Prof. Dr. V. Strehl
Informatik 8
11. Januar 2005

Übungen zur Theoretische Informatik 3 WS 2004

Aufgabe 27: Hofstadters MIU-System

Lösungshinweise

Das MIU-System definiert eine formale Sprache wie folgt:

- Als Alphabet dienen die drei Symbole M,I,U und die MIU-Sprache besteht aus genau denjenigen Wörtern, die sich aus dem einen "Axiom" MI mit Hilfe der vier folgenden Ableitungsregeln herleiten lassen:

  1. Ist xI ein Wort der MIU-Sprache, dann ist auch xIU ein Wort der MIU-Sprache.
  2. Ist Mx ein Wort der MIU-Sprache, so gilt dies auch für Mxx.
  3. Ist xIIIy ein Wort der MIU-Sprache, so auch xUy.
  4. Ist xUUy ein Wort der MIU-Sprache, so auch xy.

  Dabei sind x und y Variable für Wörter über dem Alphabet {M,I,U}.

Eine typische Ableitung

$$\mathsf{MI} \vdash_2 \mathsf{MII} \vdash_2 \mathsf{MIIII} \vdash_3 \mathsf{MUI} \vdash_1 \mathsf{MUIU} \vdash_2 \mathsf{MUIUUIU} \vdash_4 \mathsf{MUIIU}$$

zeigt, dass alle die in ihr vorkommenden Wörter zur MIU-Sprache gehören.

- Bevor Sie weiterlesen, sollten Sie sich mit dem System praktisch vertraut machen und versuchen herauszufinden, ob das Wort MIUI oder das Wort MU zur MIU-Sprache gehören oder nicht, und wenn ja, wie man sie generiert,

Das MIU-System ist keine Grammatik im üblichen Sinne (wegen der zweiten Regel), und da die Regeln 3. und 4. verkürzend sind, ist nicht einmal klar, dass die so definierte MIU-Sprache überhaupt entscheidbar ist. Natürlich ist sie rekursiv-aufzählbar, denn man kann alle MIU-Wörter systematisch generieren. Tatsächlich, und das ist schon etwas überraschend, ist die MIU-Sprache nicht nur entscheidbar, sondern sogar regulär! Der erste Schritt ist nicht schwer. Zeigen sie folgende notwendige Bedingung:

- Wenn $x \in \{\mathsf{M}, \mathsf{I}, \mathsf{U}\}^*$ zur MIU-Sprache gehört, so beginnt x mit einem M, enthält kein weiteres M und die Anzahl der x vorkommenden I's ist nicht durch 3 teilbar.

  ▶ Das beweist man durch Induktion:

  – die Aussage trifft offensichtlich zu für das Axiom MI;

  – sie bleibt unter den vier Ableitungsregeln erhalten.

  ◀

Tatsächlich ist diese Bedingung sogar hinreichend:

– Wenn $x \in \{\mathsf{M}, \mathsf{I}, \mathsf{U}\}^*$ mit einem M beginnt, kein weiteres M enthält und die Anzahl der x vorkommenden I's nicht durch 3 teilbar ist, so gehört x zur MIU-Sprache.

  ▶ Das beweist man mittels folgender Schritte:

  – Für jedes $n \geq 0$ gilt $\mathsf{MI}^{2^n} \in \mathsf{MIU}$: das ergibt sich induktiv aus dem Axiom und iterierter Anwendung der 2. Regel:

$$\mathsf{MI}^{2^n} \vdash_2 \mathsf{MI}^{2^{n+1}}$$

  – Aus $x\mathsf{I}^3 \in \mathsf{MIU}$ folgt $x \in \mathsf{MIU}$ für beliebige $x \in \{\mathsf{M},\mathsf{I},\mathsf{U}\}^*$, denn

$$x\mathsf{I}^3 \vdash_1 x\mathsf{I}^3\mathsf{U} \vdash_3 x\mathsf{UU} \vdash_4 x$$

  – Ist $k \not\equiv 0 \bmod 3$ und ist $k \leq 2^n$ mit $k \equiv 2^n \bmod 3$, so folgt aus den beiden vorhergenden Bemerkungen

$$\mathsf{MI} \vdash^* \mathsf{MI}^{2^n} \vdash^* \mathsf{MI}^k \in \mathsf{MIU}$$

