1. Klausur zur Vorlesung
Theoretische Grundlagen der Informatik
Wintersemester 2016/2017

Hier Aufkleber mit Name und Matrikelnummer anbringen

Vorname: ______________________________

Nachname: ______________________________

Matrikelnummer: ______________________________

Beachten Sie:

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

Aufgabe	Mögliche Punkte						Erreichte Punkte					
	a	b	c	d	e	Σ	a	b	c	d	e	Σ
1	2	2	2	–	–	6				–	–	
2	1	6	2	~~3~~	–	9					–	
3	2	4	2	–	–	8				–	–	
4	2	4	2	3	–	11					–	
5	1	3	2	3	–	9					–	
6	4	–	–	–	–	4		–	–	–	–	
7	10×1					10						
Σ						57						

Problem 1: Automaten $2 + 2 + 2 = 6$

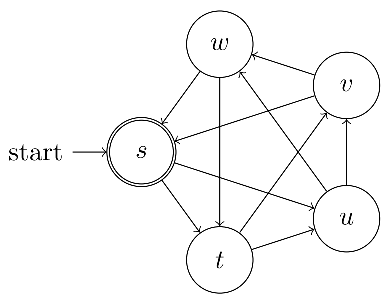
(a) Sei die Sprache

$$L = \{w \in \{\mathtt{a}, \mathtt{b}\}^\star \mid w \text{ enthält } \mathtt{ab} \text{ als Teilwort} \vee |w|_\mathtt{a} \text{ ist gerade}\}$$

gegeben. Hierbei bezeichnet $|w|_a$ die Anzahl der Vorkommen von a in w. Geben Sie den Übergangsgraphen eines nichtdeterministischen endlichen Automaten, der L erkennt, an. Verwenden Sie höchstens sieben Zustände.

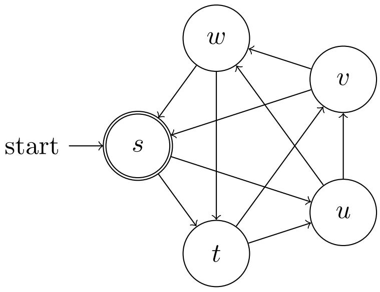
Hinweis: Einen Fehlerzustand dürfen Sie als implizit gegeben annehmen. Es wird angenommen, dass alle nicht explizit angegebenen Übergänge in diesen Fehlerzustand führen. Der Fehlerzustand zählt nicht zu den sieben Zuständen, die Sie verwenden dürfen.

(b) Geben Sie die Sprache $L(\mathcal{A})$ über dem Alphabet $\{\mathtt{a}\}$ an, die der folgende nichtdeterministische endliche Automat $\mathcal{A}$ erkennt. Alle Übergänge sind für das Zeichen a angegeben.



$L(\mathcal{A}) =$

(c) Geben Sie einen zu dem Automaten $\mathcal{A}$ aus Teilaufgabe (b) äquivalenten vollständigen deterministischen endlichen Automaten tabellarisch an. Verwenden Sie hierfür die Potenzmengenkonstruktion, und geben Sie deren Schritte explizit an.



Zur Erinnerung: der Automat $\mathcal{A}$ aus Teilaufgabe (b).

Zustand	Übergang

Problem 2: Berechenbarkeit $1 + 6 + 2 + 3 = 12$

Eine Instanz I des Postschen Korrespondenzproblem (PKP) ist eine endliche Menge $I \subset \Sigma^\star \times \Sigma^\star$, wobei Σ ein endliches Alphabet ist. Eine Folge von Tupeln $t_1, t_2, \ldots, t_k$ mit $k \in \{1, 2, \ldots\}$ und $t_i = (x_i, y_i) \in I$ ist eine Lösung von I genau dann, wenn $x_1 \cdot x_2 \cdot \ldots \cdot x_k = y_1 \cdot y_2 \cdot \ldots \cdot y_k$ gilt.

(a) Geben Sie seine Lösung der folgenden PKP-Instanz an:

$$\left\{ \binom{a}{rant}, \binom{am}{m}, \binom{at}{tata}, \binom{par}{pa}, \binom{ntat}{a} \right\}$$

Sei $\mathcal{G}$ die Menge aller kontextfreien Grammatiken über einem festen endlichen Alphabet Σ, und sei $\langle G \rangle$ eine geeignete Kodierung einer kontextfreien Grammatik $G \in \mathcal{G}$. Dann ist die Sprache L definiert als

$$L = \left\{ \langle G \rangle \mid G \in \mathcal{G},\ \exists w \in L(G) : w = w^R \right\},$$

wobei w^R das Spiegelwort eines Wortes $w \in \Sigma^\star$ bezeichnet.

