2. Klausur zur Vorlesung
Theoretische Grundlagen der Informatik
Wintersemester 2018/2019

Hier Aufkleber mit Name und Matrikelnummer anbringen

Vorname: _____

Nachname: _____

Matrikelnummer: _____

Beachten Sie:

- Bringen Sie den Aufkleber mit Ihrem Namen und Matrikelnummer auf diesem Deckblatt an und beschriften Sie jedes Aufgabenblatt mit Ihrem Namen und Matrikelnummer.
- Schreiben Sie die Lösungen auf die Aufgabenblätter und Rückseiten. Zusätzliches Papier erhalten Sie bei Bedarf von der Aufsicht.
- Es sind keine Hilfsmittel zugelassen.

	Mögliche Punkte					Erreichte Punkte				
	a	b	c	d	Σ	a	b	c	d	Σ
Aufg. 1	3	3	2	–	8				–	
Aufg. 2	2	3	3	–	8				–	
Aufg. 3	2	3	3	–	8				–	
Aufg. 4	3	4	2	–	9				–	
Aufg. 5	10				10					
Aufg. 6	2	3	2	2	9					
Aufg. 7	4	3	1	–	8				–	
Σ					60					

Problem 1: Endliche Automaten $3 + 3 + 2 = 8$ Punkte

In der Vorlesung haben Sie fehlertolerante Kodierungen kennengelernt. Diese können auch dann noch dekodiert werden, wenn einige Bits falsch übertragen wurden. In diesem Kontext wurde die Hammingdistanz $d(x, y)$ definiert. Für zwei Wörter x, y über demselben endlichen Alphabet Σ und mit derselben Länge ist $d(x, y)$ die Anzahl der Zeichen, die sich in x und y unterscheiden.

Sei L eine reguläre Sprache über dem Alphabet Σ. Definiere dazu die Sprache $B_k(L)$ wie folgt:

$$B_k(L) = \{x \in \Sigma^\star \mid \exists y \in L : |x| = |y| \text{ und } d(x, y) \leq k\}$$

Die Sprache $B_k(L)$ enthält also die Wörter, die aus Wörtern in L durch Ersetzen von höchstens k Zeichen entstehen.

(a) Sei $\Sigma = \{\mathtt{0}, \mathtt{1}\}$ und $L_{\mathtt{01}} = \{\mathtt{01}\}^\star$ und $k = 1$. Geben Sie den Zustandsübergangsgraphen eines deterministischen endlichen Automaten an, der $B_1(L_{\mathtt{01}})$ entscheidet. Sie dürfen einen impliziten Fehlerzustand benutzen. Ohne Fehlerzustand soll Ihr Automat 4 Zustände haben.

(b) Sei weiterhin $k = 1$. Zeigen Sie, dass die regulären Sprachen unter Anwendung von B_1 abgeschlossen sind. Beschreiben Sie dazu, wie aus einem endlichen Automaten für eine reguläre Sprache L ein endlicher Automat für $B_k(L)$ konstruiert werden kann. Dabei muss Ihr Automat für $B_k(L)$ nicht notwendigerweise deterministisch sein.

(c) Nun sei k beliebig, aber fest. Wie muss Ihre Konstruktion aus Aufgabenteil (b) angepasst werden, um zu zeigen, dass die regulären Sprachen unter Anwendung von B_k abgeschlossen sind?

Problem 2: Reguläre Sprachen $2 + 3 + 3 = 8$ Punkte

Gegeben seien das Alphabet $\Sigma = \{\mathtt{a}, \mathtt{b}\}$ und die Sprache $L = \{w \in \Sigma^\star \mid |w|_\mathtt{a} = |w|_\mathtt{b}\}$, wobei $|w|_x$ die Anzahl der Vorkommen des Symbols x in w bezeichnet.

(a) Beweisen Sie mithilfe des Pumping-Lemmas für reguläre Sprachen, dass L nicht regulär ist.

Wir betrachten nun folgende abgeschwächte Variante des Pumping-Lemmas für reguläre Sprachen:

> Sei L eine reguläre Sprache. Dann existiert eine Zahl $n \in \mathbb{N}$, so dass für jedes Wort $w \in L$ mit $|w| > n$ eine Darstellung
>
> $$w = uvx \text{ mit } |v| \leq n, v \neq \varepsilon$$
>
> existiert, bei der auch $uv^ix \in L$ ist für alle $i \in \mathbb{N}_0$.

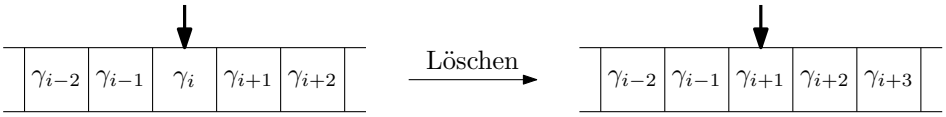
Im Gegensatz zum Pumping-Lemma, das Sie aus der Vorlesung kennen, wird hier also statt $|uv| \leq n$ nur $|v| \leq n$ gefordert.

