Prof. Javier Esparza
Philipp Czerner, Kevin Kappelmann, Valentin Krasotin

Technische Universität München
Lehrstuhl für Theoretische Informatik

Einführung in die Theoretische Informatik

Sommersemester 2024 – Übungsblatt 13

- Dieses Übungsblatt besteht aus einer Auswahl von (Klausur)aufgaben aus vorherigen Jahren und beeinhaltet keinen neuen Stoff. Verwenden Sie die letzte Übung für allgemeine Fragen zum Vorlesungsstoff in Hinblick auf die Klausurvorbereitung. (siehe Aufgabe Ü13.1)
- Das Übungsblatt ist in zwei Teile gegliedert: den Vorbereitungsteil, den Sie vor der Übung selbstständig bearbeiten sollen, und den Übungs-/Nachbereitungsteil, der Aufgaben enthält, die in der Übung besprochen werden und von Ihnen anschließend zur Nachbereitung verwendet werden können.

Vorbereitung (vor der Übung selbständig zu bearbeiten)

Vorbereitungsaufgabe Ü13.1. (*letztes Tutorium ⇒ stellt eure Fragen*)

Diese Woche findet die letzte Übung statt. Bitte lassen Sie deshalb Ihrem Tutor rechtzeitig (also so früh wie möglich) Fragen auf Zulip zukommen, die sie noch haben. Z.B. könnten das allgemeine Fragen zu Konzepten wie *Reduktion* oder Techniken wie *Pumping Lemma* sein oder aber auch Fragen zu Übungsaufgaben oder Hausaufgaben, die Sie gerne besprechen würden. Bitte geben Sie in solchen Fällen immer das Kürzel der Aufgabe UND die konkrete Frage an.

Bitte geben Sie Ihrem Tutor genügen Zeit, sich auf die Fragen vorzubereiten. Falls Ihr Tutor es nicht mehr schaffen sollte, sich vorzubereiten, oder Sie nach dem letzten Tutorium noch Fragen haben, finden Sie natürlich auf Zulip noch Unterstützung.

Vorbereitungsaufgabe Ü13.2. (*Wichtige Begriffe*)

Überprüfen Sie, dass Sie die folgenden Begriffe oder Notationen korrekt definieren können.

- Satz von Cook-Levin
- Die Probleme 2KNF-SAT, COL,MENGENÜBERDECKUNG, CLIQUE, RUCKSACK, TSP,BIN PACKING, PARTITION

Vorbereitungsaufgabe Ü13.3. (*P-Reduktion*)

Zeigen Sie 2COL $\in P$ indem Sie 2COL $\leq_P$ 2KNF-SAT zeigen.

Lösungsskizze. (Analog zur Vorlesung (3COL zu SAT))

Gegeben $G = (V, E)$, konstruieren wir die $2KNF$-Formel F:

$$\bigwedge_{v \in V} ((x_{v,1} \vee x_{v,2}) \wedge (\neg x_{v,1} \vee \neg x_{v,2})) \wedge \bigwedge_{(u,v) \in E, i \in \{1,2\}} (\neg x_{u,i} \vee \neg x_{v,i})$$

Die Reduktion ist polynomiell: Es werden $\mathcal{O}(|V| + |E|)$ viele Klauseln erzeugt.

Die Reduktion ist korrekt: Sei $f : V \to \{1, 2\}$ eine 2-Färbung von G. Wir konstruieren die Belegung

$$\sigma(x_{v,i}) := \begin{cases} 1, & \text{falls } f(v) = i \\ 0, & \text{sonst} \end{cases}.$$

Somit $\sigma(x_{v,1}) \leftrightarrow \neg\sigma(x_{v,2})$. Da f eine 2-Färbung ist, gilt $f(u) \neq f(v)$ für $(u, v) \in E$. Somit $\neg\sigma(x_{u,i}) \vee \neg\sigma(x_{v,i})$ für $(u, v) \in E$. Somit ist σ eine erfüllende Belegung für F.