(b) Zeigen Sie, dass L nicht entscheidbar ist.

Hinweis: Konstuieren Sie als Teil Ihres Beweises zu einer PKP-Instanz I eine kontextfreie Grammatik G, so dass $\langle G \rangle$ genau dann in L enthalten ist, wenn I eine Lösung besitzt.

(c) Zeigen Sie, dass L semi-entscheidbar ist.

(d) Beweisen Sie mit Hilfe der Aussagen von Teilaufgabe (b) und (c), dass die Menge der kontextfreien Sprachen unter Komplementbildung nicht abgeschlossen ist.

Problem 3: $\mathcal{NP}$-Vollständigkeit 2 + 4 + 2 = 8

Sei für $k \in \{1, 2, \ldots\}$ folgende Familie von Problemen gegeben.

Problem SAT$_{\geq k}$

Gegeben: Menge U von n Variablen, Menge C von m Klauseln über U, Parameter k.
Frage: Existieren mindestens k unterschiedliche wahrheitserfüllende Belegungen von C?

(a) Geben Sie für ein beliebiges aber festes k eine polynomiale Transformation SAT $\propto$ SAT$_k$ an.

(b) Zeigen Sie mit Ihrer polynomialen Transformation $\propto$ aus Teilaufgabe (a), dass das Problem SAT$_{\geq k}$ für ein beliebiges aber festes k $\mathcal{NP}$-vollständig ist.

(c) Für $k \in \{0, 1, 2, \ldots\}$ fragt die Familie von Problemen $\textsc{Sat}_{=k}$ nach *genau* k erfüllenden Belegungen. Geben Sie von den aus Vorlesung und Übung bekannten Komplexitätsklassen möglichst scharf diejenige an, in der das Problem $\textsc{Sat}_{=0}$ liegt. Definieren Sie diese Komplexitätsklasse außerdem formal.

Problem 4: Approximationsalgorithmen $2 + 4 + 2 + 3 = 11$

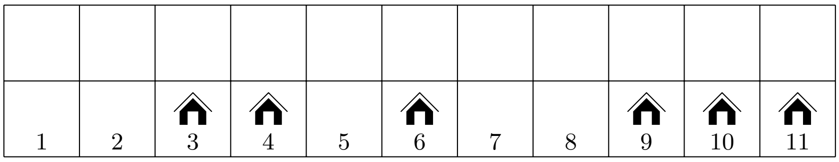
Der ebenso selbstbewusste wie fürsorgliche Superbösewicht Doktor Meta ist erzürnt. Die mediale Berichterstattung über ihn ist schrecklich unehrlich! Um die Bewohner des Metaversiums vor solchen Nachrichten zu beschützen, möchte Doktor Meta alle Straßen mit Störsendern ausstatten.

Eine gerade Straße besteht aus insgesamt $n \in \{1, 2, \dots\}$ Zellen. In manchen Zellen befindet sich ein Haus, in anderen nicht. Die Störsender können ausschließlich auf Häusern installiert werden. Ein Störsender hat eine Reichweite von $k \in \{0, 1, 2, \dots\}$. Ein Haus mit Position i liegt genau dann in Reichweite eines Störsenders mit Position j, wenn $|i - j| \leq k$ gilt. Jedes Haus soll innerhalb der Reichweite mindestens eines Störsenders liegen. Insgesamt sollen möglichst wenige Störsender installiert werden.

Dazu wird folgender Algorithmus $\mathcal{A}$ verwendet. Speichere die Position i des Hauses mit der kleinsten Position, das noch nicht abgedeckt ist. Finde dann das Haus mit der größten Position j im Intervall $\{i, i + 1, \dots, i + k\}$ und platziere dort den nächsten Störsender. Wiederhole diesen Prozess solange, bis alle Häuser innerhalb der Reichweite eines Störsenders liegen.

(a) Führen Sie diesen Algorithmus für $k = 2$ auf folgendem Beispiel einer geraden Straße aus, indem Sie genau die Positionen markieren, an denen Störsender installiert werden.

Hinweis: Markieren Sie eine Position, indem Sie das Kästchen über dem entsprechenden Haus ankreuzen.

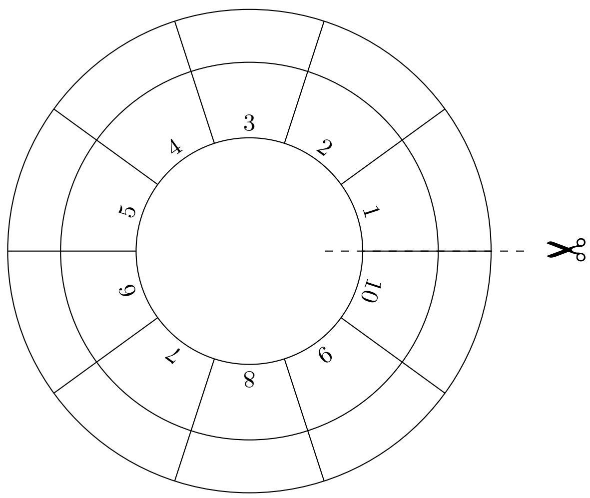


(b) Zeigen Sie, dass der Algorithmus für allgemeine k immer eine optimale Lösung liefert.

Hinweis: Vergleichen Sie die Lösung, die der Algorithmus $\mathcal{A}$ berechnet, mit einer optimalen Lösung.

Der Algorithmus soll nun erweitert werden, um nicht nur gerade Straßen mit Störsendern ausstatten zu können, sondern auch ringförmige Straßen. Dazu wird eine ringförmige Straße zwischen zwei beliebigen benachbarten Zellen „aufgeschnitten“ und dann so bearbeitet, wie eine gerade Straße.

(c) Zeigen Sie, dass dieser Ansatz dazu führen kann, dass zu viele Störsender installiert werden. Vervollständigen Sie dazu folgende Instanzen mit Häusern und markieren Sie die optimalen Lösungen der ringförmigen und der aufgeschnittenen geraden Straßen. Die Reichweite der Störsender ist $k = 2$.



1	2	3	4	5	6	7	8	9	10

(d) Bei dem Algorithmus $\mathcal{A}$ handelt es sich für ringförmige Straßen um einen Approximationsalgorithmus. Bestimmen Sie die Gütegarantie des Algorithmus. Geben Sie dazu möglichst scharf eine Funktion f an, so dass für alle Instanzen I gilt $\mathcal{A}(I) \leq f(\mathrm{OPT}(I))$. Beweisen Sie die Gütegarantie.

Diskussion: Auch bei diesem Beweis ist es wichtig, von einer optimalen Lösung der ringförmigen Straße auszugehen. Man kann nicht ohne Weiteres davon ausgehen, dass das „Zusammenkleben" einer optimalen Lösung für gerade Straßen zu einer 1-Approximation im ringförmigen Fall führt. Außerdem wurde häufig nur argumentiert, dass es „Probleme im Randbereich" gibt, die mit einem zusätzlichen Störsender gelöst werden können. Ohne das „wie" und „wieso" ist so etwas allerdings kein Beweis.

Problem 5: Maschinenmodelle $1 + 3 + 2 + 3 = 9$

Es sei der durch leeren Stack erkennende Kellerautomat $\mathcal{A} = (Q, \Sigma, \Gamma, s, Z_0, \delta, F = \emptyset)$ gegeben mit der Zustandsmenge $Q = \{s, q_0, q_1, s', q_0', q_1'\}$, dem Eingabealphabet $\Sigma = \{\mathtt{0}, \mathtt{1}, \mathtt{\#}\}$, dem Stackalphabet $\Gamma = \{Z_0, A_0, A_1, A_2, A_3\}$, Startzustand s und Stackinitialisierung Z_0. Die Übergangsfunktion δ ist folgendermaßen definiert:

$\forall Z \in \Gamma:$	$\delta(s, \mathtt{0}, Z) = \{(q_0, Z)\}$	$\forall Z \in \Gamma:$	$\delta(s', \mathtt{0}, Z) = \{(q_0', Z)\}$
$\forall Z \in \Gamma:$	$\delta(s, \mathtt{1}, Z) = \{(q_1, Z)\}$	$\forall Z \in \Gamma:$	$\delta(s', \mathtt{1}, Z) = \{(q_1', Z)\}$
$\forall Z \in \Gamma:$	$\delta(s, \mathtt{\#}, Z) = \{(s', Z)\}$		$\delta(s', \varepsilon, Z_0) = \{(s', \varepsilon)\}$
$\forall Z \in \Gamma:$	$\delta(q_0, \mathtt{0}, Z) = \{(s, ZA_0)\}$		$\delta(q_0', \mathtt{0}, A_0) = \{(s', \varepsilon)\}$
$\forall Z \in \Gamma:$	$\delta(q_0, \mathtt{1}, Z) = \{(s, ZA_1)\}$		$\delta(q_0', \mathtt{1}, A_1) = \{(s', \varepsilon)\}$
$\forall Z \in \Gamma:$	$\delta(q_1, \mathtt{0}, Z) = \{(s, ZA_2)\}$		$\delta(q_1', \mathtt{0}, A_2) = \{(s', \varepsilon)\}$
$\forall Z \in \Gamma:$	$\delta(q_1, \mathtt{1}, Z) = \{(s, ZA_3)\}$		$\delta(q_1', \mathtt{1}, A_3) = \{(s', \varepsilon)\}$

(a) Liegt das Wort `01#10` in der Sprache $L(\mathcal{A})$? Liegt das Wort `0010#1000` in der Sprache $L(\mathcal{A})$?