(b) Zeigen Sie, dass L die Aussage des abgeschwächten Pumping-Lemmas erfüllt.

(c) Beweisen Sie mithilfe der Nerode-Relation, dass L nicht regulär ist.

Problem 3: Turingmaschinen $2 + 3 + 3 = 8$ Punkte

Eine *löschende Turingmaschine* (LTM) ist eine Turingmaschine, der zusätzlich zu den Kopfbewegungen *links*, *keine Bewegung* und *rechts* die Aktion *löschen* zur Verfügung steht. Beim Löschen wird das Feld, auf dem der Kopf der Turingmaschine steht, aus dem Band gelöscht und der Kopf bewegt sich auf das Feld, das zuvor der rechte Nachbar des nun gelöschten Feldes war:



Sei $L = \{w\#w^R \mid w \in \{0,1\}^*\}$.

(a) Welche Laufzeit braucht eine Turingmaschine mit einem Band, um L zu entscheiden? Wählen Sie aus den folgenden Möglichkeiten die schärfste untere Schranke für die Worst-Case-Laufzeit aus.

- ☐ $\Omega(\log n)$
- ☐ $\Omega(n)$
- ☐ $\Omega(n^2)$
- ☐ $\Omega(2^n)$

Begründen Sie Ihre Entscheidung kurz (ohne formalen Beweis).

(b) Beschreiben Sie, wie man L mit einer LTM in Linearzeit entscheiden kann. Belegen Sie die Laufzeit.

(c) Sind Turingmaschinen und löschende Turingmaschinen gleich mächtig? Beweisen Sie.

Problem 4: Aufzählbarkeit $3 + 4 + 2 = 9$ Punkte

Ein Dieb ist auf der Flucht, und Sie sollen ihn fangen! Der Dieb befindet sich auf dem Zahlenstrahl der ganzen Zahlen $\mathbb{Z}$ an einem Ihnen unbekannten Punkt $x_0 \in \mathbb{Z}$. Zunächst gehen wir davon aus, dass der Dieb sich nicht bewegt. Ihre Aufgabe ist es, den Dieb zu fangen. Dazu dürfen Sie an jedem Tag an genau einem Punkt nachsehen, ob der Dieb sich dort befindet. Falls ja, haben Sie ihn gefangen. Falls nicht, müssen Sie am nächsten Tag weitersuchen.

Mit einer *Strategie* zum Fangen des Diebes bezeichnen wir eine unendliche Folge von Punkten $s_1, s_2, \ldots$, so dass der Dieb nach endlich vielen Tagen gefangen wird.

(a) Geben Sie eine Strategie zum Fangen des Diebes an. Begründen Sie, dass der Dieb tatsächlich nach endlich vielen Tagen gefangen wird.

Nun darf sich der Dieb nachts bewegen. Dafür gehen wir davon aus, dass er an der Position $x_0 = 0$ startet. In jeder Nacht bewegt der Dieb sich um einen festen, aber Ihnen wiederum unbekannten Wert $v \in \mathbb{Z}$ fort. Tagsüber schläft der Dieb. Am ersten Tag schläft der Dieb also am Ort v, am zweiten Tag am Punkt $2v$ und so weiter. Beachten Sie, dass v positiv, negativ oder 0 sein kann.

(b) Geben Sie eine Strategie zum Fangen des Diebes an. Begründen Sie, dass der Dieb tatsächlich nach endlich vielen Tagen gefangen wird.

Achtung: Wenn Sie die erste Position überprüfen, hat sich der Dieb bereits einmal bewegt. Er befindet sich also nicht mehr unbedingt an der Position 0*, sondern an der Position* v*.*

Nun sind sowohl x_0 als auch v fest, aber Ihnen unbekannt. Am ersten Tag schläft der Dieb also am Ort $x_0 + v$, am zweiten Tag am Punkt $x_0 + 2v$ und so weiter.

(c) Geben Sie eine Strategie zum Fangen des Diebes an. Begründen Sie, dass der Dieb tatsächlich nach endlich vielen Tagen gefangen wird.

