Klausur zur Vorlesung

Theoretische Informatik I

Prof. Dr. Christoph Kreitz
Universität Potsdam, Theoretische Informatik, Wintersemester 2008/2009

Termin: 19.12.2008

Dies ist eine **Auswahlklausur**, *d.h. sie enthält mehr Aufgaben als sie in der vorgesehenen Zeit bearbeiten können. Zur Bearbeitung sind 120 Minuten vorgesehen. Zum Bestehen der Klausur wären 40 der 100 möglichen Punkte erforderlich. Für je 9 Punkte erhalten Sie einen von 9 möglichen Hausaufgabenpunkten. Es wird kaufmännisch gerundet.*

Schreiben Sie bitte deutlich und verwenden Sie keinen Bleistift oder Rotstift. Einsicht in die eigenen Unterlagen ist gestattet. Technische Hilfsmittel wie z.B. Taschenrechner, Mobiltelefone, PDAs oder Laptops sind nicht zugelassen.

Erläutern Sie alle Ihre Schritte und begründen Sie alle Ihre Aussagen. Es genügt nicht, nur Lösungen anzugeben, ohne dass Ihre Vorgehensweise klar erkennbar ist.

Sollten Ihre Lösungen nicht auf den Aufgabenblättern Platz finden, bearbeiten Sie bitte auf jedem Zusatzblatt nur eine Aufgabe. Heften Sie die Blätter bei der Abgabe zusammen und tragen Sie auf jedem Blatt Ihren Namen, Matrikelnummer und Übungsgruppe ein.

Name:

Matrikelnummer:

Übungsgruppe:

Punkte:

Aufgabe	1	2	3	4	5	6	7	gesamt
erreichbare Punktzahl	21	17	12	15	15	10	10	100
erreichte Punktzahl								

Anmerkungen zu den “Musterlösungen”.

Eine Musterlösung ist ein Hinweis auf eine mögliche Lösung. Es gibt oft viele andersartige Lösungen, die genauso gut und manchmal vollständig unvergleichbar mit der Musterlösung sind.

Der Text der Musterlösung ist zuweilen ausführlicher als nötig. Ich kann auch nicht garantieren, daß keine Tippfehler mehr vorkommen. Wir haben allerdings keine mehr gesehen.

Die Punkte, die neben den Lösungsfragmenten stehen, sollen zeigen, was die einzelnen Aussagen “wert” sind, wenn man denselben Weg verfolgt. Bei andersartigen Lösungen mussten die Korrekteure einschätzen, was wohl auf welchen Teil “angerechnet” werden kann. So steckt z.B. oft eine fehlende Erläuterung nachher im Beweis. Deswegen bleibt die Punktbewertung z.T. immer noch Ermessenssache.

Die Korrektur hat gezeigt, daß der fettgedruckte Hinweis “Erläutern Sie alle Ihre Schritte und begründen Sie alle Ihre Aussagen” leider selten beachtet wurde. Ich werde ihn demnächst direkt bei jeder Einzelaufgabe plazieren und die Punkte ausweisen, die dafür vorgesehen sind.

Aufgabe 1 (Einfache Fragen zu bisher behandelten Themen, 21 Punkte)

Beantworten Sie die folgenden Fragen knapp aber präzise in Ihren eigenen Worten.

1. Definieren Sie: L ist eine kontextfreie, aber keine reguläre Sprache genau dann, wenn ... **(2 Punkte)**

2. Geben Sie 4 Formalismen an, mit denen genau die regulären Sprachen beschrieben werden können. **(2 Punkte)**

3. Geben Sie 4 Operationen an, unter denen die regulären Sprachen abgeschlossen sind. **(2 Punkte)**

4. Geben Sie drei Techniken an, mit denen bewiesen werden kann, dass eine Sprache nicht regulär ist. **(3 Punkte)**

5. Es sei $h : \{a, b, c\}^* \to \{0, 1\}^*$ ein Homomorphismus mit $h(a)=0$, $h(b)=10$ und $h(c)=1$ und $L=\{0100\}$. Beschreiben Sie die Menge $h^{-1}(L) = \{w \in \{x, y\}^* \mid h(w) \in L\}$. **(1 Punkt)**

6. Geben Sie drei Eigenschaften an, die sich für DEAs automatisch prüfen lassen! **(3 Punkte)**

7. Beweisen Sie durch ein Gegenbeispiel, daß die Äquivalenz $(RS + S)^*RS \cong R(SR + R)^*S$ nicht für beliebige reguläre Ausdrücke R und S gilt. **(1 Punkt)**

8. Beschreiben Sie, was einen Mealy-Automaten von einem DEA unterscheidet! **(1 Punkt)**

9. Definieren Sie: Der ε-NEA $A = (Q, \Sigma, \delta, q_0, F)$ akzeptiert ein Wort $w \in \Sigma^*$ genau dann, wenn ... **(1 Punkt)**