Sei F erfüllbar mit σ. Wir definieren

$$f(v) := \begin{cases} 1, & \text{falls } \sigma(x_{v,1}) \\ 2, & \text{sonst} \end{cases}.$$

Da $\neg\sigma(x_{u,1}) \vee \neg\sigma(x_{v,1})$ für $(u, v) \in E$, gilt $f(u) = 2 \vee f(v) = 2$. Da $\neg\sigma(x_{u,2}) \vee \neg\sigma(x_{v,2})$ für $(u, v) \in E$ und $\sigma(x_{w,1}) \leftrightarrow \neg\sigma(x_{w,2})$ für $w \in V$, gilt $f(u) = 1 \vee f(v) = 1$. Somit $f(u) \neq f(v)$. Somit ist f eine 2-Färbung von G.

Übung und Nachbereitung

Bestimmen Sie für jede der folgenden Aussagen, ob sie wahr oder falsch ist. Falls die Aussage wahr ist, geben Sie eine kurze Begründung an. Andernfalls widerlegen Sie die Aussage, gegebenenfalls mit einem geeigneten Gegenbeispiel und Begründung, dass das Gegenbeispiel korrekt ist.

Übungsaufgabe Ü13.4. (*Quiz*)

(a) **Aussage:** Für alle Sprachen A, B, C gilt $(A \setminus C)B = AB \setminus CB$.

(b) **Aussage:** Wenn $L \subseteq \Sigma^*$ regulär ist und $a \in \Sigma$, dann ist die Sprache aller Wörter aus L, die nicht mit a enden, regulär.

(c) Definition: Analog zu rechtslinearen Grammatiken ist eine Grammatik G *linkslinear*, wenn jede Produktion von G die Gestalt $X \to Ya$ oder $X \to a$ hat.

Aussage: Die Sprache $L(G)$ einer linkslinearen Grammatik G ist regulär.

Hinweis: Sie dürfen verwenden, dass reguläre Sprachen unter Spiegelung abgeschlossen sind, d.h. für jede reguläre Sprache L ist auch L^R regulär.

(d) **Aussage:** Wenn die Sprache einer kontextfreien Grammatik regulär ist, dann ist die Grammatik nicht mehrdeutig.

(e) **Aussage:** Wenn $A, B \subseteq \Sigma^*$ kontextfrei sind und $\varepsilon \notin A$, dann sind alle Lösungen der Gleichung $X = AX \cup B$ kontextfrei.

Sie dürfen das folgende Theorem für diese Aufgabe verwenden: Für alle Sprachen $X, A, B \subseteq \Sigma^*$ mit $\varepsilon \notin A$ gilt: $X = AX \cup B \Longrightarrow X = A^*B$.

Lösungsskizze.

(a) Die Aussage ist falsch. Sei $A = \{aa\}$, $B = \{\varepsilon, a\}$ und $C = \{aaa\}$. Dann gilt $(A \setminus C)B = \{aa, aaa\}$ und $AB \setminus CB = \{aa\}$.

(b) Die Aussage ist wahr. Die Sprache $\Sigma^*\{a\}$ ist regulär, da Σ^* und $\{a\}$ regulär sind und reguläre Sprachen unter Konkatenation abgeschlossen sind. Durch die Abschlusseigenschaften der regulären Sprachen ist damit auch $L \setminus \Sigma^*\{a\}$ regulär, was die gesuchte Sprache ist.

<u>Alternative:</u> Sei $M = (Q, \Sigma, \delta, q_0, F)$ ein DFA für L. Sei $q_f \notin F$. Der NFA $M' := (Q \cup \{q_f\}, \Sigma, \delta', q_0, \{q_f\})$ mit $\delta' := \delta \cup \{(q, x, q_f) \mid \delta(q, x) \in F \wedge x \in (\Sigma \setminus \{a\})\}$ akzeptiert die gewünschte Sprache.

(c) Die Aussage ist wahr. Wir spiegeln zunächst die Produktionen, d.h. wir wandeln jede Produktion der Form $X \to Ya$ in $X \to aY$ um, erhalten wir eine rechtslineare Grammatik G'. Wir haben also $L(G')^R = L(G)$. Da rechtslineare Grammatiken reguläre Sprachen erzeugen und reguläre Sprachen unter Spiegelung abgeschlossen sind, muss auch $L(G)$ regulär sein.

(d) Die Aussage ist falsch. Sei $G = (S, \{S \to SS \mid a\})$. Dann ist $L(G) = \{a^n \mid n \in \mathbb{N}_+\}$ regulär. Die Grammatik ist jedoch mehrdeutig, da man aaa sowohl mit $S \to SS \to aS \to aSS \to^2 aaa$ als auch mit $S \to SS \to Sa \to SSa \to^2 aaa$ herleiten kann.

(e) Die Aussage ist wahr. Laut dem Theorem ist die einzige Lösung der Gleichung $X = A^*B$. Da A und B CFLs sind und CFLs unter Konkatenation und Stern abgeschlossen sind, muss auch X eine CFL sein.

Übungsaufgabe Ü13.5. (*Kontextfreie Sprachen*)

Sei $\Sigma = \{a, b, c\}$.

(a) Geben Sie für die Sprache $L = \{(aba)^n b^m (ba)^n b\}$ eine kontextfreie Grammatik G an, sodass $L(G) = L$ gilt.

(b) Sei $G = (\{S\}, \Sigma, \{S \to \varepsilon \mid aSc \mid aSbSc\}, S)$ eine kontextfreie Grammatik. Zeigen Sie, dass für alle in G ableitbaren Wörter w die Eigenschaft $|w|_a = |w|_c$ gilt.

Lösungsskizze.

(a) $G = (\{S, X, B\}, \Sigma, P, S)$ mit P: $S \to Xb \qquad X \to abaXba \mid B \qquad B \to bB \mid \varepsilon$

(b) Sei $u \in L(G)$. Wir zeigen $|w|_a = |w|_c$ mit Induktion über die Erzeugung von u.

- Fall ε: $|\varepsilon|_a = 0 = |\varepsilon|_c$.
- Fall aSc: Betrachte $S \to aSc \to^* awc$ mit $|w|_a = |w|_c$ per IH. Es gilt $|awc|_a = 1 + |w|_a = 1 + |w|_c = |awc|_c$.
- Fall $aSbSc$: Betrachte $S \to aSbSc \to^* awbvc$ mit $|w|_a = |w|_c$ und $|v|_a = |v|_c$ per IH. Es gilt $|awbvc|_a = 1 + |w|_a + |v|_a = 1 + |w|_c + |v|_c = |awbvc|_c$.

Übungsaufgabe Ü13.6. (*Pumping Lemma*)

Sei $\Sigma = \{a, b, c\}$. Zeigen Sie, dass die Sprache $L = \{wc^n w^R \mid w \in \{a, b\}^n \wedge n \in \mathbb{N}_0\}$ nicht kontextfrei ist. Führen Sie einen Widerspruchsbeweis unter Verwendung des Pumping Lemmas für kontextfreie Sprachen.

Lösungsskizze.

- Angenommen, L wäre kontextfrei. Dann können wir das PL für kontextfreie Sprachen anwenden.
- Sei $n \in \mathbb{N}_+$ eine PL-Zahl für L.

- Wähle $z = a^n c^n a^n$. Es gilt $|z| = 3n \geq n$ und $z \in L$.
- Nach PL gibt es eine Zerlegung $z = uvwxy$ mit 1. $vx \neq \varepsilon$, 2. $|vwx| \leq n$ und 3. $\forall i \in \mathbb{N}_0 . uv^i wx^i y \in L$.
- Sei $z' = uv^2wx^2y$. Da wegen (2) nur maximal einer der beiden a-Blöcke gepumpt werden kann, gilt, dass $z' = a^n c z''$ oder $z' = z'' c a^n$. Da mindestens einer der beiden a-Blöcke noch Länge n hat, muss $|z'| = 3n$ gelten. Es ist aber wegen (1) auch $|z'| > 3n$. Dies ist ein Widerspruch zu (3).

  Alternativ: Fallunterscheidung über $|vx|_a > 0$ oder $|vx|_c > 0$.
- Also haben wir einen Widerspruch zur Annahme hergeleitet und somit ist L nicht kontextfrei.

Übungsaufgabe Ü13.7. (*Entscheidbarkeit*)

Betrachten Sie die Menge $A = \{w \in \{0,1\}^* \mid |L(M_w)| \leq 42\}$.

(a) Beweisen Sie, dass die Menge nicht semi-entscheidbar ist, indem Sie $\overline{H_0} \leq A$ zeigen.

(b) Geben Sie einen Semientscheidungsalgorithmus für $\overline{A}$ an.

Lösungsskizze.

(a) A ist die Menge der Wörter, deren korrespondierenden Turingmaschine auf maximal 42 Eingaben hält. Wir zeigen $\overline{H_0} \leq A$ mittels folgender Reduktion:

Reduktionsfunktion: Gegeben ein Wort w, gib die Kodierung w' einer TM zurück, die zuerst das Band löscht und dann M_w simuliert. Sobald M_w hält, geht $M_{w'}$ in einen Endzustand und hält.

Totalität: Die Reduktion gibt für jede Eingabe w ein entsprechendes w' zurück.

Berechenbarkeit: Die Reduktion ist berechenbar, da nach VL Löschen des Bandes und Simulation einer TM berechenbar sind.

Korrektheit:

- $w \in \overline{H_0}$: M_w hält nicht auf dem leeren Band, folglich hält $M_{w'}$ auf keinem Input, da es immer das Band löscht und M_w auf dem leeren Band simuliert. Damit ist $|L(M_{w'})| = 0 \leq 42$ und somit $w' \in A$.
- $w \notin \overline{H_0}$: Da M_w auf dem leeren Band hält, akzeptiert $M_{w'}$ jeden Input, da es immer M_w auf dem leeren Band simuliert und dann in einen Endzustand übergeht. Damit ist $|L(M_{w'})| = \infty > 42$ und $w' \notin A$.

(b) $\overline{A}$ ist die Menge der Wörter, die eine TM kodieren, die auf mindestens 43 Eingaben hält. Semientscheidungsalgorithmus: Simuliere M_w mit Dove-Tailing auf allen Inputs. (Also: Simuliere für alle Eingaben der Länge $\leq$1 für 1 Schritt, dann für alle Eingaben der Länge $\leq$2 für 2 Schritte, usw.) Sobald 43 der Simulationen halten, antworte positiv.

Übungsaufgabe Ü13.8. (*REJECT*)

Mit REJECT bezeichnen wir folgendes Problem:

- **Input:** Ein ϵ-NFA N mit n Zuständen über dem Alphabet $\Sigma = \{0,1\}$ und eine Zahl $m \leq n$.

- **Frage:** Gibt es ein Wort der Länge m das von N nicht akzeptiert wird?

In Mengenschreibweise:

$$\mathsf{REJECT} := \{(N, m) \mid N = (Q, \Sigma, \delta, q_0, F) \text{ ist } \epsilon\text{-NFA und } m \leq |Q| \text{ und } \Sigma^m \setminus L(N) \neq \emptyset\}$$

(a) Zeigen Sie, dass REJECT NP-vollständig ist. Verwenden Sie hierbei eine geeignete Reduktion von 3-KNF-SAT auf REJECT.

(b) Wenden Sie Ihre Reduktion auf $(x_1 \vee \neg x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee x_3)$ an. Geben Sie den konstruierten ϵ-NFA graphisch an.