  – Ist $x \in \{\mathsf{M},\mathsf{I},\mathsf{U}\}^*$, wobei M nur als erstes Symbol vorkommt und im übrigen in x das Symbol I k-mal ($k \not\equiv 0 \bmod 3$) und das Symbol U ℓ-mal vorkommen, so gilt

$$\mathsf{MI} \vdash \mathsf{MI}^{k+3\ell} \vdash x$$

  wobei man in $\mathsf{MI}^{k+3\ell}$ passenden Dreiergruppen von I's jeweils zu U zusammenfasst.

  ◀

Die genannte notwendige und hinreichende Bedingung zeigt klar, dass die MIU-Sprache regulär ist.

- Konstruieren Sie einen endlichen Automaten, der die MIU-Sprache akzeptiert.

  ▶ Das ist eine Routinesache:

  – Man benötigt drei Zustände, um die Vorkommen von I modulo 3 zu zählen und einen weiteren Zustand, um die Sonderrolle des ersten Symbols M zu registrieren, und wenn man einev vollständigen DFA bauen will, noch einen "Müll"-Zustand. Sei also $Q = \{\triangleright, 0, 1, 2, \otimes\}$; dann ist

	M	I	U
$\triangleright$	0	$\otimes$	$\otimes$
0	$\otimes$	1	0
1	$\otimes$	2	1
2	$\otimes$	0	2
$\otimes$	$\otimes$	$\otimes$	$\otimes$

  ein vollständiger DFA, der MIU akzeptiert, wenn man noch $\triangleright$ als Anfangszustand und $\{1, 2\}$ als Endzustandsmenge kennzeichnet.

  ◀

Diese Konstruktion können und sollten Sie durchführen, auch wenn Sie den hinreichenden Teil der Charakterisierung nicht bewiesen haben. Der endliche Automat (Transfermatrix) bzw. ein aus ihm leicht herzuleitender regulärer Ausdruck für die MIU-Sprache dient nun für die nächsten Schritte:

- Sei μ_n die Anzahl der Wörter der Länge n der MIU-Sprache. Zeigen Sie, dass folgende C-Rekursion gilt:

$$\mu_{n+3} = 3\,\mu_{n+2} - 3\,\mu_{n+1} + 2\mu_n \quad (n \geq 0), \mu_0 = \mu_1 = 0, \mu_2 = 1$$

▶ Man berechnet die Transfermatrix, wobei der Müllzustand $\oplus$ irrelevant ist. Man kann auch noch den Anfangszustand weglassen, da man nach dem ersten Verlassen nie mehr dahin zurückkommt. Man muss sich nur merken, dass der Schritt $\triangleright \rightarrow_{\mathsf{M}} 0$ ein Symbol wert ist. Damit wird das Symbol M aber auch irrelevant. Man braucht also nur den kleineren Automaten $\mathcal{M}$

	I	U
0	1	0
1	2	1
2	0	2

zu betrachten (mit Anfangszustand 0 und Endzustandmenge $\{1, 2\}$.

- Die Transfermatrix dieses Automaten ist

$$M = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

- Das charakteristische Polynom von M ist

$$\chi(M) = \det(\lambda \cdot I - M) = \det \begin{bmatrix} \lambda - 1 & -1 & 0 \\ 0 & \lambda - 1 & -1 \\ -1 & 0 & \lambda - 1 \end{bmatrix}$$
$$= (\lambda - 1)^3 - 1 = \lambda^3 - 3\lambda^2 + 3\lambda - 2$$

(NB: $\chi(\lambda)$ ist auch Minimalpolynom von M)

- Damit ergibt sich die Rekursion

$$\mu_{n+3} = 3\mu_{n+2} - 3\mu_{n+1} + 2\mu_n \quad (n \geq 1)$$

- Die Anfangswerte (unter Berücksichtigung der führenden M) sind $\mu_0 = \mu_1 = 0, \mu_2 = 1, \mu_3 = 3$, so dass die Rekutsion sogar schon ab $n = 0$ gilt.

◀

- Leiten Sie die erzeugende Funktion her:

$$\sum_{n \geq 0} \mu_n\, z^n = \frac{z^2}{1 - 3z + 3z^2 - 2z^3}$$

▶ Es gibt verschiedene Möglichkeiten:

- Man kann die hergeleitete Rekursion hernehmen, weiss also, dass

$$\sum_{n\geq 0} \mu_n\, z^n = \frac{a_0 + a_1 z + a_2 z^2}{1 - 3z + 3z^2 - 2z^3}$$

und folgert $a_0 = a_1 = 0, a_2 = 1$ mittels Koeffizientenvergleich.

- Man kann das Ganze auch über reguläre Ausdrücke auufziehen. Dazu betrachtet man zum Automaten $\mathcal{M}$ die drei Sprachen L_0, L_1, L_2, die jeweils ausgehend von den Zuständen $0, 1, 2$ als Anfangszuständen mit der Endzustandsmenge $\{1, 2\}$ akzeptiert werden. Der Automat liefert ein lineares Gleichungssystem für die drei Sprachen:

$$\begin{aligned} L_0 &= \mathsf{U} \cdot L_0 + \mathsf{I} \cdot L_1 \\ L_1 &= \mathsf{U} \cdot L_1 + \mathsf{I} \cdot L_2 + \lambda \\ L_2 &= \mathsf{U} \cdot L_2 + \mathsf{I} \cdot L_0 + \lambda \end{aligned}$$

Daraus könnte man mittels iterierter Anwendung von "Ardens Lemma" reguläre Ausdrücke für die drei Sprachen gewinnen. Man kann aber auch direkt weiterfahren:

Seine $\ell_0(z), \ell_1(z), \ell_2(z)$ die erzeugenden Funktikonen der drei Sprachen. Da der Automat $\mathcal{M}$ deterministisch ist, gibt es keine Mehrdeutigkeitsprobleme und man kann direkt schliessen:

$$\begin{aligned} \ell_0(z) &= z \cdot \ell_0(z) + z \cdot \ell_1(z) \\ \ell_1(z) &= z \cdot \ell_1(z) + z \cdot \ell_2(z) + 1 \\ \ell_2(z) &= z \cdot \ell_2(z) + z \cdot \ell_0(z) + 1 \end{aligned}$$

Dieses lineare Gleichungssystem für $(\ell_0(z), \ell_1(z), \ell_2(z))$ per Elimination zu Lösen ist reine Routinesache und liefert:

$$\ell_0(z) = \frac{z}{1 - 3z + 3z^2 - 2z^3}$$

Also ist

$$\sum_{n\geq 0} \mu_n\, z^n = z \cdot \ell_0(z) = \frac{z^2}{1 - 3z + 3z^2 - 2z^3}$$

◀

- Das Nennerpolynom lässt sich leicht faktorisieren:

$$1 - 3z + 3z^2 - 2z^3 = (1 - 2z)(1 - z + z^2) = (1 - 2z)(1 - \alpha \cdot z)(1 - \widehat{\alpha} \cdot z)$$

Was sind α und $\widehat{\alpha}$? (Explizite Formel!)
Welchen Absolutbetrag haben α und $\widehat{\alpha}$?
Hinweis: α und $\widehat{\alpha}$ sind nicht reell.

▶ Das ist Routinerechnerei:

– durch Koeffizientenvergleich erhält man

$$\alpha + \widehat{\alpha} = 1, \quad \alpha \cdot \widehat{\alpha} = 1$$

und die quadratische Gleichung $\alpha^2 = \alpha - 1$ für α liefert

$$\alpha = \frac{1+\sqrt{-3}}{2}, \quad \widehat{\alpha} = \frac{1-\sqrt{-3}}{2}$$

– Die Rechnung zeigt:

$$\alpha \cdot \widehat{\alpha} = 1$$

d.h. α und $\widehat{\alpha}$ liegen auf dem Einheitskreis der komplexen Ebene. Die Rechnung

$$\alpha^3 = \alpha \cdot (\alpha^2 - 1) = -\alpha \cdot \widehat{\alpha} = -1$$

zeigt, dass $\alpha^6 = 1 = \widehat{\alpha}^6 = 1$, d.h. α und $\widehat{\alpha}$ haben die Ordnung (Periode) 6 in der mulitplikativen Gruppe $\mathbb{C}^*$. Man sagt: α (und $\widehat{\alpha}$) sind primitive sechste Einheitswurzeln.

◀

- Machen Sie Partialbruchzerlegung:

$$\frac{z^2}{1-3z+3z^2-2z^3} = \frac{a}{1-2z} + \frac{b}{1-\alpha \cdot z} + \frac{c}{1-\widehat{\alpha} \cdot z}$$

und erhalten Sie die Exponentialsummendarstellung

$$\mu_n = a \cdot 2^n + b \cdot \alpha^n + c \cdot \widehat{\alpha}^n$$

Geben Sie die Konstanten a, b, c explizit an.

▶ Das geht routinemässig:

– Koeffizientenvergleich für die ersten drei Werte führt auf ein lineares Gleichungssystem für a, b, c:

$$\begin{aligned} \mu_0 &= 0 = a + b + c \\ \mu_1 &= 0 = 2a + b\alpha + c\widehat{\alpha} = 2a + b\alpha + c(1-\alpha) \\ \mu_3 &= 1 = 4a + b\alpha^2 + c\widehat{\alpha}^2 = 4a + b(\alpha - 1) - c\alpha \end{aligned}$$

mit der Lösung

$$a = \frac{1}{3}, \quad b = \frac{1+\alpha}{3(1-2\alpha)} = -\frac{\widehat{\alpha}}{3}, \quad c = \frac{\alpha - 2}{3(1-2\alpha)} = -\frac{\alpha}{3}$$

◀

- Was können Sie über das asymptotische Verhalten von μ_n aussagen?
  Wie wirken sich die Terme $b \cdot \alpha^n + c \cdot \widehat{\alpha}^n$ aus?
  Hinweis: α^n und $\widehat{\alpha}^n$ sorgen für ein periodisches Verhalten mit der Periodenlänge 6.

▶ Aus obigen Resultaten ergibt sich:

- Da α und $\widehat{\alpha}$ den Absolutbetrag 1 haben, bleibt der Beitrag von $b \cdot \alpha^n + c \cdot \widehat{\alpha}^n$ beschränkt und ist für das asymptotische Verhalten irrelevant. Es gilt also:

$$\mu_n \sim_{n\to\infty} a \cdot 2^n = \frac{2^n}{3}$$

- Das periodische Verhalten von α^n und $\widehat{\alpha}^n$ zeigt sich konkret so:

$$\alpha^n + \widehat{\alpha}^n = \alpha^{n \bmod 6} + \widehat{\alpha}^{n \bmod 6} = \begin{cases} 1 & n \equiv 1 \bmod 6 \\ -1 & n \equiv 2 \bmod 6 \\ -2 & n \equiv 3 \bmod 6 \\ -1 & n \equiv 4 \bmod 6 \\ 1 & n \equiv 5 \bmod 6 \\ 2 & n \equiv 0 \bmod 6 \end{cases}$$

◀

- Zeigen Sie nun noch die Darstellung

$$\mu_n = \lfloor \frac{2^n}{3} \rfloor + \lfloor \frac{n \bmod 6}{3} \rfloor$$

Beachten Sie, der zweite Summand ist $= 0$ für $n \equiv 0, 1, 2 \bmod 6$ und $= 1$ für $n \equiv 3, 4, 5 \bmod 6$.

▶ Man kann das Ergebnis auch so schreiben:

$$3\,\mu_n = 2^n - (\alpha^{n-1} + \widehat{\alpha}^{n-1}) \quad (n \geq 0)$$

Daraus folgt die Behauptung. ◀

- Wenn Sie nun noch Claude Shannon nach der Coderate der MIU-Sprache fragte, was können Sie ihm antworten?

▶ Nach Definition:

$$\lim_{n\to\infty} \frac{\log_3 \mu_n}{n} = \log_3 2$$

◀