10. Sei $A = (Q, \Sigma, \delta, q_0, F)$ ein ε-NEA. Wie bezeichnet man für $p \in Q$ die Menge $\{x \mid x \in Q \wedge \hat{\delta}(p, \varepsilon) = x\}$? **(1 Punkt)**

11. Wann ist eine kontextfreie Grammatik G eindeutig? **(1 Punkt)**

12. Wo steckt der Fehler in der folgenden Aussage: "Die Sprache einer mehrdeutigen kontextfreien Grammatik ist inhärent mehrdeutig"? **(1 Punkt)**

13. Geben Sie den speziellsten Typ der Grammatik $G = (\{S, A, B\}, \{x, y, z\}, P, S)$ mit der Regelmenge $P = \{S \rightarrow AxB,\ S \rightarrow \varepsilon,\ A \rightarrow yAB,\ A \rightarrow B,\ B \rightarrow zSz\}$ an! **(1 Punkt)**

14. Worin besteht bei der Umwandlung von NEAs in DEAs mittels optimierter Teilmengenkonstruktion die „Optimierung“? **(1 Punkt)**

Lösung

1. *sie von einer kontextfreien Grammatik erzeugt wird,* •
   aber von keinem DEA akzeptiert / von keinem regulären Ausdruck beschrieben / von keiner regulären Grammatik erzeugt wird. •

2. *DEAs, NEAs, ε-NEAs, reguläre Ausdrücke, rechtslineare und linkslineare Grammatiken.*

   (je 1/2 Punkt, max. 2 Punkte, je 1/2 Punkt Abzug für eine falsche Antwort)

3. *endliche Vereinigung, endlicher Durchschnitt, Komplementbildung, endliche Differenz, Spiegelung, Hüllenbildung, Verkettung, Homomorphismen, inverse Homomorphismen, Quotientenbildung, ...*

   (je 1/2 Punkt, max. 2 Punkte, je 1/2 Punkt Abzug für eine falsche Antwort)

4. *1. Widerlegungs-/Kontrpositionsbeweis mit dem Pumping Lemma* •

   2. Transformation der betrachteten Sprache in eine bekanntermaßen nichtreguläre Sprache, wobei ausschließlich regularitätserhaltende Transformationen verwendet werden. •

   Hinweis auf Verwendung von Abschlusseigenschaften genügt

   3. Anwendung des Satzes von Myhill/Nerode. •

5. $h^{-1}(L) = \{aba, acaa\}$ •

6. *ist die Sprache des Automaten leer* •
   ist ein Wort w Element der Sprache eines Automaten •
   beschreiben zwei Automaten dieselbe Sprache •

7. *Für $R = r$, $S = s$ gilt $srs \in (rs + s)^*rs$, aber $srs \notin r(sr + r)^*s$* •

8. *Ein Mealy-Automat ist ein DEA mit Ausgabe.* •

9. $\hat{\delta}(q_0, w) \cap F \neq \emptyset$ •

10. *die ε-Hülle von p* •

11. *Wenn jedes Wort der Sprache von G genau eine Linksableitung (Ableitungsbaum) hat.* •

12. *Eine Sprache L ist inhärent mehrdeutig, wenn* jede *kontextfreie Grammatik, die L erzeugt, mehrdeutig ist (oder es keine eindeutige kfG gibt, die L erzeugt).* •

13. *Die Grammatik ist kontextfrei* •

14. *Es werden nur die Teilmengen generiert, die von der ε-Hülle des Startzustands aus erreichbar sind.* •

Aufgabe 2 (Nichtdeterministische und deterministische Automaten, 17 Punkte)

Ein (zugegebenermaßen etwas seltsamer) Parkautomat akzeptiert Münzen zu 50c, 1€ und 2€ und druckt nur dann einen Parkschein aus, wenn ein glatter Eurobetrag eingeworfen wurde, auf jedes 50c Stück ein zweites 50c Stück folgt und als letztes entweder eine 1€ Münze oder eine 1€ Münze direkt gefolgt von einer 2€ Münze eingeworfen wurde. Die Sprache L beschreibe nun alle Serien von Münzeinwürfen, die zum Ausdruck eines Parkscheins führen.

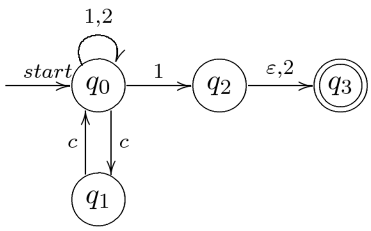
1. Erklären sie den syntaktischen Aufbau der Sprache L und geben Sie eine kurze und möglichst einfache Beschreibung der Sprache L über dem Alphabet $\{c, 1, 2\}$ (das Symbol c steht für die 50c-Münze) an, die keine natürliche Sprache verwendet. **(2 Punkte)**
2. Geben Sie einen NEA A mit 4 Zuständen, genau einem Endzustand und $L(A) = L$ an! **(4 Punkte)**
3. Zeigen Sie, dass Ihr Automat A tatsächlich die Sprache L akzeptiert! **(7 Punkte)**
4. Konstruieren Sie einen zu A äquivalenten DEA A'! **(4 Punkte)**

Lösung

1. *Die Bedingung, daß auf jedes 50c Stück ein weiteres 50c Stück folgen muß, sorgt bereits dafür, daß nur glatte Eurobeträge eingeworfen werden können um einen Parkschein zu erhalten.* •
   Es reicht also, die korrekte Reihenfolge der Münzeinwürfe einzuhalten: beliebige Folgen von cc,1 oder 2 und zum Schluß 1 oder 12.
   $L = \{uv \in \{c,1,2\}^* \mid u \in \{cc,1,2\}^* \wedge v \in \{1,12\} = L((cc+1+2)^*(1+12))$ •

2. *Bedeutung der Zustände des Automaten:*
   - q_0: *Glatter Eurobetrag ist eingeworfen – Wörter aus* $L((cc+1+2)^*)$
   - q_1: *letzte Münze war ein c – Wörter aus* $L((cc+1+2)^*)c$
   - q_2: *letzte Münze war ein 1€ – Wörter aus* $L((cc+1+2)^*)1$
   - q_3: *letzte Münze war ein 1€ oder 2€ nach 1€ – Wörter aus* $L((cc+1+2)^*(1+12))$ ••

   $A = (\{q_0,q_1,q_2,q_3\},\{c,1,2\},\delta,q_0,\{q_3\})$ ••



3. *Zu zeigen ist:* $L(A) = L$, *d.h.* $\forall w \in \{c,1,2\}^*.\, \hat{\delta}(q_0,w) = q_3 \Leftrightarrow w \in L$ •
   Wir zeigen durch simultane Induktion über die Wortlänge von $w \in \{c,1,2\}^*$:
   (1) $q_0 \in \hat{\delta}(q_0,w) \Leftrightarrow w \in L((cc+1+2)^*)$
   (2) $q_1 \in \hat{\delta}(q_0,w) \Leftrightarrow w \in L((cc+1+2)^*c)$
   (3) $q_3 \in \hat{\delta}(q_0,w) \Leftrightarrow w \in L$ •

   IAnf: *Es gilt* $q_0 \in \hat{\delta}(q_0,\varepsilon)$ *und* $\varepsilon \in L((cc+1+2)^*)$. *Damit sind linke und rechte Seite von (1) erfüllt und keine der Seiten von (2) bzw. (3). Also gelten alle drei Äquivalenzen für* $w = \varepsilon$. •

   IA: *Die drei Äquivalenzen seien für ein* $u \in \{c,1,2\}$ *bewiesen.* •

   IS: *Sei* $w = ua$ *für ein* $a \in \{c,1,2\}$. *Es gilt*

   (1) $q_0 \in \hat{\delta}(q_0,ua) \;\Leftrightarrow\; (q_0 \in \hat{\delta}(q_0,u) \wedge a \in \{1,2\}) \vee (q_1 \in \hat{\delta}(q_0,u) \wedge a = c)$
   $\Leftrightarrow (u \in L((cc+1+2)^*)\ \text{IA(1)}\ \wedge a \in \{1,2\}) \vee (u \in L((cc+1+2)^*c)\ \text{IA(2)}\ \wedge a = c)$
   $\Leftrightarrow ua \in L(((cc+1+2)^*)$ •

   (2) $q_1 \in \hat{\delta}(q_0,ua) \;\Leftrightarrow\; q_0 \in \hat{\delta}(q_0,u) \wedge a = c$
   $\Leftrightarrow (w \in L((cc+1+2)^*)\ \text{IA(1)}\ \wedge a = e$
   $\Leftrightarrow ua \in L((cc+1+2)^*c)$ •

   (3) $q_3 \in \hat{\delta}(q_0,ua) \;\Leftrightarrow\; q_2 \in (\hat{\delta}(q_0,u) \wedge a = 2) \vee (q_0 \in \hat{\delta}(q_0,u) \wedge a = 1)$
   $\Leftrightarrow (u = vb \wedge \hat{\delta}(q_0,v) = q_0 \wedge b = 1 \wedge a = 2) \vee (u \in L((cc+1+2)^*)\ \text{IA(1)}\ \wedge a = 1)$
   $\Leftrightarrow (ua = v12 \wedge v \in L((cc+1+2)^*)\ \text{IA(1)}\ ) \vee (ua \in L \wedge a = 1) \;\Leftrightarrow\; ua \in L$ •

   Fehlt das „zu zeigen: ", muß hier ein Schlusssatz kommen, der zeigt, was mit der Induktion bewiesen wurde.

