y\#YX \rightarrow 0X \mid 1X \mid 1 \mid 0 \end{aligned}$$

Aufgabe 4

(Punkte)

Sei $L \subset (\{0, 1, \dots, 9\} \cup \{+, -, *\})^*$ die Menge der Formeln in *umgekehrter Polnischen Notation*. Die Sprache L wird durch folgende Grammatik G beschrieben:

$$S \rightarrow SSo \mid zSo \mid \varepsilon \text{ mit } z \in \{0, 1, \dots, 9\}, o \in \{+, -, *\}.$$

- Bringe Sie die Grammatik G' in Greibach-Normalform. Dokumentieren Sie Ihr Vorgehen.
Lösung:

0. Chomsky-Normalform:

$$S \rightarrow SB \mid ZB \quad (1)$$

$$B \rightarrow SO \quad (2)$$

$$O \rightarrow + \mid - \mid * \quad (3)$$

$$Z \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (4)$$

(5)

1. Nummerieren der Variablen

$$A_1 \rightarrow A_1A_2 \mid A_4A_2 \quad (6)$$

$$A_2 \rightarrow A_1A_3 \quad (7)$$

$$A_3 \rightarrow + \mid - \mid * \quad (8)$$

$$A_4 \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (9)$$

(10)

2. Wende Ersetzung (ii) auf $A_1 \rightarrow A_1A_2$ an.

$$A_1 \rightarrow A_4A_2B \mid A_4A_2 \quad (11)$$

$$A_2 \rightarrow A_1A_3 \quad (12)$$

$$A_3 \rightarrow + \mid - \mid * \quad (13)$$

$$A_4 \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (14)$$

$$B \rightarrow A_1A_2 \mid A_1A_2B \quad (15)$$

3. Wende Ersetzung (i) auf $A_2 \rightarrow A_1A_3$ an.

$$A_1 \rightarrow A_4A_2B \mid A_4A_2 \quad (16)$$

$$A_2 \rightarrow A_4A_2BA_3 \mid A_4A_2A_3 \quad (17)$$

$$A_3 \rightarrow + \mid - \mid * \quad (18)$$

$$A_4 \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (19)$$

$$B \rightarrow A_1A_2 \mid A_1A_2B \quad (20)$$

4. Wende Ersetzung (i) auf A_1 an.

$$A_1 \rightarrow 0A_2B \mid 1A_2B \mid \dots \mid 9A_2B \quad (21)$$

$$\mid 0A_2 \mid 1A_2 \mid \dots \mid 0A_2 \quad (22)$$

$$A_2 \rightarrow A_4A_2BA_3 \mid A_4A_2A_3 \quad (23)$$

$$A_3 \rightarrow + \mid - \mid * \quad (24)$$

$$A_4 \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (25)$$

$$B \rightarrow A_1A_2 \mid A_1A_2B \quad (26)$$

5. Wende Ersetzung (i) auf A_2 an.

$$A_1 \rightarrow 0A_2B \mid 1A_2B \mid \dots \mid 9A_2B \quad (27)$$

$$\mid 0A_2 \mid 1A_2 \mid \dots \mid 9A_2 \quad (28)$$

$$A_2 \rightarrow 0A_2BA_3 \mid 1A_2BA_3 \mid \dots \mid 9A_2BA_3 \quad (29)$$

$$\mid 0A_2A_3 \mid 1A_2A_3 \mid \dots \mid 9A_2A_3 \quad (30)$$

$$A_3 \rightarrow + \mid - \mid * \quad (31)$$

$$A_4 \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (32)$$

$$B \rightarrow A_1A_2 \mid A_1A_2B \quad (33)$$

6. Wende Ersetzung (i) auf B an

$$A_1 \rightarrow 0A_2B \mid 1A_2B \mid \dots \mid 9A_2B \quad (34)$$

$$\mid 0A_2 \mid 1A_2 \mid \dots \mid 9A_2 \quad (35)$$

$$A_2 \rightarrow 0A_2BA_3 \mid 1A_2BA_3 \mid \dots \mid 9A_2BA_3 \quad (36)$$

$$\mid 0A_2A_3 \mid 1A_2A_3 \mid \dots \mid 9A_2A_3 \quad (37)$$

$$A_3 \rightarrow + \mid - \mid * \quad (38)$$

$$A_4 \rightarrow 0 \mid 1 \mid 2 \mid \dots \mid 9 \quad (39)$$

$$B \rightarrow 0A_2BA_2 \mid 1A_2BA_2 \mid \dots \mid 9A_2BA_2 \quad (40)$$

$$\mid 0A_2A_2 \mid 1A_2A_2 \mid \dots \mid 9A_2A_2 \quad (41)$$

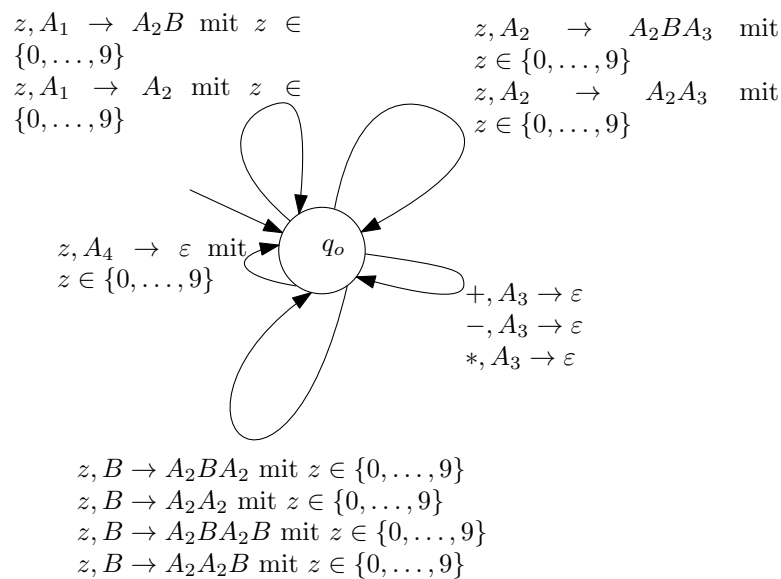
$$\mid 0A_2BA_2B \mid 1A_2BA_2B \mid \dots \mid 9A_2BA_2B \quad (42)$$

$$\mid 0A_2A_2B \mid 1A_2A_2B \mid \dots \mid 9A_2A_2B \quad (43)$$

$$(44)$$

- Konstruieren Sie mittels G' einen PDA der L erkennt.

Lösung:

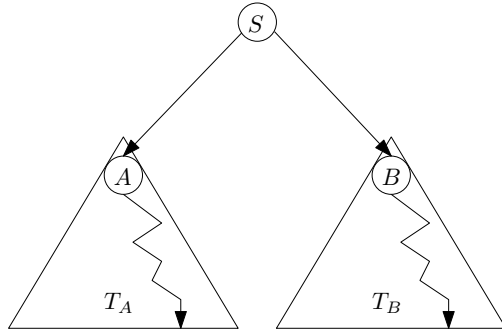


Sei G eine Grammatik in Chomsky-Normalform und T ein Syntaxbaum zu w , dessen längster Pfad die Länge n hat. Zeige, dass $|w| \leq 2^{n-1}$.

Lösung:

Induktion über n (Länge des längsten Pfades):

- $n = 1$: Ableitung der Form $A \rightarrow a$ mit $a \in \Sigma$
 $\Rightarrow w = a, |w| = 1 \leq 2^{1-1}$
- $n > 1$: An der Wurzel gibt es eine Ableitung $S \rightarrow AB$



- Längster Pfad in T_A, T_B : $n - 1$
- $\Rightarrow T_A, T_B$ haben je max. 2^{n-2} Blätter (Induktionsannahme)
- $\Rightarrow T$ hat max. $2^{n-2} + 2^{n-2} = 2^{n-1}$ Blätter
 $\Rightarrow |w| \leq 2^{n-1}$

Aufgabe 6

(Punkte)

Sei L eine kontextfreie Sprache. Die Funktionen α und β sind wie folgt definiert.

$$\alpha(L) = \{yx \mid xy \in L\} \tag{45}$$

$$\beta(L) = \{yxz \mid xyz \in L\} \tag{46}$$

Zeigen Sie:

- Die Menge der kontextfreien Sprachen ist abgeschlossen unter α .

Lösung:

- Die Menge der kontextfreien Sprachen ist nicht abgeschlossen unter β .

Lösung: