Formale Systeme: Besprechung Musterklausur

Stephan Mennicke

Wissensbasierte Systeme

01. Februar 2024

a) Geben Sie die formale Definition eines *nichtdeterministischen Kellerautomaten* an. Vervollständigen Sie hierfür den nachfolgenden Text:

```
Ein nichtdeterministischer Kellerautomat (PDA) ist ein Sechs-Tupel \mathcal{M}=\langle\ ,\ ,\ ,\ ,\ \rangle mit den folgenden Bestandteilen: . . .
```

- b) Welcher andere Akzeptanzbegriff für Kellerautomaten ist laut Anmerkung in der Vorlesung auch möglich?
- c) Benennen Sie formal die Unterschiede zwischen deterministischen und nicht-deterministischen Kellerautomaten.
- d) Welcher Typ formaler Sprachen wird durch deterministische und welcher durch nichtdeterministische Kellerautomaten charakterisiert? Benennen Sie jeweils eine Sprache genau diesen Typs.

a) Geben Sie die formale Definition eines *nichtdeterministischen Kellerautomaten* an. Vervollständigen Sie hierfür den nachfolgenden Text:

Ein nichtdeterministischer Kellerautomat (PDA) ist ein Sechs-Tupel $\mathcal{M}=\langle \ , \ , \ , \ , \ \rangle$ mit den folgenden Bestandteilen: . . . $\mathcal{M}=\langle Q, \Sigma, \Gamma, \delta, Q_0, F \rangle$ mit den folgenden Bestandteilen:

Q: endliche Menge von Zuständen

 Σ : Eingabealphabet

Γ: Kelleralphabet

 δ : Übergangsfunktion, eine totale Funktion: $Q \times \Sigma_{\epsilon} \times \Gamma_{\epsilon} \to 2^{Q \times \Gamma_{\epsilon}}$,

 Q_0 : Menge möglicher Startzustände $Q_0 \subseteq Q$

F: Menge von Endzuständen $F \subseteq Q$

b) Welcher andere Akzeptanzbegriff für Kellerautomaten ist laut Anmerkung in der Vorlesung auch möglich?

Neben der Akzeptanz über Endzustände in $F \subseteq Q$ gibt es die (äquivalente) Akzeptanz über leeren Keller.

c) Benennen Sie formal die Unterschiede zwischen deterministischen und nicht-deterministischen Kellerautomaten.

Ein deterministischer Kellerautomat (DPDA) ist ein Sechs-Tupel $\mathcal{M} = \langle Q, \Sigma, \Gamma, \delta, q_0, F \rangle$ mit den folgenden Bestandteilen:

δ: Übergangsfunktion, eine partielle Funktion $Q \times \Sigma_{\varepsilon} \times \Gamma_{\varepsilon} \to Q \times \Gamma_{\varepsilon}$, so dass für alle $q \in Q$, $a \in \Sigma$ und $A \in \Gamma$ jeweils nur eines der folgenden definiert ist:

$$\delta(q, a, A)$$
 $\delta(q, a, \varepsilon)$ $\delta(q, \varepsilon, A)$ $\delta(q, \varepsilon, \varepsilon)$

 q_0 : ein Startstand $q_0 \in Q$

d) Welcher Typ formaler Sprachen wird durch deterministische und welcher durch nichtdeterministische Kellerautomaten charakterisiert? Benennen Sie jeweils eine Sprache genau diesen Typs.

Automat	Sprachklasse (bzw. Typ)		Beispielsprache
PDA	kontextfrei	Typ 2	$L = \{a^i b^j c^k \mid i \neq j \text{ oder } j \neq k\}$
DPDA	det. kontextfrei	det. Typ 2	

Hierbei bezeichnet $|w|_x$ die Anzahl der $x \in \Sigma$ in $w \in \Sigma^*$.

6

M2 - Pumping-Lemma

a) Formulieren Sie formal präzise das *Pumping-Lemma für kontextfreie Sprachen*. Vervollständigen Sie hierfür den nachfolgenden Text:

Für jede kontextfreie Sprache L gibt es eine natürliche Zahl $n \ge 0$, so dass gilt: . . .

b) Zeigen Sie mithilfe des *Pumping-Lemmas für reguläre Sprachen*, dass die Sprache $L = \{0^p \mid p \text{ ist eine Primzahl}\}$ nicht regulär ist.

M2 - Pumping-Lemma

a) Formulieren Sie formal präzise das *Pumping-Lemma für kontextfreie Sprachen*. Vervollständigen Sie hierfür den nachfolgenden Text:

Für jede kontextfreie Sprache L gibt es eine natürliche Zahl $n \ge 0$, so dass gilt: für jedes Wort $z \in L$ mit $|z| \ge n$ gibt es eine Zerlegung z = uvwxy mit $|vx| \ge 1$ und $|vwx| \le n$, so dass für jede Zahl $k \ge 0$ gilt: $uv^k wx^k y \in L$.

b) Zeigen Sie mithilfe des *Pumping-Lemmas für reguläre Sprachen*, dass die Sprache $L = \{0^p \mid p \text{ ist eine Primzahl}\}$ nicht regulär ist.

Annahme: L erfüllt das Pumping-Lemma. D.h. es gibt $n \ge 0$, so dass für jedes $z \in L$ mit $|z| \ge n$ eine Aufteilung z = uvw existiert mit:

(1)
$$|v| \geqslant 1$$
 (2) $|uv| \leqslant n$ (3) $uv^k w \in L$ für jedes $k \geqslant 0$.

Wähle eine Primzahl $\ell > n+2$. Laut Pump-Eigenschaft finden wir eine Zerlegung von $0^\ell = uvw$, für die insbesondere gilt: $uv^k w \in L$ für k = |uw|. Aber $uv^{|uw|} w = 0^{(|v|+1)|uw|} \notin L$. Widerspruch. L ist daher nicht regulär.

9

M3 - Grammatiken

Gegeben sei die Grammatik $G = (V, \Sigma, P, S)$ mit

$$V = \{S, A, B\}, \ \Sigma = \{a, b\} \text{ und}$$

 $P = \{S \rightarrow ASB, \ S \rightarrow AB, \ AB \rightarrow BA, \ A \rightarrow a, \ B \rightarrow b\}.$

- a) Von welchem maximalen Typ ist G? Begründen Sie Ihre Antwort.
- b) Geben Sie vier Wörter $w_1, w_2, w_3, w_4 \in L(G)$ mit $|w_1| = |w_2| = |w_3| = |w_4| = 4$ an.
- c) Beschreiben Sie die durch G erzeugte Sprache L(G) in einer geeigneten Notation.
- a) G ist vom Typ 1 (kontextsensitiv) und nicht vom Typ 2 (kontextfrei).
- b) $w_1 = aabb$, $w_2 = abab$, $w_3 = abba$, $w_4 = baba$
- c) $L(G) = \{w \in \{a, b\}^+ \mid |w|_a = |w|_b\}$

M4 - Grammatiken: CNF/CYK

Gegeben sei das Wort w = abac und die Grammatik $G = (V, \Sigma, P, S)$ mit

$$V = \{S, A, B, C, D\}, \Sigma = \{a, b, c\}$$
und

$$P = \{S \rightarrow AB, \ A \rightarrow BA, \ A \rightarrow a, B \rightarrow AC, \ B \rightarrow BB, \ B \rightarrow b, C \rightarrow c, \ D \rightarrow AB\}.$$

- a) Ist die Grammatik G in Chomsky-Normalform? Begründen Sie Ihre Antwort.
- b) Entscheiden Sie mithilfe des Cocke-Younger-Kasami-Algorithmus, ob $w \in L(G)$ gilt. Transformieren Sie, falls notwendig, G in Chomsky-Normalform.
- c) Entfernen Sie in *G*, sofern vorhanden, nichtterminierende und nichterreichbare Symbole. Begründen Sie Ihr Vorgehen.

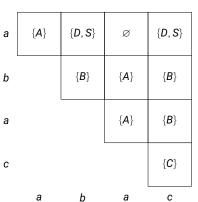
- a) ja. Jede Regel ist von der Form $A \to BC$ oder $A \to x$ ($A, B, C \in V, x \in \Sigma$).
- b) ...
- c) *D* ist nicht erreichbar, $G' = (V', \Sigma, P', S)$ mit $V' = \{S, A, B, C\}$ und $P' = \{S \rightarrow AB, A \rightarrow BA, A \rightarrow a, B \rightarrow AC, B \rightarrow BB, B \rightarrow b, C \rightarrow c\}$ hat weder nichtterminierende noch nichterreichbare Symbole.

M4 - Grammatiken: CNF/CYK

Gegeben sei das Wort w = abac und die Grammatik $G = (V, \Sigma, P, S)$ mit

$$V = \{S, A, B, C, D\}, \ \Sigma = \{a, b, c\} \text{ und }$$

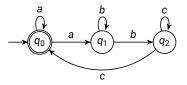
$$P = \{S \rightarrow AB, A \rightarrow BA, A \rightarrow a, B \rightarrow AC, B \rightarrow BB, B \rightarrow b, C \rightarrow c, D \rightarrow AB\}.$$



Damit ist $w = abac \in L$.

M5 - NFA/Reguläre Ausdrücke/DFA

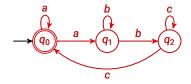
Gegeben sei der NFA $\mathcal{M} = (\{q_0, q_1, q_2\}, \{a, b, c\}, \delta, \{q_0\}, \{q_0\})$ mit δ :



- a) Berechnen Sie mithilfe des Arden-Lemmas einen regulären Ausdruck α mit $L(\alpha) = L(\mathcal{M})$.
- b) Konstruieren Sie einen zu ${\mathfrak M}$ äquivalenten DFA ${\mathfrak M}'.$ Verwenden Sie dazu die Potenzmengenkonstruktion aus der Vorlesung. Stellen Sie dabei sicher, dass der konstruierte Automat keine unerreichbaren Zustände enthält.

M5 - NFA/Reguläre Ausdrücke/DFA

Gegeben sei der NFA $\mathcal{M} = (\{q_0, q_1, q_2\}, \{a, b, c\}, \delta, \{q_0\}, \{q_0\})$ mit δ :

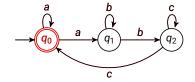


a) Berechnen Sie mithilfe des Arden-Lemmas einen regulären Ausdruck α mit $L(\alpha) = L(\mathfrak{M}).$

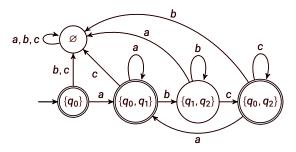
$$\begin{array}{lllll} \alpha_0 & \equiv & a\alpha_0 \mid a\alpha_1 \mid \epsilon & \equiv & a\alpha_0 \mid ab^+c^+\alpha_0 \mid \epsilon & \equiv & (a \mid ab^+c^+)^* \\ \alpha_1 & \equiv & b\alpha_1 \mid b\alpha_2 & \equiv & b^*b\alpha_2 \equiv b^+\alpha_2 \\ \alpha_2 & \equiv & c\alpha_2 \mid c\alpha_0 & \equiv & c^*c\alpha_0 \equiv c^+\alpha_0 \end{array}$$

M5 - NFA/Reguläre Ausdrücke/DFA

Gegeben sei der NFA $\mathcal{M} = (\{q_0, q_1, q_2\}, \{a, b, c\}, \delta, \{q_0\}, \{q_0\})$ mit δ :



b) DFA \mathfrak{M}' :



M6 - Nerode/Minimalautomat

Geben Sie die Nerode-Äquivalenzklassen für die nachfolgenden Sprachen an und geben Sie den Minimalautomaten für L_1 an.

$$L_1 = L((ab)^*a^* \mid b)$$
 und $L_2 = \{w \in \{a, b\}^* \mid w = w^R\}$

Für L₁ lauten die Äquivalenzklassen:

$$[\varepsilon]_{L_1} = L(\varepsilon) \qquad [a]_{L_1} = L((ab)^*a)$$

$$[b]_{L_1} = L(b) \qquad [aa]_{L_1} = L((ab)^*aaa^*)$$

$$[ab]_{L_1} = L(ab(ab)^*) \qquad [aa]_{L_1} = L((ab)^*aaa^*)$$
 Minimalautomat:

Für L_2 lauten die Äquivalenzklassen: $[w]_{L_2} = \{w\}$ mit $w \in \{a, b\}^*$.

D.h. jedes Wort aus $\{a, b\}^*$ bildet eine eigene Äquivalenzklasse.

Damit ist der Nerode-Index von L_2 unendlich, d.h. die Sprache L_2 ist nicht regulär.

M6 - Nerode/Minimalautomat

Beweis:

Angenommen, es gibt zwei Wörter $v, w \in [v]$ mit $v \neq w$ und $|v| \geqslant |w|$.

Falls |v| = |w|, so ist $vv^R \in L_2$, aber $wv^R \not\in L_2$, also $v \not\sim_{L_2} w$.

Ansonsten gilt |v| > |w|.

Falls $wv^R \not\in L_2$ ist, so gilt $v \not\sim_{L_2} w$, denn $vv^R \in L_2$.

Ist auch $wv^R \in L_2$, betrachte die Zerlegung $wv^R = wxyz$, wobei |wx| = |z| und $1 \le |y| \le 2$.

Insbesondere ist $(wxy)^R = y^R(wx)^R = yz$, also $y^R = y$.

Betrachte $wx\bar{y}z$, wobei \bar{y} aus y entsteht, indem jedes a durch b (und umgekehrt) ersetzt wird.

Dann ist auch $wx\bar{y}z \in L_2$.

Aber $v^R = xyz \neq x\bar{y}z$, also ist $vx\bar{y}z \notin L_2$.

Damit $v \not\sim_{L_2} w$.

Seien Σ ein Alphabet, $w \in \Sigma^*$ und $a \in \Sigma$. Eine *Injektion von a in w* ist ein Wort, das aus w entsteht, wenn ein a an eine beliebige Stelle in w eingefügt wird. Wir definieren die *Injektionen von a in w* als

$$w|_{a} := \{uav \mid \exists u, v \in \Sigma^* : w = uv\}.$$

Der Operator $||_a$ ($a \in \Sigma$) ist auf Wörtern definiert und wird auf natürliche Weise auf Sprachen $L \subseteq \Sigma^*$ erweitert durch:

$$L\|_a := \bigcup_{w \in L} w\|_a.$$

- a) Seien $\Sigma = \{a, b, c\}$ und $L = \{a^m b^n c^m \mid m > 0, n \geqslant 0\}$. Bestimmen Sie drei verschiedene Wörter $w_1, w_2, w_3 \in (L||_b) \setminus L$.
- b) Die regulären Sprachen sind unter $\|a\|$ (für alle $a \in \Sigma$) abgeschlossen. Begründen Sie die Korrektheit dieser Aussage.

Seien Σ ein Alphabet, $w \in \Sigma^*$ und $a \in \Sigma$. Eine *Injektion von a in w* ist ein Wort, das aus w entsteht, wenn ein a an eine beliebige Stelle in w eingefügt wird. Wir definieren die *Injektionen von a in w* als

$$w|_{a}:=\{uav\mid \exists u,v\in \Sigma^{*}: w=uv\}.$$

Der Operator $\|a\|$ ($a \in \Sigma$) ist auf Wörtern definiert und wird auf natürliche Weise auf Sprachen $L \subseteq \Sigma^*$ erweitert durch:

$$L\|_a := \bigcup_{w \in L} w\|_a.$$

a) Seien $\Sigma = \{a, b, c\}$ und $L = \{a^m b^n c^m \mid m > 0, n \ge 0\}$. Bestimmen Sie drei verschiedene Wörter $w_1, w_2, w_3 \in (L||_b) \setminus L$.

$$w_1 = bac$$
, $w_2 = acb$, $w_3 = aabac$

Nicht korrekt wären: $w_4 = abc$ or $w_5 = abbc$, obwohl $w_4 \in ac|_b$ und $w_5 \in abc|_b$.

Seien Σ ein Alphabet, $w \in \Sigma^*$ und $a \in \Sigma$. Eine *Injektion von a in w* ist ein Wort, das aus w entsteht, wenn ein a an eine beliebige Stelle in w eingefügt wird. Wir definieren die *Injektionen von a in w* als

$$w|_{a} := \{uav \mid \exists u, v \in \Sigma^* : w = uv\}.$$

Der Operator $\|a\|$ ($a \in \Sigma$) ist auf Wörtern definiert und wird auf natürliche Weise auf Sprachen $L \subseteq \Sigma^*$ erweitert durch:

$$L\|_a := \bigcup_{w \in L} w\|_a.$$

b) Die regulären Sprachen sind unter $\|a\|$ (für alle $a \in \Sigma$) abgeschlossen. Begründen Sie die Korrektheit dieser Aussage.

Wir zeigen für eine (beliebige) reguläre Sprache L und ein Symbol $a \in \Sigma$, dass $L||_a$ ebenfalls regulär ist.

Wir zeigen für eine (beliebige) reguläre Sprache $L \subseteq \Sigma^*$ und ein Symbol $a \in \Sigma$, dass $L||_a$ ebenfalls regulär ist.

Vorüberlegungen:

Wir haben *L*, eine reguläre Sprache, gegeben.

Auf *L* direkt argumentieren?

Können L auch als

- (i) reguläre Grammatik,
- (ii) DFA/NFA und
- (iii) regulären Ausdruck annehmen.

Es genügt, für eine gewählte Repräsentation zu zeigen, dass diese so umgeformt werden kann, dass

- (1) die Sprache $L||_a$ beschrieben wird und
- (2) das Resultat der Umformung die Repräsentation nicht verlässt.

Lassen Sie uns programmieren, also mit Automaten arbeiten.

Wir zeigen für eine (beliebige) reguläre Sprache $L \subseteq \Sigma^*$ und ein Symbol $a \in \Sigma$, dass $L||_a$ ebenfalls regulär ist.

Beweis:

Da L regulär ist, existiert ein DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ mit $\mathbf{L}(\mathcal{M}) = L$.

Wir konstruieren einen NFA \mathfrak{M}' ausgehend von \mathfrak{M} wie folgt:

- (i) kopiere alle Zustände und Übergänge, sodass \mathfrak{M}' aus den Zuständen und Übergängen von \mathfrak{M} besteht und einer Kopie von \mathfrak{M} .
- (ii) Startzustand bleibt q_0 . Endzustandsmenge wird die Kopie von F.
- (iii) Führe zwischem jedem (Original-)Zustand $q \in Q$ und seiner Kopie \widehat{q} einen a-Übergang ein.

$$L(\mathfrak{M}') \subseteq L|_a$$

Jeder akzeptierende Lauf ρ in \mathfrak{M}' für ein Wort w muss per Konstruktion einen der neuen a-Übergänge beinhalten.

Für Zustände $q, q' \in Q$ und $q_f \in F$ muss $\rho = q_0 \cdots q \widehat{q} \widehat{q'} \cdots \widehat{q_f}$.

Entfernen wir \widehat{q} sowie $\widehat{\cdot}$ von den übrigen Zuständen, erhalten wir einen akzeptierenden Lauf ρ' von $\mathfrak M$ für ein Wort $v \in \mathbf L(\mathfrak M)$.

Es gilt $w \in v|_a$, also $w \in L|_a$.

Wir zeigen für eine (beliebige) reguläre Sprache $L \subseteq \Sigma^*$ und ein Symbol $a \in \Sigma$, dass $L||_a$ ebenfalls regulär ist.

Beweis:

Da L regulär ist, existiert ein DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ mit $\mathbf{L}(\mathcal{M}) = L$.

Wir konstruieren einen NFA \mathfrak{M}' ausgehend von \mathfrak{M} wie folgt:

- (i) kopiere alle Zustände und Übergänge, sodass \mathfrak{M}' aus den Zuständen und Übergängen von \mathfrak{M} besteht und einer Kopie von \mathfrak{M} .
- (ii) Startzustand bleibt q_0 . Endzustandsmenge wird die Kopie von F.
- (iii) Führe zwischem jedem (Original-)Zustand $q \in Q$ und seiner Kopie \widehat{q} einen a-Übergang ein.

$$L(\mathcal{M}') \supseteq L|_{a}$$

Sei $w \in L|_a$.

Dann gibt es ein Wort $v \in L$ mit $v = u_1u_2$ und $w = u_1au_2$.

Es gibt einen akzeptierenden Lauf ρ in ${\mathfrak M}$ für v mit $\rho=\rho_1\rho_2.$

Sei q der letzte Zustand von ρ_1 und sei $\widehat{\rho_2}$ der Lauf, der aus ρ_2 entsteht, indem alle Zustände des Laufs durch ihre Kopien ausgetauscht werden.

Dann ist $\rho' = \rho_1 \widehat{q} \widehat{\rho_2}$ ein akzeptierender Lauf für w in \mathfrak{M}' . $w \in L(\mathfrak{M}')$.

Wir zeigen für eine (beliebige) reguläre Sprache $L \subseteq \Sigma^*$ und ein Symbol $a \in \Sigma$, dass $L|_a$ ebenfalls regulär ist.

Beweis:

Da L regulär ist, existiert ein DFA $\mathcal{M} = \langle Q, \Sigma, \delta, q_0, F \rangle$ mit $\mathbf{L}(\mathcal{M}) = L$.

Wir konstruieren einen NFA \mathfrak{M}' ausgehend von \mathfrak{M} wie folgt:

- (i) kopiere alle Zustände und Übergänge, sodass \mathfrak{M}' aus den Zuständen und Übergängen von \mathfrak{M} besteht und einer Kopie von \mathfrak{M} .
- (ii) Startzustand bleibt q_0 . Endzustandsmenge wird die Kopie von F.
- (iii) Führe zwischem jedem (Original-)Zustand $q \in Q$ und seiner Kopie \widehat{q} einen a-Übergang ein.

Es gilt also $\mathbf{L}(\mathfrak{M}') = L|_{a}$.

Da \mathfrak{M}' ein NFA ist, ist $L||_a$ somit regulär.

a) Zeigen Sie, dass die Formel

$$\mathit{F} := \neg b \lor \Big(\big(\neg a \lor \neg c \lor d \big) \land \big(\neg d \lor (a \land \neg c) \big) \Big)$$

äquivalent zu einer Horn-Formel ist.

b) Nutzen Sie das Resolutionsverfahren, um zu zeigen, dass

$$\{ (\neg a \lor b), (\neg a \lor d), (\neg b \lor \neg d \lor c), (\neg d \lor \neg c \lor \neg a) \} \models \neg a$$

gilt.

 Begründen Sie, warum die Hyperresolution für Horn-Formeln immer in polynomieller Zeit terminiert.

a) Zeigen Sie, dass die Formel

$$F := \neg b \lor \Big(\big(\neg a \lor \neg c \lor d \big) \land \big(\neg d \lor (a \land \neg c) \big) \Big)$$

äquivalent zu einer Horn-Formel ist.

Umwandlung in KNF ergibt:

$$F = \neg b \lor ((\neg a \lor \neg c \lor d) \land (\neg d \lor (a \land \neg c)))$$

$$\equiv (\neg a \lor \neg b \lor \neg c \lor d) \land (\neg b \lor \neg d \lor (a \land \neg c))$$

$$\equiv (\neg a \lor \neg b \lor \neg c \lor d) \land (a \lor \neg b \lor \neg d) \land (\neg b \lor \neg c \lor \neg d)$$

In keiner Klausel kommt mehr als ein positives Literal vor. *F* ist damit äquivalent zu einer Horn-Formel.

Tatsächlich gilt sogar $F \equiv (\neg a \lor \neg b \lor \neg c) \land (a \lor \neg b \lor \neg d)$.

b) Nutzen Sie das Resolutionsverfahren, um zu zeigen, dass gilt:

$$\{(\neg a \lor b), (\neg a \lor d), (\neg b \lor \neg d \lor c), (\neg d \lor \neg c \lor \neg a)\} \models \neg a$$

 $\{ (\neg a \lor b), (\neg a \lor d), (\neg b \lor \neg d \lor c), (\neg d \lor \neg c \lor \neg a), (\neg \neg a) \}$ ist unerfüllbar. Also gilt $\neg a$ in allen Modellen von $\{ (\neg a \lor b), (\neg a \lor d), (\neg b \lor \neg d \lor c), (\neg d \lor \neg c \lor \neg a) \}$. Damit gilt die Behauptung.

c) Begründen Sie, warum die Hyperresolution für Horn-Formeln immer in polynomieller Zeit terminiert.

Jede Resolvente ist von der Form $\top \rightarrow p$ für ein Atom p.

Für eine feste Formel φ gibt es nur linear viele Atome p, die in φ vorkommen.

Es gibt also nur linear viele solcher Resolventen.

Ein Hyperresolutionsschritt ist in polynomieller Zeit möglich.

Die Hyperresolution terminiert in polynomieller Zeit.