4. *Konstruktion erfolgt mit der optimierten Teilmengenkonstruktion:* •

δ'	c	1	2	
$\rightarrow Q_0$	Q_1	Q_{023}	Q_0	1/2
Q_1	Q_0	$Q_\emptyset$	$Q_\emptyset$	1/2
$*Q_{023}$	Q_1	Q_{023}	Q_{03}	1/2
$*Q_{03}$	Q_1	Q_{023}	Q_0	1/2
$Q_\emptyset$	$Q_\emptyset$	$Q_\emptyset$	$Q_\emptyset$	1/2

$A' = (\{Q_0,q_1,Q_{023},Q_{03},Q_\emptyset\},\{c,1,2\},\delta',Q_0,\{Q_{023},Q_{03}\})$ 1/2

Aufgabe 3 (Grammatiken, NEAs und reguläre Ausdrücke, 12 Punkte)

Gegeben sei die rechtslineare Grammatik $G = (\{S, X, Y, Z\}, \{0, 1\}, P, S)$ mit der Produktionenmenge $P = \{S{\rightarrow}0X|1Y,\ X{\rightarrow}0S|1Y|0Z,\ Y{\rightarrow}1S|0X|1Z,\ Z{\rightarrow}0X|1Y|\varepsilon\}$

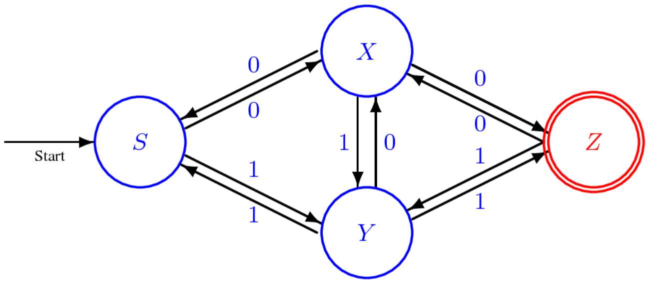
1. Wandeln Sie die Grammatik G mit dem aus der Vorlesung bekannten Verfahren in einen NEA A um, in dem keine ε-Übergänge vorkommen. **(4 Punkte)**
2. Wandeln Sie den Automaten A mit Hilfe der Zustandselemination in den regulären Ausdruck R um! Kürzen Sie dabei alle entstehenden regulären Ausdrücke unter Angabe einer Begründung so weit wie möglich. **(8 Punkte)**

Lösung

1. *Das bekannte Verfahren wandelt eine Grammatik* $G = (V,T,P,S)$ *um in einen Automaten* $A := (V, T, \delta, S, F)$ *mit* $\delta(X,a) = \{X' \mid X{\to}aX' \in P\}$ *und* $F = \{X \in V \mid X{\to}\varepsilon \in P\}$.

   Für die in der Aufgabenstellung gegebene Grammatik entsteht dabei der Automat

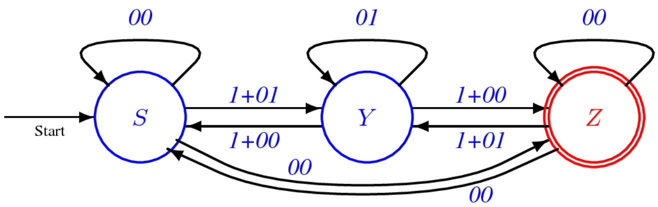
   $A = (\{S,X,Y,Z\},\{0,1\},\delta,S,\{Y,Z\})$, •
   wobei δ *durch das folgende Diagramm beschrieben ist.* • • •

   Jeder falsche Pfeil – ob nun zuviel, zuwenig, oder falsch markiert – ergibt einen Punkt Abzug. Wenn zuvor das allgemeine Verfahren zitiert wurde, ist es offensichtlich, daß Flüchtigkeitsfehler vorliegen und es gibt nur jeweils einen halben Punkt Abzug.

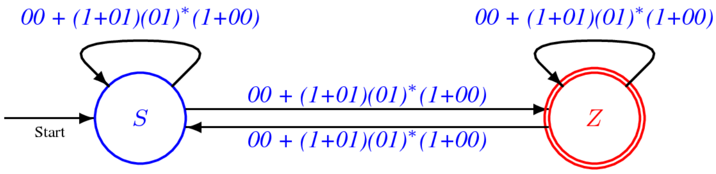
2. *Wir wandeln den NEA zunächst in einen VNEA um, bei dem die Kanten mit regulären Ausdrücken markiert sind. Da an jeder Kante nur ein Symbol steht ändert sich hierdurch gar nichts.* •

   Es liegt nahe, als erstes den Nicht-Endzustand X *(oder* Y*) zu eliminieren. Dies liefert* • • •

   Je ein Punkt für die richtigen Pfeile bei S, Y und Z Jeder falsche Pfeil ergibt einen halben Punkt Abzug.

   Als nächstes eliminieren wir den Zustand Y*. Dies ergibt* ••

   Je ein Punkt für die richtigen Pfeile bei S und Z Jeder falsche Pfeil ergibt einen halben Punkt Abzug.

   Wir lesen hieraus einen regulären Ausdruck ab, wie in Einheit 2.3, Folie 21 beschrieben:

   $E = (R^* + SU^*T)^*SU^*$ *mit* $R = S = U = T = 00 + (1+01)(01)^*(1+00)$ •

   Da alle Komponenten identisch sind, vereinfacht sich $E = (R^* + RR^*R)^*RR^*$ *zu dem regulären Ausdruck* $E = R^+ = (00 + (1+01)(01)^*(1+00))^+$ •

Aufgabe 4 (Kontextfreie Grammatiken, Ableitungsbäume, 15 Punkte)

Sei L die Menge der regulären Ausdrücke über den Terminalsymbolen $\Sigma = \{a, b, (,), +, ^*, \emptyset, e\}$. Dabei sollen a und b die einzigen Zeichen sein, die von regulären Ausdrücken in L erzeugt werden können. Um Verwechslungen vorzubeugen, wird in L das leere Wort durch e statt durch ε dargestellt. Ein wohlgeformter regulärer Ausdruck aus L ist zum Beispiel $(a + b + e)^*ab$.

1. Geben Sie eine einfache kontextfreie Grammatik mit Terminalalphabet Σ und nur einer Variablen an, die genau die regulären Ausdrücke über dem Alphabet $\{a, b\}$ erzeugt. Begründen Sie kurz die Korrektheit Ihrer Grammatik. **(3 Punkte)**
   Da die Grammatik möglichst einfach sein soll, wird sie wahrscheinlich mehrdeutig sein.
2. Geben Sie eine linksseitige und zwei rechtsseitige Ableitungen für den Ausdruck $(a + b + e)^*ba$ an. Geben Sie für eine dieser Ableitungen einen Ableitungsbaum an. **(7 Punkte)**
3. Geben Sie eine eindeutige kontextfreie Grammatik für L an, welche die üblichen Vorrangs- und Assoziativitätsregeln beachtet (Klammern binden am stärksten, dann folgen *, Verkettung und $+$; Verkettung und $+$ sind beide linksassoziativ). Begründen Sie die Eindeutigkeit indem Sie den Zweck der von Ihnen verwendeten Variablen erklären. **(5 Punkte)**

Lösung

1. *Die Definition der Syntax regulärer Ausdrücke auf Folie 5 von Einheit 2.3 läßt sich (unter Wegfall des Konkatenationssymbols $\circ$) direkt in die folgende Grammatik umsetzen:* •

   $G = (\{E\}, \Sigma, P, E)$ *mit* $P = E \rightarrow \emptyset \mid e \mid a \mid b \mid (E) \mid E^* \mid E + E \mid EE.$ ••

2. *Linksseitige Ableitung:*
   $E \rightarrow EE \rightarrow E^*E \rightarrow (E)^*E \rightarrow (E + E)^*E \rightarrow (E + E + E)^*E \rightarrow (a + E + E)^*E$
   $\rightarrow (a+b+E)^*E \rightarrow (a+b+e)^*E \rightarrow (a+b+e)^*EE \rightarrow (a+b+e)^*bE \rightarrow (a+b+e)^*ba.$ ••

   rechtsseitige Ableitungen:
   $E \rightarrow EE \rightarrow EEE \rightarrow EEa \rightarrow Eba \rightarrow E^*ba \rightarrow (E)^*ba \rightarrow (E + E)^*ba$
   $\rightarrow (E + e)^*ba \rightarrow (E + E + e)^*ba \rightarrow (E + b + e)^*ba \rightarrow (a + b + e)^*ba.$ ••
   $E \rightarrow EE \rightarrow Eb \rightarrow EEa \rightarrow Eba \rightarrow E^*ba \rightarrow (E)^*ba \rightarrow (E + E)^*ba$
   $\rightarrow (E + e)^*ba \rightarrow (E + E + e)^*ba \rightarrow (E + b + e)^*ba \rightarrow (a + b + e)^*ba.$ •

   Nur ein Punkt für die zweite Ableitung, weil sie der ersten sehr ähnlich ist. Zwei Punkte für den Baum. ••

3. *Die **A**usdrücke niedrigster Prioriät stehen "außen" und werden linkssassoziativ aus **T**ermen aufgebaut. Wenn sie innerhalb von Termen erscheinen, müssen sie geklammert sein.* •

   *Terme bestehen aus linkssassoziativen Verkettungen von **K**leene-Stern Operationen. Wenn sie innerhalb von Stern-Ausdrücken erscheinen, müssen sie geklammert sein.* •

   *Stern-Ausdrücken sind einfache oder gesternte **B**asis-Ausdrücke, die wiederum aus $\emptyset, e, a, b$ oder geklammerten Ausdrücken bestehen.* •

