Marco Kunze (makunze@cs.tu-berlin.de)
Sebastian Nowozin (nowozin@cs.tu-berlin.de)
Philip Voss (vossy@cs.tu-berlin.de)

WS 2001/2002
21. 11. 2001

Lösungen zur Projektaufgabe TheGI1

Ansatz für $L_1 \circ L_2$

Behauptung: Seien L_1, L_2 regulär, dann ist auch $L_1 \circ L_2 = L = \{wv \mid w \in L_1, v \in L_2\}$ regulär.
Beweis: Da L_1, L_2 regulär sind, existieren die regulären Grammatiken $G_1 = (N_1, T_1, S_1, P_1)$ und $G_2 = (N_2, T_2, S_2, P_2)$, mit $L(G_1) = L_1$ und $L(G_2) = L_2$. Dabei sind die Elemente der N und P Mengen jeweils eindeutig mit einem Index gekennzeichnet, so dass $N_1 \cap N_2 = \emptyset$ und $P_1 \cap P_2 = \emptyset$. Sei nun $S' \in (N_1 \cup T_1 \cup N_2 \cup T_2)$ ein neues Startsymbol. Wir erzeugen eine neue Grammatik G' mit:

$$G' = (N_1 \cup N_2 \cup \{S'\}, T_1 \cup T_2, S', P')$$

$$S' = S_1$$

$$P' \quad = \quad \{Q \to pU, R \to tS_2 \mid Q \to pU, R \to t \in P_1\} \cup P_2$$

Dabei werden alle Produktionen aus P_1 nach P' übernommen. Genau diejenigen, die kein Nonterminal produzieren werden ergänzt, indem das Startsymbol S_2 angehängt wird.
Wir zeigen:

(a) G' ist regulär

Offensichtlich sind N' und T' disjunkt: $(N_1 \cup N_2 \cup \{S'\}) \cap (T_1 \cup T_2) = \emptyset$.

Weiter sind alle Produktionen P' aus G' von der erlaubten Bauart für reguläre Grammatiken, da $P_1 \cup P_2$ erlaubt sind und die zusätzlichen Produktionen auch regulär sind.

(b) $L(G') = L_1 \circ L_2 = L$

Es sind zwei Richtungen zu beweisen

- $L(G') \supseteq L$
  Sei $v \in L$. Dann existieren $u \in L_1$, $w \in L_2$ mit $v = uw$. Also $S' \Rightarrow^{\star}_{G'} v$.
  Dann existiert eine Ableitung

  $$S' \Rightarrow^{\star}_{G'} uS_2 \Rightarrow^{\star}_{G'} uw$$

  Da $S_1 \Rightarrow \overset{\star}{G_1}\, u$ immer mit einer Ableitung der Produktion der Form $(R \to t) \in P_1$ enden würde, genau diese jedoch so verändert wurden, dass sie S_2 produzieren, erhält man als produziertes Wort von G': uS_2.
  Also: $v \in L(G')$.
- $L(G') \subseteq L$
  Sei $v \in L(G')$, $u \in L(G_1)$, $w \in L(G_2)$. Dann existiert eine Ableitung

  $$S' \Rightarrow_{G_1} w_1 \Rightarrow_{G_1} w_2 \Rightarrow_{G_1} \ldots \Rightarrow_{G_1} w_n \Rightarrow_{G'} w_n u_1 \Rightarrow_{G_2} w_n u_2 \Rightarrow_{G_2} \ldots \Rightarrow_{G_2} w_n u_m = v$$

  Da $P_2 \in P'$ und $(P_1 = \{Q \to pU, R \to tS_2 \mid Q \to pU, R \to t \in P_1\}) \in P'$.
  Also $v = w_n u_m$, mit $w_n \in L_1$ und $u_m \in L_2$. Also $v \in L$.

Ansatz für $L_1 \cup L_2$

Behauptung: Seien L_1, L_2 regulär, und $w \in \Sigma^\star$. Wir zeigen, dass auch $L = \{w \mid w \in L_1 \vee w \in L_2\}$ regulär ist.
Beweis: Da L_1, L_2 regulär sind, existieren die regulären Grammatiken $G_1 = (N_1, T_1, S_1, P_1)$ und $G_2 = (N_2, T_2, S_2, P_2)$, mit $L(G_1) = L_1$ und $L(G_2) = L_2$. Dabei sind die Elemente der N und P Mengen jeweils eindeutig mit einem Index gekennzeichnet, so dass $N_1 \cap N_2 = \emptyset$ und $P_1 \cap P_2 = \emptyset$. Sei nun $S' \notin (N_1 \cup T_1 \cup N_2 \cup T_2)$ ein neues Startsymbol. Wir erzeugen eine neue Grammatik G' mit:

$$G' = (N_1 \cup N_2 \cup \{S'\}, T_1 \cup T_2, S', P_1 \cup P_2 \cup \{S' \to S_1 \mid S_2\})$$
$$G' = (N', T', S', P')$$

Wir zeigen:

(a) G' ist regulär

Offensichtlich sind N' und T' disjunkt: $(N_1 \cup N_2 \cup \{S'\}) \cap (T_1 \cup T_2) = \emptyset$.

Weiter sind alle Produktionen P' aus G' von der erlaubten Bauart für reguläre Grammatiken, da $P_1 \cup P_2$ erlaubt sind und die zusätzliche Regel $S' \to S_1 \mid S_2$ auch regulär ist. Also ist G' regulär.

(b) $L(G') = L$

Es sind zwei Richtungen zu beweisen

- $L(G') \supseteq L$
  Sei $v \in L$. Es ist zu zeigen, dass dann auch $S_1 \Rightarrow^\star_{G_1} v$ oder $S_2 \Rightarrow^\star_{G_2} v$ existieren.

  Da S' genau zwei mögliche Ableitungen, nämlich $S' \to S_1$ und $S' \to S_2$ besitzt, betrachten wir beide Fälle getrennt.
  Fall "$S' \to S_1$": $S' \Rightarrow_{G'} S_1 \Rightarrow^\star_{G_1} v$, da $P_1 \subseteq P'$.
  Fall "$S' \to S_2$": $S' \Rightarrow_{G'} S_2 \Rightarrow^\star_{G_2} v$, da $P_2 \subseteq P'$.
  Also ist $v \in L(G')$.
- $L(G') \subseteq L$
  Sei $v \in L(G')$. Es ist zu zeigen, dass dann auch eine Ableitung existiert, der Form:

  $$S' \Rightarrow_{G'} v$$

  Da S' nur einmal existiert, nämlich als Startproduktion $S' \to S_1 \mid S_2$, *muss* diese Regel im ersten Schritt, und danach nie wieder angewandt worden sein.
  Wir betrachten die zwei möglichen Ableitungen getrennt.
  Fall "S_1" ($P_1 \subseteq P'$):

  $$S' \Rightarrow_{G'} S_1 \Rightarrow_{G'} v$$

  Fall "S_2" ($P_2 \subseteq P'$):

  $$S' \Rightarrow_{G'} S_2 \Rightarrow_{G'} v$$

  Also: $S_1 \Rightarrow^\star_{G_1} v$ und $S_2 \Rightarrow^\star_{G_2} v$. Also $v \in L$.

Ansatz für $L^\star$

Behauptung: Sei L regulär. Dann ist auch auch $L^\star$ regulär.
Beweis: Wenn L regulär ist, so existiert eine reguläre Grammatik G.

$$G = (N, T, S, P)$$

Diese wird folgendermassen in eine neue Grammatik G' überführt:

$$G' = (N', T, S', P')$$

$$N' = N \cup \{S'\}$$

$$P' = P \cup \{S' \to S\} \cup \{S' \to \lambda\} \cup \{(R \to tS) \mid (R \to t) \in P\}$$

Wir zeigen:

(a) G' ist regulär

N' und T' disjunkt: $(N \cup \{S'\}) \cap T = \emptyset$.

Weiter sind alle Produktionen P' aus G' von der erlaubten Bauart für reguläre Grammatiken, da die zusätzliche Produktion der Form $(R \to tS)$ auch regulär sind. Also ist G' regulär.

(b) $L(G') = L$

Es sind zwei Richtungen zu beweisen

- $L(G') \supseteq L^\star$
  Sei $v \in L^\star$. Dann gibt es eine Ableitung der neuen Grammatik, so dass $S' \Rightarrow^\star_{G'} v$.
  Für $v = \lambda$ ($v \in L^0$):
  $$S' \Rightarrow \lambda$$
  Für $|v| > 0$ ($v \in L^n, n \geq 1$):
  $$S' \Rightarrow_{G'} S \Rightarrow^\star_{G'} w_1$$
  terminiert, wenn $w_1 = v$ (also $v \in L^1$), oder aber die Produktion terminiert das Wort nicht:
  $$S' \Rightarrow_{G'} S \Rightarrow^\star_{G'} w_1 S \Rightarrow^\star_{G'} w_1 w_2 S \Rightarrow^\star_{G'} \ldots \Rightarrow^\star_{G'} v \qquad \text{mit } w_n \in L, n \geq 1$$
  Damit ist gezeigt, dass G' alle Wörter in $L^\star$ produzieren kann, mit der verlangten Form $w_1 w_2 w_3 \ldots w_n$, $n \in \mathbb{N}$. Also ist $v \in L(G')$.
- $L(G') \subseteq L^\star$
  Sei $v \in L(G')$. Dann ist zu zeigen, dass dann auch eine Ableitung
  $$S' \Rightarrow_{G'} v$$
  existiert. Diese gibt es:
  Für $v \in L^0$:
  $$S' \Rightarrow_{G'} S \Rightarrow^\star_{G'} \lambda$$
  Für $v \in L^n, n \geq 1$:
  $$S' \Rightarrow_{G'} S \Rightarrow^\star_{G'} w_1 S \Rightarrow^\star_{G'} w_1 w_2 S \Rightarrow^\star_{G'} \ldots \Rightarrow^\star_{G'} v$$
  Da alle aus G' produzierten Teilwörter $w_1, w_2, \ldots, w_n \in L$ sind, ist $v \in L^\star$.

Ansatz für $\overline{L}$

Ein Automat M', der die Komplementärsprache $L'(= \overline{L})$ des Automaten M mit der Sprache L produziert, muß folgende Bedingungen erfüllen:

1.) Er darf keine Wörter akzeptieren, die in L sind

2.) Er muß alle restlichen Wörter akzeptieren.

Kostruktion des neuen Automaten:

Zu 1.) Formal ergibt sich: Wenn $M = (Q, \Sigma, q_A, \delta, E)$, dann ergibt sich für den Automaten, der die Komplementärsprache $\overline{L}$ produziert:

$$E' = Q \cap E$$

Zu 2.) Für jeden Zustand Q muß eine Zustandsüberführung bezüglich jeden Elements aus Σ erstellt werden, die in den neu definierten Zustand q'_{neu} führt. q'_{neu} ist akzeptierend, und führt bei jedem Terminal aus Σ zu sich selbst zurück.

$$Q' = Q \cup q'_{neu}$$

$$\delta' = \delta \cup \{(q'_{neu}, x, q'_{neu}) \mid x \in \Sigma\} \cup \{(q_E, x, q'_{neu}) \mid (q_E, x, q_F) \notin \delta, x \in \Sigma, q_E, q_F \in Q\}$$

$$E' = (Q \cap E) \cup q'_{neu}$$

Das Alphabet Σ und der Anfangszustand q_A ist für beide Automaten gleich.

Beweis der Konstruktion:

Zu 1.) Die Zustandsüberführungen des deterministischen Automaten M sind unverändert übernommen worden. Die ehemals akzeptierten Zustände sind nicht mehr akzeptierend. Daher führt jede Eingabe eines ehemals akzeptierten Wortes zwangsläufig in einen nicht akzeptierten Zustand. Also akzeptiert der Automat keines der ursprünglich akzeptierten Wörter.

Zu 2.) Jeder ehemals nicht akzeptierende Zustand ist nun akzeptierend. Für jede nicht vorhandene Kombination aus Zustand und gelesenem Zeichen ist eine Zustandsüberführung (q_E, x, q'_{neu}) hinzugefügt worden, die in einen neuen, akzeptierenden Zustand führt. Der neu eingeführte Zustand q'_{neu} führt nun bei jeder Eingabe zu sich selbst. Damit ist der Automat vollständig und jedes vorher nicht akzeptierte Wort wird nun akzeptiert.