Problem 5: NP-Vollständigkeit 10 Punkte

Sie sind Aufseher eines internationalen Gefängnisses mit $3n$ Gefangenen und n Zellen für jeweils drei Insassen. Jeder Gefangene beherrscht eine bestimmte Menge an Sprachen, wobei es insgesamt m Sprachen gibt. Die Sprachen, die die Gefangenen sprechen, sind durch die $3n \times m$-Matrix M gegeben:

$$M[i][j] := \begin{cases} 1 & \text{falls Gefangener } i \text{ Sprache } j \text{ spricht} \\ 0 & \text{sonst.} \end{cases}$$

Sie wollen verhindern, dass die Gefangenen untereinander Informationen austauschen können. Dazu wollen Sie die Gefangenen so auf die Zellen aufteilen, dass innerhalb jeder Zelle die Gefangenen nicht miteinander kommunizieren können, also kein Insasse eine gemeinsame Sprache mit einem anderen Insassen hat.

Wir formulieren dieses Szenario als Entscheidungsproblem MUTE PRISON. Gegeben sind dabei die Menge P der Gefangenen, die Menge S der Sprachen und die Matrix M. Gesucht ist eine Einteilung der Gefangenen in die Zellen, sodass kein Insasse mit einem anderen kommunizieren kann.

Das NP-schwere Entscheidungsproblem PARTITION INTO TRIANGLES ist wie folgt definiert:

Gegeben: Ungerichteter einfacher Graph $G = (V, E)$ mit $|V| = 3n$.
Frage: Gibt es eine Partitionierung $V_1 \cup V_2 \cup \cdots \cup V_k$ von V, sodass jedes V_i ein Dreieck in G induziert?

Zeigen Sie, dass MUTE PRISON NP-vollständig ist. Geben Sie bei Ihrer Reduktion explizit an, von welchem Problem auf welches reduziert wird!

Hinweis: Beachten Sie, dass ein Dreieck eine Clique der Größe 3 *ist, d.h. alle drei Knoten sind paarweise miteinander verbunden. Was bedeutet das für den Komplementgraphen von G?*

Problem 6: Approximationsalgorithmen $2 + 3 + 2 + 2 = 9$ Punkte

Eine *Knotenüberdeckung* (engl. *Vertex Cover*) für einen ungerichteten Graph $G = (V, E)$ ist eine Knotenmenge $V' \subseteq V$, sodass für jede Kante in E mindestens einer der beiden Endpunkte in V' enthalten ist. Das NP-schwere Optimierungsproblem VERTEX COVER fragt nach einer minimalen Knotenüberdeckung und ist wie folgt definiert:

Gegeben: Ungerichteter Graph $G = (V, E)$
Gesucht: Knotenmenge $V' \subseteq V$ mit $|V'|$ minimal, sodass für alle Kanten $(u, v) \in E$ gilt:
$u \in V'$ oder $v \in V'$.

In der Vorlesung haben wir gesehen, dass sich jedes NP-schwere Problem als *ganzzahliges lineares Programm* (engl. *Integer Linear Program*, ILP) darstellen lässt. Für VERTEX COVER können wir also das äquivalente Problem VC-$\mathbb{N}$ betrachten:

Gegeben:
Ungerichteter Graph $G = (V, E)$
Variablen:
Für jeden Knoten $v \in V$ eine Variable x_v
Nebenbedingungen:
(I) Für jeden Knoten $v \in V$: $x_v \in \{0, 1\}$
(II) Für jede Kante $(u, v) \in E$: $x_u + x_v \geq 1$
Zielfunktion:
Minimiere $f_{\mathbb{N}}(G) = \sum_{v \in V} x_v$

Aus einer Lösung für VC-$\mathbb{N}$ lässt sich eine Lösung für VERTEX COVER konstruieren, indem man $V' = \{v \in V \mid x_v = 1\}$ setzt. Der Wert von x_v gibt also an, ob v in der Knotenüberdeckung enthalten ist oder nicht.

Eine gängige Methode, ILPs zu approximieren, ist die *LP-Relaxierung*. Dabei wird die Beschränkung aufgehoben, dass die Variablen ganzzahlige Werte haben müssen. Das dadurch entstehende *lineare Programm* (LP) lässt sich in Polynomialzeit lösen. Im Fall von VERTEX COVER erhalten wir das Problem VC-$\mathbb{R}$. Der einzige Unterschied gegenüber VC-$\mathbb{N}$ ist in Nebenbedingung (I): Die x_v dürfen nun beliebige reelle Zahlen aus dem Intervall $[0, 1]$ sein. Dementsprechend ist auch der Wert der Zielfunktion $f_{\mathbb{R}}(G) = \sum_{v \in V} x_v$ reellwertig.

Sei im Folgenden $OPT_{\mathbb{R}}(G)$ der Wert von $f_{\mathbb{R}}(G)$ für eine optimale Lösung von VC-$\mathbb{R}$ und $OPT_{\mathbb{N}}(G)$ der Wert von $f_{\mathbb{N}}(G)$ einer optimalen Lösung von VC-$\mathbb{N}$.

(a) Zeigen Sie, dass für jeden Graphen G gilt: $OPT_{\mathbb{R}}(G) \leq OPT_{\mathbb{N}}(G)$.