   Durch diese Auftrennung entsteht eine eindeutige Grammatik $G = (\{E, T, K, B\}, \Sigma, P, E)$, wobei P folgende Regeln enthält: ••

$$
\begin{aligned}
E &\rightarrow E + T \mid T \\
T &\rightarrow TK \mid K \\
K &\rightarrow B^* \mid B \\
B &\rightarrow \emptyset \mid e \mid a \mid b \mid (E)
\end{aligned}
$$

   (ohne tieferen Beweis der Eindeutigkeit)

Aufgabe 5 (Nichtreguläre Sprachen, 15 Punkte)

1. Beweisen mit Hilfe der Abschlusseigenschaften regulärer Sprachen, daß die Sprache $L_1 = \{(ab)^n c^m (ba)^{2n} \mid m,n \in \mathbb{N}\}$ nicht regulär ist. **(4 Punkte)**
2. Zeigen Sie, daß die Sprache L_1 unendlich viele Äquivalenzklassen hat **(5 Punkte)**
3. Beweisen Sie mit Hilfe des Pumping Lemmas, dass die Sprache $L_2 = \{0^a 1^b 2^c \mid a,b,c \in \mathbb{N}^+ \wedge b{=}a^c\}$ nicht regulär ist! **(6 Punkte)**

Lösung

1. *Wir versuchen, die Sprache L_1 mit der bekannten nicht-regulären Sprache $L_{01} = \{0^n 1^n \mid n \in \mathbb{N}\}$ in Verbindung zu bringen. Dazu wählen wir einen Homomorphismus $h : \{0,1\}^* \to \{a,b,c\}^*$ mit $h(0) = ab$ und $h(1) = baba$.* •
   Man beachte, daß mit diesem Homomorphismus kein 'c' erzeugt wird, also nur auf Wörter über $\{a,b\}^*$ abgebildet wird. •
   Deshalb ist $h^{-1}(L_1) = \{w \in \{0,1\}^* \mid h(w) \in L_1\} = \{w \in \{0,1\}^* \mid \exists n \in \mathbb{N}.h(w) = (ab)^n (ba)^{2n}\}$
   $= \{w \in \{0,1\}^* \mid \exists n \in \mathbb{N}.w = 0^n 1^n\} = L_{01}$ •
   Da L_{01} nicht regulär ist und die regulären Sprachen unter inversen Homomorphismen abgeschlossen sind, kann L_1 nicht regulär sein. •
   Es muß nicht so ausführlich sein, aber eine Begründung, warum das c wegfällt, muß erkennbar sein. Man kann übrigens auch zuerst mit $L((a+b)^*)$ oder dergleichen schneiden.
2. $\Sigma^*/L_1 = \{[\varepsilon]_{L_1}, [a]_{L_1}, [ab]_{L_1}, [aba]_{L_1}, [abab]_{L_1}, \ldots\}$, *denn:* •
   - $[\varepsilon]_{L_1}$ *enthält genau die Wörter der Sprache L_1* 1/2
   - $[a]_{L_1}$ *enthält Wörter, die um $b(ab)^n c^m (ba)^{2n+2}$ ergänzt werden müssen um ein Wort der Sprache zu erzeugen. Dies ist nur das Wort a, also $[a]_{L_1} = \{a\}$* 1/2
   - $[b]_{L_1}$ *enthält alle Wörter, die sich nicht zu einem Wort aus L_1 fortsetzen lassen, also z.B. $b, aa, ba, bb, aab, ..$* 1/2
   - $[ab]_{L_1}$ *enthält Wörter, die genau um Wörter der Form $(ab)^n c^m (ba)^{2n+2}$ ergänzt werden dürfen, um ein Wort der Sprache zu erzeugen. Dies gilt nur für ab (läßt sich um $baba$ und $abbababa$ ergänzen, was für kein anderes Wort gilt).* 1/2
   - $[aba]_{L_1}$ *enthält Wörter, die um Wörter der Form $a(ab)^n c^m (ba)^{2n+4}$ ergänzt werden dürfen, um ein Wort der Sprache zu erzeugen. Dies ist nur das Wort aba, also $[a]_{L_1} = \{aba\}$* 1/2
   - $[abab]_{L_1}$ *enthält Wörter, die genau um Wörter der Form $(ab)^n c^m (ba)^{2n+4}$ ergänzt werden dürfen, um ein Wort der Sprache zu erzeugen. Dies gilt nur für $abab$.* 1/2

   $\vdots$

   Setzt man dieses Argument fort so sieht man, daß jedes $(ab)^k$ und jedes $(ab)^k a$ eine eigene Äquivalenzklasse bildet. Damit gibt es unendlich viele Klassen. •

   Man kann das auch präziser formulieren, aber das ist nicht notwendig. Die Klassen $[aba]_{L_1}$ und $[abab]_{L_1}$ müssen nicht explizit aufgelistet werden, wenn erklärt wird, daß jedes $(ab)^k$ und jedes $(ab)^k a$ eine eigene Äquivalenzklasse bildet.
3. *Sei $n \in \mathbb{N}$ beliebig. Wir wählen $w = 0^a 1^b 2^c \in L_2$ mit $b{=}a^c$, $c{>}0$ und $a{>}n$. Also gilt $|w| > n$.* •

