Rechtslineare Sprachen $r - LIN(X^*)$ (6)

Korollar: $REG(X^*) = l - LIN(X^*)$.

Beweis: Die Definition von linkslinear ist hinsichtlich der Produktionen $z \longrightarrow z \cdot x$ spiegelbildlich aufgebaut. Ersetzt man in dem Beweis zu obigem Satz überall $z \longrightarrow z' \cdot x$ durch $z \longrightarrow x \cdot z'$, dann ergibt sich der Beweis für das Korollar.

Zusammenfassung

Hauptsatz (1)

$REG(X^*) = RAT(X^*) = REC(X^*) = r - LIN(X*) = l - LIN(X^*)$. $REG(X^*)$ ist abgeschlossen unter

$$\cup, \cdot, *, ^{-}, \cap, hom, hom^{-1}$$

und

$$hom : X^* \longrightarrow REG(X^*).$$

Jede dieser Sprachklassen legt andere Verallgemeinerungen nahe.

Zusammenfassung (1)

(1) REG : Hier liegen Verallgemeinerungen in zwei Richtungen nahe:

(a) Man ersetzt das Monoid X^* durch freie Gruppen oder freie kommutative Monoide oder andere Monoide.
Berstel, J.: Transductions and Contextfree Languages. Teubner, Stuttgart (1979), Hotz, G./Estenfeld, K.: Formale Sprachen. B I Wissenschaftsverlag.

(b) Man verallgemeinert die Wegekategorie. Anstelle der Wege treten z.B. Bäume oder allgemeiner Netze.
Hotz, G.: Hierarchischer Entwurf komplexer Systeme, In: Wegener, I. (Ed.): Highlights aus der Informatik. Springer-Verlag (1996)

(2) RAT : Man setzt anstelle von X^* beliebige Monoide oder Algebren.

Zusammenfassung (2)

(3) REC : Man ersetzt in den Darstellungen $hom^{-1}(M)$, das endliche Monoid durch einfach berechenbare unendliche Monoide. Das syntaktische Monoid $synt(D(Z))$ der Dycksprache ist dazu ein Beispiel.

(4) $r - LIN$: Hier geht man von der rechts- oder linkslinearen Chomsysprache zu allgemeineren Grammatikformen über.

Die für $REG(X^*)$ festgestellten Gleichheiten zwischen den aufgeführten Konzepten zur Definition von Sprachen bleiben bei diesen Verallgemeinerungen i.a. nicht erhalten.

Komplexitätsaspekte (1)

Wir haben gezeigt, daß $REG(X^*) = RAT(X^*)$ ist. Also kann man jede reguläre Sprache durch eine Folge rationaler Ausdrücke definieren.

Formal ist ein rationaler Ausdruck über einer Kleene-Algebra ein Wort über dem Alphabet

$$A = X \cup \{\emptyset\} \cup \{(,)\} \cup \{\cup, \cdot, *\},$$

das induktiv wie folgt definiert wird.

- $\emptyset, x$ für $x \in X$ sind rationale Ausdrücke.

- Sind ϱ_1, ϱ_2 rationale Ausdrücke, dann auch $(\varrho_1 \cup \varrho_2), (\varrho_1 \cdot \varrho_2), (\varrho_1^*)$.

- Jeder rationale Ausdruck läßt sich aus den elementaren Ausdrücken durch die Anwendung der im vorigen Punkt angegebenen Schritte erzeugen.

Komplexitätsaspekte (2)

Wir messen die Größe der rationalen Ausdrücke durch die Anzahl der Positionen in dem Ausdruck, auf denen Operationszeichen stehen.

Formal: Wir definieren einen Homomorphismus $\lambda : A^* \longrightarrow A^*$ durch

$$\lambda(\alpha) = \begin{cases} 1 & \text{für} \quad \alpha \in X \cup \{\emptyset\} \cup \{(,)\} \\ \alpha & \text{für} \quad \alpha \in \{\cup, \cdot, *\}. \end{cases}$$

Nun definieren wir für rationale Ausdrücke u

$$|u|_\lambda := |\lambda(u)|.$$

Uns interessiert nun der Zusammenhang zwischen $|u|_\lambda$ und $|a|$, wenn u und a die gleiche Sprache definieren. Hierbei legen wir unsere Konstruktionen in den Beweisen zu

$REG(X^*) \subset RAT(X^*)$ und $RAT(X^*) \subset REG(X^*)$

zugrunde. Man erkennt

$$\begin{aligned} |a \cup b| &= |a| + |b| + 2, \\ |a \cdot b| &= |a| + |b|, \\ |a^*| &= |a| + 2. \end{aligned}$$

Komplexitätsaspekte (3)

Hieraus folgt:

Die angegebene Konstruktion von Akzeptoren a_u zu vorgegebenen rationalen Ausdrücken erfüllt

$$|a_u| \leq 2 \cdot |u|_\lambda.$$

Wir betrachten nun den Beweis zu $REG \subset RAT$, und geben eine Abschätzung für die Größe des in dem Beweis zu a konstruierten rationalen Ausdrucks u_a an.

Zu jeder Kante s aus dem Graphen von a haben wir einen rationalen Ausdruck der Form

$$\mathcal{W}_{k+1} = \mathcal{W}_k(S,T) \cup \mathcal{W}_k(S,P) \cdot (s \cdot \mathcal{W}_k(R,P))^* \cdot s \cdot \mathcal{W}_k(R,T)$$

angegeben. Dieser Ausdruck v_{k+1} hat die Größe

$$|v_{k+1}|_\lambda = 6,$$

wenn wir die $\mathcal{W}_k(P,R)$ einmal als Alphabetelemente ansehen.

Komplexitätsaspekte (4)

Indem man von $v_N, N = \#E(a)$ ausgehend rekursiv die jeweils benötigten Ausdrücke einsetzt, dann erhält man eine Folge von Ausdrücken, deren Größe exponentiell mit der Rekursionstiefe zunimmt.

Zur Erläuterung betrachten wir ein einfaches Beispiel. Sei a definiert durch

$$\begin{tikzcd} S \arrow[r, "s_1"] \arrow[rr, bend right, "s_4"'] & Q \arrow[loop above, "s_3"] \arrow[r, "s_2"] & T \end{tikzcd}$$

mit $\mu(s_1) = \mu(s_2) = x$, $\mu(s_3) = s(s_4) = y$.

Komplexitätsaspekte (5)

Wir nehmen der Reihe nach die Kanten s_1, s_2, s_3, s_4 auf. Wir benötigen

$$\begin{aligned}
\mathcal{W}_4(S,T) &= \mathcal{W}_3(S,T) \cup \mathcal{W}_3(S,S) \\
&\quad \cdot (s_4 \cdot \mathcal{W}_3(T,S))^* \cdot s_4 \cdot \mathcal{W}_3(T,T) \\
\mathcal{W}_3(S,S) &= \mathcal{W}_2(S,S) \cup \mathcal{W}_2(S,Q) \\
&\quad \cdot (s_3 \cdot \mathcal{W}_2(Q,Q))^* \cdot s_3 \cdot \mathcal{W}_2(Q,T) \\
\mathcal{W}_3(S,T) &= \mathcal{W}_2(S,T) \cup \mathcal{W}_2(S,Q) \\
&\quad \cdot (s_3 \cdot \mathcal{W}_2(Q,Q))^* \cdot s_3 \cdot \mathcal{W}_2(Q,T) \\
\mathcal{W}_3(T,S) &= \mathcal{W}_2(T,S) \cup \mathcal{W}_2(T,Q) \\
&\quad \cdot (s_3 \cdot \mathcal{W}_2(Q,Q))^* \cdot s_3 \cdot \mathcal{W}_3(Q,S) \\
\mathcal{W}_3(T,T) &= \mathcal{W}_2(T,T) \cup \mathcal{W}_2(T,Q) \\
&\quad \cdot (s_3 \cdot \mathcal{W}_2(Q,Q))^* \cdot s_3 \cdot \mathcal{W}_2(Q,T).
\end{aligned}$$

Wenn wir nun die für den Index $k = 3$ berechneten Ausdrücke in $\mathcal{W}_4(S,T)$ einsetzen und diese Konstruktion rekursiv fortsetzen, dann erhalten wir rationale Ausdrücke von exponentiell wachsender Größe.

Komplexitätsaspekte (6)

Betrachten wir aber für $k = 1$

$$\begin{aligned} \mathcal{W}(S,Q) &= \mathcal{W}_0(S,Q) \cup \mathcal{W}_0(S,S) \\ &\quad \cdot (s_1 \cdot \mathcal{W}_0(Q,S)) \cdot s_1 \cdot \mathcal{W}_0(Q,Q) \\ &= \emptyset \cup \emptyset^* \cdot (s_1 \cdot \emptyset)^* \cdot s_1 \cdot \emptyset^*, \end{aligned}$$

dann sehen wir, daß schließlich Ausdrücke entstehen, die sehr viele überflüssige Bestandteile enthalten. Und es ist leicht zu sehen, daß

$$\mathcal{W}_a(S,T) = s_1 \cdot s_3^* \cdot s_2 \cup s_4$$

ist. Unsere Konstruktion eines rationalen Ausdrucks, die einen leicht verständlichen Beweis liefert, taugt also nicht für praktische Anwendungen, und sie gibt uns keine interessante Auskunft über die Ausdruckskraft von rationalen Ausdrücken unter komplexitätstheoretischen Aspekten.

Komplexitätsaspekte (7)

Wir beobachten aber in dem Beispiel $\mathcal{W}(S,Q)$, daß man in den rationalen Ausdrücken Kürzungen zulassen sollte, die durch die Axiome

$$\emptyset \cup u = u, \quad \emptyset \cdot u = \emptyset, \quad \emptyset^* = 1, \quad 1 \cdot u = u$$

beschrieben werden und durch assoziative „Umklammerungen“

$$(u \cdot (v \cdot w)) = ((u \cdot v) \cdot w)$$

und distributives Ausmultiplizieren

$$(u \cdot (v \cup w)) = ((u \cdot v) \cup (u \cdot w)).$$

Diese hier nur angedeutete Algebra der rationalen Ausdrücke ist noch nicht hinreichend untersucht worden.

Komplexitätsaspekte (8)

Eine Bemerkung zur Erweiterung der Kleene-Algebra durch Hinzunahme der Komplementbildung als zusätzliche Operation.

Wir betrachten also $(REG(X^*), \cup, \cdot, *, ^{-})$, worin

$$\overline{L} = X^* - L$$

ist.

Was erwarten wir hier, wenn wir unseren Beweis zu $L \in REG(X^*) \Rightarrow \overline{L} \in REG(X^*)$ anschauen:

Ist a gegeben und nichtdeterministisch, dann müssen wir damit rechnen, daß für jeden Akzeptor $\overline{a}$ für $\overline{L}$

$$|\overline{a}| \geq c \cdot 2^{|a|}$$

gilt, wie in den Übungen gezeigt wurde.

Komplexitätsaspekte (9)

Unsere Konstruktion für $a \cup b$ führt zu einem nichtdeterministischen Akzeptor. Iteriert man also diese Konstruktion nach der Idee

$$a,\ b,\ \overline{a},\ \overline{b},\ \overline{a} \cup \overline{b},\ \overline{\overline{a} \cup \overline{b}}, \ldots\ ,$$

dann liefert unsere Konstruktion Akzeptoren a_i, deren Größe wie

$$n, 2^n, 2^{2^n}, \ldots, 2^{2^{\cdot^{\cdot^{\cdot^{2^n}}}}}$$

wächst.

Man kann Beispiele dafür angegeben, daß dieses Wachstum von Akzeptoren relativ zu rationalen Ausdrücken über $(REG(X^*), \cup, \cdot, *, \_)$ unvermeidlich ist.