(b) Geben Sie einen Graphen G mit höchstens vier Knoten an, für den $OPT_{\mathbb{R}}(G) < OPT_{\mathbb{N}}(G)$ gilt. Geben Sie für beide Probleme eine optimale Lösung an.

Wir betrachten nun folgenden Algorithmus $\mathcal{A}$, der eine (nicht notwendigerweise optimale) Lösung für VC-$\mathbb{N}$ berechnet:

1. Berechne eine optimale Lösung $X = (x_1, \ldots, x_n)$ mit $x_i \in [0, 1]$ für VC-$\mathbb{R}$.
2. Generiere eine Lösung $X' = (x'_1, \ldots, x'_n)$ mit $x'_i \in \{0, 1\}$ für VC-$\mathbb{N}$, indem wir jedes x'_i wie folgt wählen:

$$x'_i = \begin{cases} 1 & \text{falls } x_i \geq \frac{1}{2} \\ 0 & \text{sonst.} \end{cases}$$

(c) Zeigen Sie, dass Algorithmus $\mathcal{A}$ eine Lösung für VC-$\mathbb{N}$ berechnet, also dass die Nebenbedingungen erfüllt sind.

(d) Zeigen Sie, dass $\mathcal{A}$ ein Approximationsalgorithmus für VC-$\mathbb{N}$ mit einer relativen Gütegarantie von 2 ist.

Hinweis: Zeigen Sie dafür zunächst, dass für beliebige Graphen G gilt: $\mathcal{A}(G) \leq 2 \cdot OPT_{\mathbb{R}}(G)$.

Problem 7: Chomsky-Normalform und CYK-Algorithmus $4 + 3 + 1 = 8$ Punkte

(a) Gegeben sei die Grammatik $G_1 = (\Sigma_1, V_1, S, R_1)$ mit Terminalen $\Sigma_1 = \{\mathtt{a}, \mathtt{b}\}$, Nichtterminalen $V_1 = \{S, A, B\}$, Startsymbol S und Produktionen

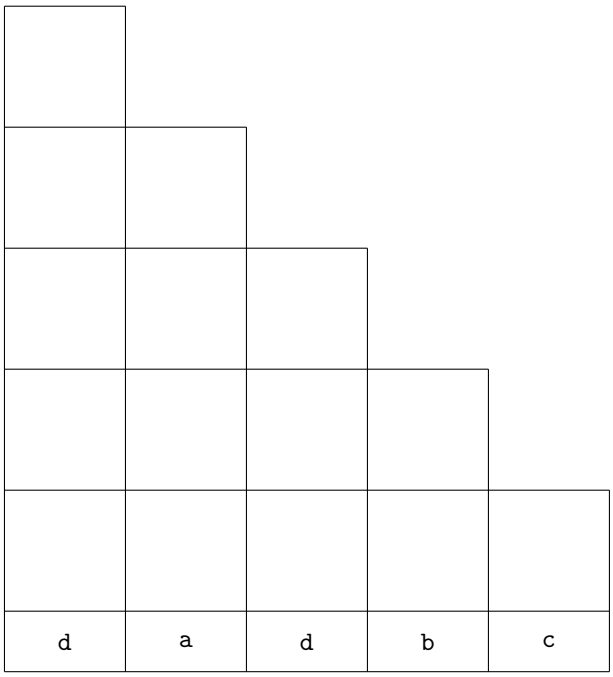
$$\begin{aligned} R_1 = \{ & S \to A\mathtt{b}B \\ & A \to A\mathtt{a} \mid \varepsilon \\ & B \to S \mid \mathtt{b}\} \end{aligned}$$

Transformieren Sie G_1 in Chomsky-Normalform. Geben Sie bei Ihrer Umformung die Zwischenschritte an.

(b) Gegeben sei die Grammatik $G_2 = (\Sigma_2, V_2, S, R_2)$ mit Terminalen $\Sigma_2 = \{\texttt{a}, \texttt{b}, \texttt{c}, \texttt{d}\}$, Nichtterminalen $V_2 = \{S, A, B, C, D\}$, Startsymbol S und Produktionen

$$\begin{aligned} R_2 = \{ & S \to AC \mid BD \\ & A \to BC \mid \texttt{a} \\ & B \to DA \mid \texttt{b} \mid \texttt{c} \\ & C \to CB \mid AD \mid \texttt{c} \mid \texttt{d} \\ & D \to BD \mid DC \mid \texttt{d}\} \end{aligned}$$

Überprüfen Sie mit dem CYK-Algorithmus, ob das Wort dadbc in der Sprache $L(G_2)$ enthalten ist.



(c) Geben Sie einen Syntaxbaum für das Wort dadbc in G_2 an.