Matthias Jauernig
Kurs: 03In1

01.11.03

4. Aufgabenserie zu den Grundlagen der Informatik

Abgabetermin: Mi, 05.11.03

Zu 10.) Aufgabenstellung lautet jeweils „geben Sie an“, „wie lauten“ und „bestimmen Sie“; daher reicht das bloße Erwähnen der Endergebnisse aus:

(a) 43.34375_{10}
$= 101011.01011_2$
$= 53.26_8$
$= 2B.58_{16}$
$= 0100\ \ 0011\ .\ 0011\ \ 0100\ \ 0011\ \ 0111\ \ 0101_{\ BCD}$

(b) 0.1_{10}
$= 0.0\overline{0011}_2$ (Periode: 0011)
$= 0.0\overline{6314}_8$ (Periode: 6314)
$= 0.1\overline{9}_{16}$ (Periode: 9)

- 0.1_{10} ist im Dual-, Oktal- und Hexadezimalsystem nicht eindeutig, sondern nur periodisch darstellbar

(c) (i) $11011.1101_2 = 27.8125_{10}$
(ii) $472_8 = 314_{10}$
(iii) $1AFD.82_{16} = 6909.5078125_{10}$

Zu 11.) ***Pseudozufallszahlengenerator***

Hierin benutzte Begründungen:

(B1) Begründung, dass m die maximale Periodenlänge ist
(B2) Begründung, dass Periodenlänge = m = maximal für Theorem A

Allgemeine Vorbetrachtungen:

- $a, b \in \mathbb{N}_0$, $m \in \mathbb{N}$; $x_0 \in \mathbb{N}_0$, $x_0 < m$ als Anfangswert
- Iterationsverfahren zum Erzeugen einer Folge von Zufallszahlen:

$$\mathbf{x_{i+1} := (a*x_i+b) \bmod m} \qquad (i=0,1,2,...)$$

→ dieses Verfahren bezeichnet man als die „*Methode der linearen Kongruenz*“, die heute sehr häufig zum Einsatz kommt und schon Inhalt von mehrere Bücher füllenden mathematischen Betrachtungen war

→ diese Methode stellt einen (wenn geschickt eingesetzt) effizienten Algorithmus zur Ermittlung von Pseudozufallszahlen dar, bei welchem das „Zufallselement“ durch die Modulooperation mit m gegeben ist; damit entsteht eine Zahlenfolge, die mit gesundem Menschenverstand nicht mehr nachvollziehbar, also scheinbar „zufällig“ ist (in Wirklichkeit ist sie das natürlich nicht, da ein fest definierter Algorithmus eingesetzt wird → daher auch die Bezeichnung „Pseudozufallszahlengenerator“)

→ das Verfahren hat für alle a, b, m, x_0 eine endlich lange Periode n, was typisch ist für diese Art rekursiver Funktionen

→ die maximale Länge der Periode beträgt m;
Begründung: da der Term $a*x_i+b$ modulo m genommen wird, kann der Wert von x_{i+1} nicht größer als m-1 sein ($0 \leq x_{i+1} < m$) → das bedeutet, dass im Maximalfall m verschiedene Zufallszahlen erzeugt werden, bis eine Zahl generiert wird, welche schon einmal in der Folge vorkam; da keine Zahl der rekursiven Folge zweimal in einer Periode auftauchen kann, beginnt für den Fall der Wiederholung einer Zahl der Folge eine neue Periode → *das bedeutet, dass bei m Zufallszahlen die Periode höchstens die Länge m haben kann*

(B1)

- allgemein lässt sich zu den einzelnen Variablen in der Iterationsvorschrift Folgendes sagen:
  → b als Summand macht die Periode länger, wenn $b \neq 0$
  → m sollte generell groß gewählt werden, um die Periodenlänge zu erhöhen; mathematische Analysen haben gezeigt, dass m auf …x21 enden sollte, wobei x eine gerade Zahl ist → dadurch wird das Auftreten bestimmter problematischer Fälle verhindert (vgl. Sedgewick: „Algorithmen in C“)
  → a als Faktor sollte weder zu groß noch zu klein sein, am besten eine Ziffer weniger als m

Theorem A: Folge mit maximaler Periodenlänge m:

- mathematisch bewiesen ist, dass eine Periode ihre maximale Länge m (vgl. ***(B1)***) erreicht, wenn folgende Bedingungen erfüllt sind (vgl. D.E. Knuth: „The Art of Computer Programming“, Vol. 2: „Seminumerical Algorithms“):

(A1)	$m=2^k$	(d.h. m ist Zweierpotenz und damit Vielfaches von 4)
(A2)	$a \equiv 1 \pmod 4$	(d.h. a=4i+1, also ein um 1 inkrementiertes Vielfaches von 4 → a ist demnach ungerade)
(A3)	$ggt(b,m)=1$	(d.h. b ist relative Primzahl von m, da beide als gemeinsamen Teiler nur die 1 haben)

Bestätigung des Theorems für: m=16, a=9, b=3, x_0<m beliebig

- Es sei n die Periode der Folge; eben aufgrund der Periodizität dieser kann gesagt werden: ist n=m für **ein** x_0<m, so ist n=m ebenso für alle anderen x_0<m (vgl. ***(B2)***)
- Repräsentativ wähle ich x_0=5 aus. Ist die Periode der entstehenden Zahlenfolge nun n=m=16, so gilt dies für alle anderen x_0<16 gleichermaßen (vgl. ***(B2)***). Mit der gegebenen Iterationsvorschrift erhält man für x_0=5 folgende Wertetabelle:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	…
i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	…
x_i	5	0	3	14	1	12	15	10	13	8	11	6	9	4	7	2	5	0	…

→ Aus der Tabelle lässt sich leicht entnehmen, dass n=16 ist, wenn x_i alle Werte von 0 bis 15 angenommen hat. An dieser Stelle ist die maximale Periodenlänge erreicht, da diese höchstens so groß wie m sein kann (vgl. ***(B1)***). Da $x_{16}=x_0$ ist, beginnt für i=16 eine neue Periode.
→ Daraus folgt, dass die Periode für diese speziellen Werte von m, a, b, x_0 die Länge n=m=16 besitzt, welches nach ***(B1)*** die maximal erreichbare Länge der Folge darstellt.

→ Für x_0=5 ist *Theorem A also erfüllt*. Zu zeigen ist nun, dass daraus folgt, dass Theorem A auch für alle anderen x_0<16 erfüllt ist. Dies ist nämlich der Fall.
Begründung: setzt man m=16, a=9 und b=3 in die Iterationsvorschrift des Pseudozufallsgenerators ein, erhält man: $\mathbf{x_{i+1}:=(9x_i+3) \bmod 16}$.
Da die Folge wie mit obiger Wertetabelle gezeigt für x_0=5 periodisch mit m=16 als Periode ist, kommt jede der 16 möglichen Zahlen („k = i mod 16“ erzeugt in jedem Fall eine Zahl k, welche die Bedingung $0 \leq k < 16$ erfüllt, egal welchen Term oder Wert i enthält) einmal pro Periode vor.
Wählt man ein x_0<16, so muss dieser Wert demnach genau einmal in einer Periode der Folge mit x_0=5 vorkommen.

(B2)

Demnach entspricht jedes x_0<16 einem x_i aus der Folge mit x_0=5.
Die Bestätigung von Theorem A folgt nun daraus, dass sich die Zahlenfolge, wie ich sie für x_0=5 aufgestellt habe, für ein anderes x_0<16 nicht ändert. Es kommt lediglich zu einer *Verschiebung* jedes Folgenelements x_i, aufgrund der Periodizität der Folge mit der Periode m=16, wie ich sie für x_0=5 gezeigt habe, bleibt die Reihenfolge der Elemente allerdings erhalten, woraus folgt, dass sich jede Folge mit x_0<16 äquivalent zu der Folge mit x_0=5 verhält und somit auch dieselbe Periode enthält.

- ***Das bedeutet: Für jedes x_0<16 ist die Periodenlänge im Fall m=16, a=9 und b=3 gleich m; allgemein formuliert: Theorem A ist erfüllt für jedes x_0<m.***

→ Bemerkung: interessant ist auch zu sehen, dass die Bedingungen aus Theorem A zu einer Folge führen, welche alternierend gerade und ungerade Zahlen erzeugt. Dies lässt sich allerdings sehr einfach erklären: aus (A2) folgt, dass a ungerade ist; aus (A1) folgt, dass m gerade ist und aus (A3) folgt, dass b ungerade ist, wenn (A1) gilt. Demnach wird mit $(a*x_i+b)$ eine ungerade Zahl generiert, wenn x_i gerade ist. Durch (mod m) ändert sich daran nichts, da m gerade ist: die erzeugte Zahl x_{i+1} bleibt ungerade. Das Gegenteil ist der Fall, wenn x_i ungerade ist. $(a*x_i+b)$ wird gerade, ebenso $((a*x_i+b) \bmod m)$: die generierte Zahl x_{i+1} ist ebenfalls gerade.

→ **C-Programm zur Erzeugung von Zufallszahlen mithilfe der *Methode der linearen Kongruenz:*** **siehe Anhang**

Zu 12.) Lexikographische Ordnung

Gegeben: A … Alphabet
A^+ … Menge der nichtleeren Zeichenketten über A

Vorbetrachtungen:

- A besteht aus einer endlichen Menge von Symbolen
- B sei Wort bzw. Zeichenkette, d.h. eine endliche Aneinanderreihung (Verkettung) von Symbolen des zugrunde liegenden Alphabets A:
  $B^0 := \{\varepsilon\}$ (ε … leere Zeichenkette)
  $B^{i+1} := B^i \circ A$ für jedes $i \in \mathbb{N}$
  $\Rightarrow$ dann ist $A^+ := \bigcup_{i \in \mathbb{N}} B^i$

  A^+ bezeichnet somit die Menge aller Worte (Zeichenketten) über dem Alphabet A

- Es seien a, b, c $\subseteq A^+$ Zeichenketten über A und die Relation $R_{\leq}$ mit: a $R_{\leq}$ b $\Leftrightarrow$ a $\leq$ b
- Dann gilt: (1) a $\leq$ a, da a=a $\Rightarrow$ Relation ist reflexiv
  (2) a $\leq$ b $\wedge$ b $\leq$ c $\equiv$ a $\leq$ c $\Rightarrow$ Relation ist transitiv
  (3) a $\leq$ b $\wedge$ b $\leq$ a $\equiv$ a=b $\Rightarrow$ Relation ist anti-symmetrisch
  $\Rightarrow$ aus (1)-(3) folgt: $R_{\leq}$ ist Halbordnungsrelation über A^+
  (4) a $\leq$ b $\vee$ b $\leq$ a = "wahr" für alle a, b $\subseteq A^+$
  $\Rightarrow$ aus (4) folgt: Halbordnung $R_{\leq}$ ist Ordnung (-srelation) über A^+

Zu definieren: die zu der $\leq$-Relation gehörende lexikographische Ordnung in der Menge A^+ der nichtleeren Zeichenketten über A
(lexikographische Ordnung: Anordnung der Zeichenketten wie im Lexikon üblich)

Definitionsfestsetzungen:
- Es seien B, C $\subseteq A^+$ Zeichenketten über A
- Wenn ω ein leeres Zeichen und a ϵ A, dann gilt: $\omega \leq$ a
- Elemente der Menge A^+ aller nichtleeren Zeichenketten über A sind eindeutig, d.h. sie kommen nur genau einmal in A^+ vor.

Definition (verbal):
- Alle Elemente (sprich: Zeichenketten) aus A^+ werden durch $\leq$ lexikographisch geordnet. Dabei sind Zeichen (Buchstaben bzw. Elemente aus A), die in einem geordneten Alphabet A weiter links stehen, in Relation $\leq$ zu Zeichen, die in A an selber Position oder weiter rechts stehen.
- Die lexikographische Ordnung mit $\leq$ ergibt sich durch zeichenweises Vergleichen einer Zeichenkette mit den anderen Zeichenketten aus A^+ (um alle Zeichenketten aus A^+ lexikographisch zu ordnen, muss eines von vielen bekannten Sortierverfahren angewendet werden (Mergesort, Heapsort, Quicksort, Radix-Sort-Methoden etc.). Für 2 Zeichenketten bedeutet das: prüfen der Aussage $b_i \leq c_i$ mit b_i, c_i Zeichen aus B bzw. C ($b_i \epsilon$ B, $c_i \epsilon$ C) und i $\epsilon \mathbb{N}$, solange $b_i \wedge c_i \neq \omega$ und die Aussage wahr ist, $b_i \leq c_i$.
  Steht ein Zeichen $b_i \epsilon$ B in <-Relation zu einem Zeichen $c_i \epsilon$ C, so ist B lexikographisch weiter links einzuordnen als C; ist $b_i=c_i$, so wird i inkrementiert; steht b_i in >-Relation zu c_i, so ist B lexikographisch rechts von C einzuordnen.
  Da wie oben erwähnt keine Zeichenkette in A^+ doppelt vorkommt, wird A^+ über A mit der Ordnungsrelation $\leq$ lexikographisch geordnet. Ist $A_i \subseteq A^+$ mit i $\epsilon \mathbb{N}$, so gilt für eine lexikographische Ordnung allgemein: $A_0 \leq A_1 \leq \ldots \leq A_n$.