Hinweis: Betrachten Sie die Menge aller nichterfüllenden Belegungen für die Formel F, die genau m Variablen hat, als Wörter der Länge m.

Lösungsskizze.

(a)
- REJECT $\in NP$:
  Ein Zertifikat ist ein Wort $w \in \Sigma^m \setminus L(N)$. Da $m \leq |Q|$, ist $|w| = m$ polynomiell bezüglich $|N|$. Das Wortproblem (und damit $w \notin L(N)$) kann nun in polynomieller Zeit gelöst werden (vgl. Vorlesung).
- Idee für Reduktion: Der ϵ-NFA N soll alle nicht erfüllenden Belegungen von F akzeptieren und alle erfüllenden Belegungen ablehnen. Hierfür rät der NFA zunächst die Klausel, welche nicht erfüllt ist und dann die unerfüllende Belegung.
- Reduktion: Sei $F = \bigwedge_{i=1}^{k} C_i$ eine Formel in 3-KNF mit m Variablen $(x_1, x_2, \ldots, x_m)$ und mit $C_i = \bigvee_{j=1}^{l_i} L_{i,j}$ für jedes i. Der ϵ-NFA $N := \Big(Q, \Sigma, \delta, q_0, Q_F\Big)$ wird wie folgt definiert, wobei $[i, j] := \{i, i+1, \ldots, j-1, j\}$:

$$\begin{aligned} Q &:= \{q_0\} \cup \{q_{i,j} \mid i \in [1,k] \wedge j \in [0,m]\} \\ Q_F &:= \{q_{i,m} \mid i \in [1,k]\} \\ \delta &:= \{(q_0, \epsilon, q_{i,0}) \mid i \in [1,k]\} \\ &\quad \cup \{(q_{i,j-1}, 0, q_{i,j}) \mid j \in [1,m] \wedge \neg x_j \notin \{L_{i,1}, \ldots, L_{i,l_i}\}\} \\ &\quad \cup \{(q_{i,j-1}, 1, q_{i,j}) \mid j \in [1,m] \wedge x_j \notin \{L_{i,1}, \ldots, L_{i,l_i}\}\} \end{aligned}$$

  Die Reduktion bildet dann F auf (N, m) ab.
- Polynomielle Zeit:
  Der konstruierte ϵ-NFA hat $\mathcal{O}(km)$ Zustände und Transition und ein konstantes Alphabet. Somit ist seine Größe polynomiell in F. Die Komponenten der Übergangsfunktion können während des Lesens einer Klausel C direkt konstruiert werden. Somit bedarf die Reduktion polynomielle Zeit.
- Korrektheit:
  Wir zeigen nun, dass die von N akzeptierten Wörter der Länge m genau die unerfüllenden Belegungen von F sind. Dann gilt:

$$\begin{aligned} F \in \mathsf{3\text{-}KNF\text{-}SAT} &\iff F \text{ ist erfüllbar} \\ &\iff \text{nicht alle Belegungen für } F \text{ sind unerfüllend} \\ &\iff N \text{ lehnt ein Wort der Länge } m \text{ ab} \\ &\iff (N, m) \in \mathsf{REJECT} \end{aligned}$$

Sei $\sigma(F) = 0$ eine nicht erfüllende Belegung. Dann existiert insbesondere eine Klausel C_i mit $\sigma(C_i) = \sigma(L_{i,1}) + \sigma(L_{i,2}) + \sigma(L_{i,3}) = 0$. Das Wort $w = \sigma(x_1)\sigma(x_2)...\sigma(x_m)$ wird dann von $q_{i,0}$ aus akzeptiert und es gilt somit $w \in L(N)$.

Besteht andererseits ein Lauf in N über q_{i_0}, so erfüllt die Belegung, beschrieben durch den Lauf, kein Literal der Klausel C_i. Damit ist die Belegung unerfüllend für F.

(b) $(N, 4)$ mit N: