Theoretische Informatik 1
Übungsblatt 4

René Maseli
Prof. Dr. Roland Meyer

TU Braunschweig
Wintersemester 2022/23

Ausgabe: 13.12.2022 Abgabe: 23.12.2022, 23:59

Geben Sie Ihre Lösungen bis Freitag, 23.12 23:59 Uhr, im Vips-Verzeichnis der StudIP-Veranstaltung ab. Fertigen Sie dazu ihre Hausaufgaben direkt in .pdf Form an oder scannen ihre handschriftlichen Hausaufgaben ein. Geben Sie in Gruppen von **4 Personen** ab.

Definition: Finite-state Transducer

Ein endlicher Transduktor über einem endlichen Eingabe-Alphabet Σ und einem endlichen Ausgabe-Alphabet Γ ist formal ein 4-Tupel $T = \langle Q, q_0, \rightarrow, Q_F \rangle$ bestehend aus

1. einer endlicher Menge an Zuständen Q,
2. einen initialen Zustand $q_0 \in Q$
3. eine Transitionsrelation $\rightarrow \subseteq Q \times (\Sigma \cup \{\tau\}) \times (\Gamma \cup \{\tau\}) \times Q$,
4. und eine Menge akzeptierender Zustände $Q_F \subseteq Q$

Man kann sich einen Transduktor als einen NFA mit spontanen Transitionen vorstellen, welcher nicht nur Eingabewörter akzeptiert, sondern dabei auch Wörter in einer Ausgabe hinterlässt. Es übersetzt die Eingaben aus Σ^* in Ausgaben aus Γ^*. Das macht Transduktoren zu geeigneten Werkzeugen in der Linguistik und im Verarbeiten natürlicher Sprachen.

Im Folgenden werden weitere Notationen und wichtige Definitionen aufgelistet:

1. $\langle p, a, x, q \rangle \in \rightarrow$ wird durch $p \xrightarrow{a/x} q$ ausgedrückt. Wenn im Zustant p ein a gelesen wird, gibt der Transduktor ein x aus und wechselt in Zustand q. Intuitiv werden den Transitionen mit $a = \tau$ gefolgt, ohne ein Eingabesymbol zu konsumieren, und die Transitionen mit $x = \tau$ erzeugen keine Ausgabe.
2. $\rightarrow^* \subseteq Q \times (\Sigma \cup \{\tau\})^* \times (\Gamma \cup \{\tau\})^* \times Q$ bezeichnet die reflexive transitive Hülle von $\rightarrow$. Es erfüllt insbesondere $q \xrightarrow{\varepsilon/\varepsilon}{}^* q$ und $q \xrightarrow{w/o}{}^* q_n \iff \exists q_1, \ldots, q_{n-1} : q \xrightarrow{w_1/o_1} q_1 \xrightarrow{w_2/o_2} \cdots \xrightarrow{w_{n-1}/o_{n-1}} q_{n-1} \xrightarrow{w_n/o_n} q_n$.
3. T induziert eine Relation $[\![T]\!] \subseteq \Sigma^* \times \Gamma^*$ wiefolgt:

   $$w \,[\![T]\!]\, o \iff \exists w' \in (\Sigma \cup \{\tau\})^*, o' \in (\Gamma \cup \{\tau\})^* : q_0 \xrightarrow{w'/o'}{}^* q_f \in Q_F \quad \text{and} \quad \pi_\Sigma(w') = w, \pi_\Gamma(o') = o$$

   Hereby, $\pi_\Sigma \colon (\Sigma \cup \{\tau\}) \rightarrow \Sigma^*$ with $\pi_\Sigma(\tau) = \varepsilon$ and $\forall a \in \Sigma \colon \pi_\Sigma(a) = a$ induces a homomorphism, which deletes τ from a word. τ denotes either spontaneous transitions or empty output and hence should not be visible.

We say that $o \in \Gamma^*$ is an output of T on $w \in \Sigma^*$.

4. T does not only transduce single words, but whole languages. We define for any language $L \subseteq \Sigma^*$ the translation under T as $T(L) = \{o \in \Gamma^* \mid \exists w \in L : w \llbracket T \rrbracket\ o\} \subseteq \Gamma^*$.

Aufgabe 1: Endliche Transduktoren [20 Punkte]

a) [3 Punkte] Konstruieren Sie einen Transduktor T, der in einem Wort $w \in \{a, b, c\}^*$ ein b vor jedes Vorkommen von a vorhängt und jedes zweite c löscht. Geben Sie einen regulären Ausdruck für $T((ac)^*)$ an.
Ein Korrektheitsbeweis ist nicht nötig.

b) [3 Punkte] Wir nennen einen Transduktor deterministisch, wenn es in jedem Zustand zu jeder Eingabe **maximal eine** mögliche, und damit auch **eindeutige**, Transition gibt; diese darf auch spontan sein. Beispielsweise darf ein Zustand mit einer a-Transition keine weitere a-Transition oder spontane Transition besitzen. In beiden Fällen würde es 2 mögliche Transitionen geben, die der Transduktor beim Lesen eines a's nehmen könnte.
Zeigen Sie, dass es **nicht** möglich ist Transduktoren zu determinisieren, d.h. es gibt nicht zu jedem Transduktor T einen deterministischen Transduktor T^{det} mit $T(L) = T^{\text{det}}(L)$ für alle Sprachen $L \subseteq \Sigma^*$.

c) [3 Punkte] Sei $h : \Sigma^* \to \Gamma^*$ ein Homomorphismus zwischen Sprachen. Konstruieren Sie einen Transduktor T_h, sodass $T_h(L) = h(L)$ für jede Sprache $L \subseteq \Sigma^*$ gilt. Beweisen Sie die Korrektheit ihrer Konstruktion.

d) [3 Punkte] Zeigen Sie nun, dass es auch einen Transduktor $T_{h^{-1}}$ gibt, sodass $T_{h^{-1}}(L) = h^{-1}(L)$ für alle $L \subseteq \Gamma^*$ gilt. Beweisen Sie die Korrektheit ihrer Konstruktion.

e) [2 Punkte] Zeigen Sie, dass es zu jeder regulären Sprachen M einen Transduktor T_M gibt mit $T_M(L) = L \cap M$.
Bemerkung: Hiermit haben Sie gezeigt, dass Transduktoren in der Lage sind, viele übliche Operationen auf Sprachen darzustellen. Ist eine Sprachklasse abgeschlossen unter Übersetzungen von Transduktoren, dann folgt nun direkt, dass diese auch unter den oben genannten Operationen abgeschlossen ist.

f) [6 Punkte] Zeigen Sie nun, dass dies auch in umgekehrter Richtung gilt. Das heißt, eine Sprachklasse, die unter den oben aufgeführten Operationen abgeschlossen ist, ist auch unter Übersetzungen von Transduktoren abgeschlossen.
Bemerkung: Sie müssen also zeigen, dass Sie die Übersetzung einer beliebigen Sprache L unter einem Transduktor T mit Hilfe von den drei oben genannten Operation darstellen können.

Aufgabe 2: Einzigartige Minimale DFAs [9 Punkte]

Sei $L \subseteq \Sigma^*$ eine reguläre Sprache und $A = \langle Q_A, i_A, \rightarrow_A, F_A \rangle$ ihr Äquivalenzklassen-DFA mit $Q_A = \Sigma^*|_{\equiv_L}$, $i_A = [\varepsilon]_{\equiv_L}$, $[v]_{\equiv_L} \xrightarrow{s}_A [w]_{\equiv_L} \iff v.s = w$ und $F_A = \{[w]_{\equiv_L} \mid w \in L\}$. Dabei ist bereits bekannt, dass $\mathcal{L}(A) = L$ gilt und dass A minimal für L ist. Sei $B = \langle Q_B, i_B, \rightarrow_B, F_B \rangle$ ein weiterer DFA mit $\mathcal{L}(B) = L$ und $|Q_B| = [\equiv_L]$.

a) [2 Punkte] Zeigen Sie, dass alle Zustände in B erreichbar sind, also dass für jeden Zustand $q \in Q_B$ ein Wort $w \in \Sigma^*$ existiert, sodass $i_B \xrightarrow{w}{}^*_B q$ gilt.

Sei $f : Q_B \rightarrow Q_A$ induktiv definiert mit $f(i_B) = [\varepsilon]_{\equiv_L}$ und $\forall p \xrightarrow{s}_B q \colon f(q) = [w]_{\equiv_L}.s$.

Es ist zu zeigen, dass B zu A isomorph ist ($B \sim A$).

b) [3 Punkte] Zeigen Sie induktiv für alle $w \in \Sigma^*$, dass $f(q_w) = [w]_{\equiv_L}$ gilt, wobei $q_w \in Q_B$ der einzigartige Zustand mit $i_B \xrightarrow{w}{}^*_B q_w$.

c) [2 Punkte] Zeigen Sie, dass f bijektiv ist.

d) [2 Punkte] Sei $q \in Q_B$ ein Zustand von B. Zeigen Sie, dass $q \in F_B \iff f(q) \in F_A$ gilt.

Aufgabe 3: Kosten der Determinisierung [6 Punkte]

In dieser Aufgabe wollen wir zeigen, dass sich manche Sprachen mit einem kleinen NFA beschreiben lassen, jeder DFA dafür jedoch zwangsweise riesig ist.

Für eine Zahl $k \in \mathbb{N}, k > 0$ sei $L_{a@k} = \Sigma^*.a.\Sigma^{k-1}$ die Sprache der Wörter über $\Sigma = \{a, b\}$, deren k-ter Buchstabe von rechts ein a ist.

a) [1 Punkt] Zeigen Sie, wie man zu jedem $k \in \mathbb{N}, k > 0$ einen NFA $A_k = \langle Q_k, q_0, \rightarrow, F_k \rangle$ mit $\mathcal{L}(A_k) = L_{a@k}$ und $|Q_k| = k + 1$ konstruiert. Geben Sie diesen Automaten formal als Tupel an. Sie müssen die Sprachgleichheit nicht beweisen.

b) [2 Punkte] Zeichnen Sie A_3 und bestimmen Sie seine Determinisierung A_3^{det} mittels der Rabin-Scott-Potenzmengenkonstruktion.
Vergleichen Sie die Zustandsanzahl von A_3 und A_3^{det}.

c) [3 Punkte] Sei $k \in \mathbb{N}, k > 0$ beliebig. Beweisen Sie, dass es für $L_{a@k}$ keinen DFA B mit weniger als 2^k Zuständen gibt, so dass $\mathcal{L}(B) = L_{a@k}$ gilt.
Hinweis: Gehen Sie wie folgt vor:

1. Angenommen es gäbe $B = (Q', q_0', \rightarrow', Q_F')$ mit $\mathcal{L}(B) = L_{a@k}$ und $|Q'| < 2^k$.
2. Betrachten Sie die Menge Σ^k der Wörter der Länge k. Wie viele solcher Wörter gibt es?
3. Betrachten Sie zu jedem Wort $w \in \Sigma^k$ den (eindeutigen) Zustand q_w, in dem DFA B ist, nachdem er w gelesen hat.
4. Leiten Sie nun einen Widerspruch her.