   Sei $w = xyz$ eine beliebige Zerlegung mit $y{\neq}\varepsilon$ und $|xy|{\leq}n$. •

   Dann gilt $x = 0^i$, $y = 0^j$ und $z = 0^{a-i-j} 1^b 2^c$ für ein $j{>}0$ und $i{+}j{\leq}n$. •

   Wir wählen $k = 2$. Dann ist $xy^2 z = 0^{a+j} 1^b 2^c$ •
   und es folgt $xy^2 z \notin L_2$, denn für $j{>}0$ ist $b{=}a^c < (a{+}j)^c$. •

   Damit ist die im Pumping Lemma beschriebene Eigenschaft regulärer Sprachen verletzt. Aufgrund des Pumping Lemmas kann L_2 also nicht regulär sein. •

Aufgabe 6 (Abschlusseigenschaften, 10 Punkte)

Sei L eine reguläre Sprache. Zeigen Sie, dass die Sprache

$$\mathrm{noprefix}(L) = \{w \in L \mid \forall u, v \in \Sigma^*. (w{=}uv \wedge v{\neq}\varepsilon) \Rightarrow u \notin L\}$$

regulär ist. Konstruieren Sie dazu aus einem endlichen Automaten A für L einen endlichen Automaten A', der die Sprache $\mathrm{noprefix}(L)$ akzeptiert und begründen Sie die Korrektheit Ihrer Konstruktion.

Konstruktion und Begründungen sollten präzise sein, aber ein formaler Beweis ist nicht erforderlich.

Lösung *Die Sprache ist identisch mit der Sprache $min(L)$ aus Übung 6.1. Wer das erkennt, ist um einiges weiter. Allerdings ist der Hinweis darauf alleine keine akzeptable Lösung. Es müßte relativ genau gezeigt werden, daß dies wirklich dasselbe ist. Die Erklärung unten benutzt zwar auch "echter Präfix", aber die eigentliche Konstruktion fällt immer auf die Definition von $\mathrm{noprefix}(L)$ zurück.*

Die Eigenschaft $w{=}uv \wedge v{\neq}\varepsilon$, die wir im folgenden zuweilen mit $u \sqsubset w$ abkürzen, drückt aus, daß u ein echter Präfix von w ist. Ein Wort w gehört zur Sprache $\mathrm{noprefix}(L)$, wenn es selbst zu L gehört, aber kein echter Präfix u von w in L liegt. Anders ausgedrückt: wenn $w \in L$ zu $\mathrm{noprefix}(L)$ gehört, dann darf keine echte Erweiterung $w' \in L$ von w zu $\mathrm{noprefix}(L)$ gehören (ansonsten wäre ja $w \in L$ ein echter Präfix von diesem Wort). ••

Ein Automat $A' = (Q', \Sigma, \delta', q_0', F')$ für $\mathrm{noprefix}(L)$ darf nur Wörter aus L akzeptieren, aber er darf keine Wörter akzeptieren, die Erweiterungen von bereits akzeptierten Wörtern sind. Wenn also die Verarbeitung eines Wortes einen Endzustand von A' passiert hat, darf sie danach nie wieder in einen Endzustand geraten, d.h. aus $\delta'(q_0', w) \in F'$ folgt $\delta(q_0', wv) \notin F$ für alle $v{\neq}\varepsilon$. ••

Dies kann erreicht werden, indem man δ so verändert, daß aus einem Endzustand kein anderer Endzustand erreicht werden kann. Hierzu entfernt man aus δ einfach alle Übergänge, die in einem Zustand aus F beginnen, oder – wenn der Automat deterministisch bleiben soll, lenkt man sie alle in einen neuen Zustand q_f um. Ansonsten bleibt der Automat unverändert. ••

Es sei also $A = (Q, \Sigma, \delta, q_0, F)$ ein DEA mit $L(A) = L$. Wir konstruieren $A' = (Q' \cup \{q_f\}, \Sigma, \delta', q_0, F)$

mit $\delta'(q,a) = \begin{cases} q_f & \text{falls } q \in F \cup \{q_f\} \\ \delta'(q,a) & \text{sonst} \end{cases}$ ••

Aus dieser Definition folgt $\delta'(q,v) = q_f$ für alle $q \in F \cup \{q_f\}$ und $v{\neq}\varepsilon$ und $\delta'(q_0, w_1..w_n) = \delta(q_0, w_1..w_n)$, falls $\delta(q_0, w_1..w_i) \notin F$ für alle $i{<}n$. Damit folgt insbesondere aus $\delta'(q_0, w) \in F$, daß $\delta'(q_0,u){=}\delta(q_0,u) \notin F$ für alle $u \sqsubset w$ gilt. Keine Beweise erforderlich

Damit gilt $w \in L(A')$
$\Leftrightarrow\ \delta'(q_0', w) \in F$
$\Leftrightarrow\ \delta'(q_0', w) \in F\ \wedge\ \delta'(q_0', u) \notin F$ *für alle* $u \sqsubset w$
$\Leftrightarrow\ \delta(q_0', w) \in F\ \wedge\ \delta(q_0', u) \notin F$ *für alle* $u \sqsubset w$
$\Leftrightarrow\ w \in L\ \wedge\ u \notin L$ *für alle* $u \sqsubset w$
$\Leftrightarrow\ w \in \mathrm{noprefix}(L)$ ••

Es gibt viele Punkte für eine gründliche verbale Erklärung der Konstruktion und nur je 2 Punkte für die tatsächliche Beschreibung von A' und den Beweis. Fehlt die verbale Erklärung zu Beginn und der Beweis, so kann man für eine korrekte Konstruktion maximal 5 Punkte geben. Der Text darf natürlich viel knapper sein.

Die folgende Lösung wäre auch akzeptabel: Man bestimmt im Automaten A alle von F erreichbaren Endzustände und entfernt diese aus F um F' zu erhalten. δ selbst bleibt unverändert.
Wenn dabei $q_0 \in F$ war und entfernt werden müßte, gilt $\varepsilon \in L$ und somit $\mathrm{noprefix}(L) = \{\varepsilon\}$. In diesem Fall ist $\mathrm{noprefix}(L)$ offensichtlich regulär, aber es müßte ein Automat für $\{\varepsilon\}$ konstruiert werden.

Aufgabe 7 (Minimierung von DEAs, 10 Punkte)

Gegeben sei der durch die folgende Übergangstabelle definierte DEA.

	0	1
$\rightarrow A$	C	G
B	F	B
$*C$	E	C
D	A	D
E	A	H
F	J	G
G	D	A
H	A	D
I	C	G
J	B	D

1. Entfernen Sie die unerreichbaren Zustände aus dem Automaten. **(3 Punkte)**
2. Ermitteln Sie mit Hilfe des Table-Filling Verfahrens alle Paare äquivalenter Zustände. Um die einzelnen Runden erkennbar zu machen, kennzeichnen Sie bitte in jeder Runde die neuen Einträge mit einem neuen Symbol (verwenden Sie der Reihe nach $\times$, $+$, $*$, $\circ$, $\#$, $\ldots$) **(5 Punkte)**
3. Geben Sie einen äquivalenten minimalen DEA an. **(2 Punkte)**

Lösung

1. *A ist der Startzustand. Erreichbar sind der Reihe nach: C, G; E, D; H.* ••
   Damit sind B, F, I, J unerreichbar und können entfernt werden. •
2. *Wir (stutzen die Tabelle und) wenden auf die restlichen Zustände das Table-Filling Verfahren an.*
   C ist der einzige Endzustand – alle anderen sind hiervon unterscheidbar •

	A	C	D	E	G	H
A	\	×				
C	×	\	×	×	×	×
D		×	\			
E		×		\		
G		×			\	
H		×				\

.

	0	1
→A	C	G
*C	E	C
D	A	D
E	A	H
G	D	A
H	A	D

Eingabe 0: nur A führt zu akzeptierenden Zuständen. D, E, G und H sind hiervon unterscheidbar. •

	A	C	D	E	G	H
A	\	×	+	+	+	+
C	×	\	×	×	×	×
D	+	×	\			
E	+	×		\		
G	+	×			\	
H	+	×				\

Eingabe 1: nur C führt zu akzeptierenden Zuständen. Alle anderen sind hiervon unterscheidbar (keine neue Information).

In der nächsten Runde überprüfen wir die verbleibenden 6 Paare aus D, E, G und H.

Eingabe 0: G führt zu D, alle anderen zum nichtäquivalenten A •

	A	C	D	E	G	H
A	\	×	+	+	+	+
C	×	\	×	×	×	×
D	+	×	\		*	
E	+	×		\	*	
G	+	×	*	*	\	*
H	+	×			*	\

Eingabe 1: G führt zu A, alle anderen zu D oder H (keine neue Information).

In der nächsten Runde überprüfen wir die verbleibenden 3 Paare $\{D, E\}$, $\{D, H\}$ und $\{E, H\}$. Dies ergibt weder bei 0 noch bei 1 neue Unterschiede. •

Damit ist die obige Tabelle endgültig. Die Äquivalenzklassen sind $\{A\}$, $\{C\}$, $\{D, E, H\}$ und $\{G\}$. •

3. *Der resultierende minimale Automat ist $A_{min} = (\{A, C, DEH, G\}, \{0, 1\}, \delta', A, \{C\})$, wobei δ' durch folgende Tabelle definiert ist.* ••

	0	1
→A	C	G
*C	DEH	C
DEH	A	DEH
G	DEH	A