Prof. Javier Esparza
Philipp Czerner, Martin Helfrich

Technische Universität München
Lehrstuhl für Theoretische Informatik

Einführung in die Theoretische Informatik

Sommersemester 2021 – Hausaufgabenblatt 4

Abgabe: 17.05.2021, 12:00 CEST

- Beachten Sie die Abgabemodalitäten auf der Vorlesungswebsite!
- Sei $\Phi := \{\{1\}, \{2\}, \{3, 4\}, \{5, 6\}, \{7\}\}$. Nach dem Abgabedatum werden wir für jede Menge $A \in \Phi$ eine zufällige Aufgabe $a \in A$ wählen und korrigieren.
- Es werden diese Aufgaben korrigiert: H4.1, H4.2, H4.3, H4.6, H4.7
- Wenn Sie einen Beweis aufstellen, von dem Sie wissen, das einzelne Schritte problematisch oder unvollständig sind, merken Sie dies bitte in Ihrer Lösung an, damit wir das bei der Korrektur positiv berücksichtigen können.

Aufgabe H4.1. (*vero nihil verius*) 0.5+0.5+0.5+0.5+1 Punkte

Bestimmen Sie für folgende Aussagen, ob sie wahr oder falsch sind. Geben sie eine kurze Begründung für wahre Aussagen an und ein Gegenbeispiel für falsche.

Sei $\Sigma \neq \emptyset$ ein Alphabet, $L_1, L_2 \subseteq \Sigma^*$ reguläre Sprachen, und $L_3, L_4 \subseteq \Sigma^*$ nicht-reguläre Sprachen.

(a) $(L_1 \cap L_2)^2$ ist regulär.

(b) $L_3 \cup L_4$ ist nicht regulär.

(c) L_4 erfüllt die Eigenschaft des Pumping-Lemmas nicht.

(d) Jede Teilmenge $L \subseteq L_1$ ist regulär.

(e) $(L_3)^*$ ist regulär.

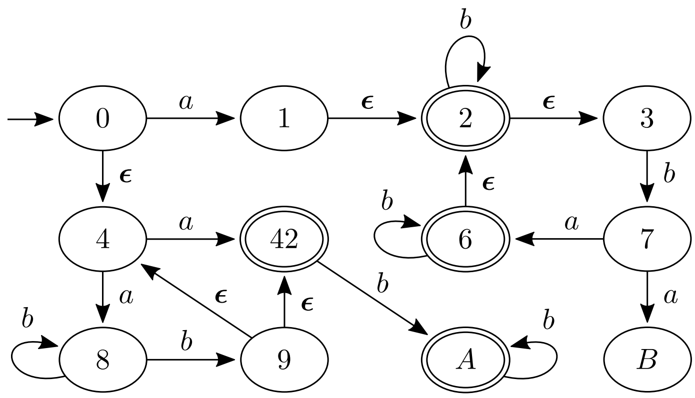
Aufgabe H4.2. (*in nuce*) 1+1 Punkte

(a) Vereinfachen Sie den folgenden ϵ-NFA M, konstruieren Sie also einen äquivalenten ϵ-NFA mit höchstens 6 Zuständen.

(b) Finden Sie einen regulären Ausdruck r mit $|r| \leq 6$ und $L(r) = L(M)$.

Beschreiben Sie bei beiden Teilaufgaben ihr Vorgehen.

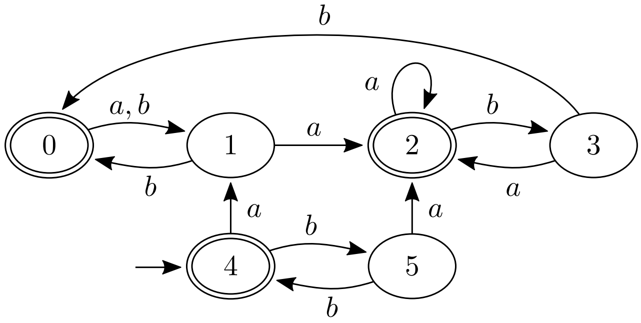
Hinweise: Bei der Länge eines regulären Ausdrucks zählen wir Klammern nicht, die Länge von $(ab \mid \epsilon)^*$ wäre also 5. Zur Lösung dieser Aufgabe sind keine Kenntnisse über minimale DFAs vonnöten.



Aufgabe H4.3. (*a maiore ad minus*) 2+1 Punkte

(a) Bestimmen Sie für folgenden DFA M die Äquivalenzklassen von $\equiv_M$ mit der Erweiterung des in der Vorlesung vorgestellten Verfahrens aus Ü4.4. Geben Sie also insbesondere für jedes Paar an Zuständen (q, r), das sie „markieren“, ein Wort w an, das sie unterscheidet (d.h. von $\delta(q, w), \delta(r, w)$ ist genau einer akzeptierend). Es steht ihnen frei, auch die optimierte Version des Verfahrens mit Abhängigkeitsanalyse zu verwenden.

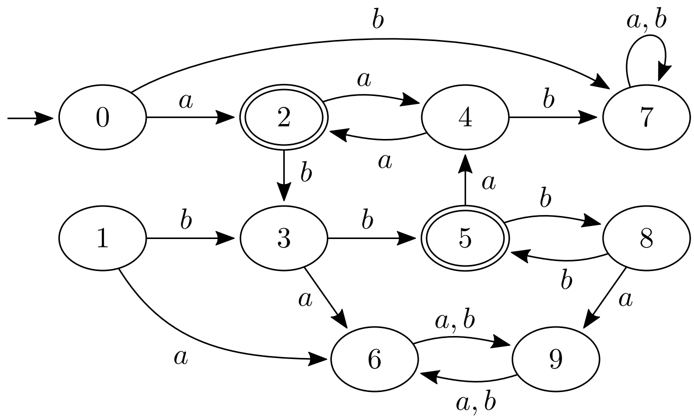
(b) Erzeugen Sie einen minimalen DFA äquivalent zu M, indem sie die äquivalenten Zustände von M kollabieren.



Aufgabe H4.4. (*a maximo ad minimum*) 2+1 Punkte

(a) Bestimmen Sie für folgenden DFA M' die Äquivalenzklassen von $\equiv_{M'}$. Sie müssen hierzu nicht nach dem in der Vorlesung vorgestellten Verfahren vorgehen.

(b) Minimieren Sie nun M' mit dem Minimierungsalgorithmus aus der Vorlesung und geben Sie einen regulären Ausdruck für $L(M')$ an.



Aufgabe H4.5. (*ejusdem generis*) 1+3 Punkte

(a) Sei $\Sigma := \{a\}$, $k \in \mathbb{N}_{>0}$ beliebig und $L := \{a^{ik} : i \in \mathbb{N}\}$ die Sprache der Wörter deren Länge ein Vielfaches von k ist. Bestimmen sie die Äquivalenzklassen von $\equiv_L$ (s. Definition 3.55), und zeigen Sie so, dass der minimale DFA für L genau k Zustände hat.

(b) Sei $r := ab^* \mid ba^*$ ein regulärer Ausdruck über $\Sigma := \{a, b\}$. Konstruieren Sie den kanonischen Minimalautomaten für $L(r)$ direkt nach Definition 3.59. Beschriften Sie hierfür jeden Zustand mit einem regulären Ausdruck, dessen erzeugte Sprache der Äquivalenzklasse des Zustandes entspricht. (Der Startzustand wäre z.B. mit ϵ beschriftet.) Beschreiben Sie ihr Vorgehen. Sie müssen nicht beweisen, dass Sie jeder Äquivalenzklasse einen passenden regulären Ausdruck zugeordnet haben.

Achtung: Wir betrachten die Äquivalenzklassen bezüglich $\equiv_{L(r)}$, nicht die der Zustände des Automaten!

Aufgabe H4.6. (*sic parvis magna*) 1+2+1 Punkte

Dora hat ihren Kindergartenprofessor noch einmal gefragt, und weiß nun, wie die Produkt-Konstruktion funktioniert. Jetzt ist ihr aber aufgefallen, dass der entstandene Produktautomat nicht minimal sein muss! Entrüstet möchte sie zu ihrem Kindergartenprofessor stampfen, ihn mit Sand bewerfen, und sich darüber beschweren, dass er ihr ein offensichtlich schlechtes Verfahren beigebracht hat. Können Sie Dora beschwichtigen, indem Sie demonstrieren, dass die ihr bekannten Verfahren, um Mengenoperationen auf DFAs auszuführen, zumindest manchmal ein bestmögliches Ergebnis liefern?

(a) Sei L eine reguläre Sprache und $M = (Q, \Sigma, \delta, q_0, F)$ ein minimaler DFA für L. Zeigen Sie, dass es einen minimalen DFA für $\overline{L}$ mit $|Q|$ Zuständen gibt.

(b) Zeigen Sie, dass es für jede Zahl $n \in N$ zwei DFAs $M_1 = (Q_1, \Sigma, \delta_1, q_{01}, F_1)$ und $M_2 = (Q_2, \Sigma, \delta_2, q_{02}, F_2)$ gibt, sodass $|Q_1|, |Q_2| \geq n$ und der minimale Automat für $L(M_1) \cap L(M_2)$ genau $|Q_1| \cdot |Q_2|$ Zustände hat.

(c) Beweisen Sie, dass es für jede Zahl $n \in N$ zwei DFAs $M_3 = (Q_3, \Sigma, \delta_3, q_{03}, F_3)$ und $M_4 = (Q_4, \Sigma, \delta_4, q_{04}, F_4)$ gibt, sodass $|Q_3|, |Q_4| \geq n$ und der minimale Automat für $L(M_3) \cup L(M_4)$ genau $|Q_3| \cdot |Q_4|$ Zustände hat.

Hinweis: Es mag hilfreich sein, Ergebnisse aus Aufgabe H4.5 zu verwenden.

Bonusaufgabe H4.7. (*inveniet quod quisque velit*) 3 Bonuspunkte

Der Superschurke Dr. Evilsparza hat jüngst ein Pharmaunternehmen (Evillest Evilness *Inc.*) gegründet, und behauptet nun, einen Impfstoff gegen schlechtes Wetter entwickelt zu haben. Die Öffentlichkeit ist begeistert – Sie trauen dem angeblich reformierten Sünder allerdings nicht und versuchen, in der DNA-Sequenz nach Hinweisen zu suchen.

Bekanntermaßen werden in der DNA vier verschiedene Basen kodiert: Evalcyclohexandifluorit, Isophoronditrigocyanat, Vinylcyclohexenquadroxid, und Levomethagameorphan. Diese werden mit ihrem Initialbuchstaben abgekürzt, sodass letztendlich ein Wort über dem Alphabet $\Sigma := \{e, i, l, v\}$ entsteht. Sie sind in Besitz einer Sequenz, mit der Dr. Evilsparza in 2006 versuchte, die gesamte Studierendenschaft zu mathematischen Zombies zu wandeln, um die Universitätsherrschaft an sich zu reißen. Sie vermuten, dass er diese Sequenz wieder verwendet – aber Sie müssen Beweise finden, und die Zeit drängt!

Unter diesem Link finden Sie sowohl die Impfstoff-Sequenz w, als auch die Zombie-Sequenz s. Bestimmen Sie, ob s in w enthalten ist, und geben sie den kleinsten Index i an, sodass $w_1 w_2 ... w_i$ bereits s enthält. Beschreiben Sie ihr Vorgehen.

Hinweise: Verwenden Sie zur Lösung dieser Aufgabe einen Computer. Beschreiben Sie bitte Ihren Ansatz in *natürlicher Sprache* und illustrieren die wesentlichen Schritte ihrer Lösung mit geeigneten Codefragmenten. Sie können (müssen aber nicht), ihrer Lösung ihren vollständigen Programmcode beifügen, jedoch steht es den Korrektoren frei, diesen zu ignorieren. Wenn Sie existierende Algorithmen verwenden, die nicht Teil der Vorlesung sind, beschreiben Sie bitte ausführlich, weshalb und auf welche Weise diese funktionieren.

Wir sind, wie immer, an Lösungen interessiert, die sich verallgemeinern lassen. Ihre Lösung sollte also für beliebige Wörter w und s funktionieren. Zusätzlich soll ihre Lösung effizient sein (im Bezug auf die theoretische Komplexität des Algorithmus), und innerhalb von wenigen Sekunden (aber auf jeden Fall innerhalb von einer Minute) terminieren. Sie dürfen Standard-Datenstrukturen (wie etwa Hashtabellen oder binäre Bäume) verwenden, ansonsten beschränken Sie sich bitte auf grundlegende Funktionalitäten verbreiteter Programmiersprachen.

Highscores: Die Musterlösung wurde nicht optimiert und terminiert in etwa 750ms. Falls ihr Programm schneller ist, können Sie gerne eine Mail an Philipp Czerner mit dem Betreff „THEO H4.7“ schreiben, in der Sie ihr Programm und Anweisungen, es auszuführen, anhängen. Wir veröffentlichen dann eine Bestenliste auf Zulip. (Um die Zeiten für die Bestenliste zu ermitteln, werden wir etwas andere $w, s \in \Sigma^*$ verwenden.)