(b) Geben Sie die Sprache $L(\mathcal{A})$ an, die $\mathcal{A}$ erkennt.

$L(\mathcal{A}) =$

Sei nun $k \in \{0, 1, 2, \ldots\}$ eine beliebige aber feste Konstante. Das Berechnungsmodell Kellerautomat wird nun so eingeschränkt, dass auf dem Stack zu jedem Zeitpunkt nur noch höchstens k Symbole liegen können. Für jede Konfiguration (q, w, α) gilt also $\alpha \in \Gamma^k$.

(c) Argumentieren Sie, dass der Kellerautomat $\mathcal{A}$ dieser Einschränkung nicht genügt, indem Sie ein Wort angeben, bei dessen Abarbeitung durch $\mathcal{A}$ zu einem Zeitpunkt mindestens $k + 1$ Symbole auf dem Stack liegen.

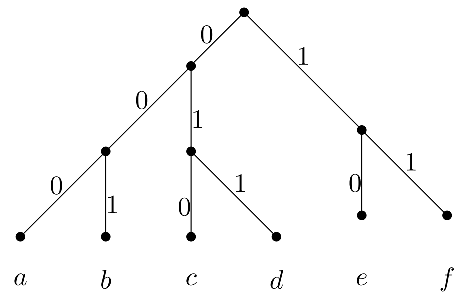
(d) Können derart eingeschränkte nichtdeterministische Kellerautomaten genau die regulären Sprachen erkennen? Beweisen Sie Ihre Behauptung!

Problem 6: Informationstheorie 4

Die Häufigkeitsverteilung H der Symbole $\Sigma = \{a, b, c, d, e, f\}$ ist durch folgende Tabelle gegeben.

Zeichen	a	b	c	d	e	f
Wahrscheinlichkeit	$\frac{2}{47}$	$\frac{8}{47}$	$\frac{11}{47}$	$\frac{4}{47}$	$\frac{9}{47}$	$\frac{13}{47}$

Gegeben ist folgender Codierungsbaum T zur Häufigkeitsverteilung H. Zeigen oder widerlegen Sie, dass T eine minimale mittlere Codewortlänge hat.



Problem 7: Gemischtes 10

Sind die folgenden Aussagen korrekt? Begründen Sie jeweils kurz.

(a) Sei ein deterministischer endlicher Automat $\mathcal{A} = (Q, \Sigma, \delta, s, F)$ gegeben. Dann akzeptiert der Automat $(Q, \Sigma, \delta, s, Q \setminus F)$ die Sprache $L(\mathcal{A})^c$.

(b) Sei $L \subseteq \Sigma^*$ eine Sprache über einem endlichen Alphabet Σ. Existiert für jedes $w \in L$ ein endlicher Automat, der w erkennt, so ist L regulär.

(c) Jede Sprache über einem endlichen Alphabet ist rekursiv aufzählbar.

(d) Ein 2-Approximationsalgorithmus liefert mit Wahrscheinlichkeit $\frac{1}{2}$ eine optimale Lösung.

(e) Eine Grammatik, die mindestens eine Ableitungsregel enthält, erzeugt mindestens ein Wort.

(f) Das Pumping-Lemma ist für die Sprache

$$L = \{a^i b^j \mid i, j > 0, i \text{ gerade}, j \text{ prim}\}$$

erfüllt.

(g) Es existiert eine Sprache in $\mathcal{NP}$, die in Polynomialzeit entschieden werden kann.

(h) Es existiert ein Polynom p, sodass die Anzahl n der Äquivalenzklassen eines minimalen nichtdeterministischen Automaten mit m Zuständen durch $p(m)$ beschränkt ist.

(i) Der Schnitt einer kontextfreien Sprache mit einer regulären Sprache ist regulär.

(j) Sei $K \in \{1, 2, \ldots\}$ eine beliebige aber feste Konstante. Das Problem Knapsack mit in K beschränkten Gewichten und Kosten liegt in $\mathcal{P}$.