Programmiermethoden in der Mathematik Wintersemester 06/07 TU Berlin, Institut für Mathematik MA 4-5 Thomas Slawig, Uwe Prüfert

Zahlendarstellung auf dem Rechner (1)

Zahlensysteme

b-adisches Zahlensystem: $b \in \mathbb{N}, b > 1$.

Alphabet $\Sigma_b := \{0, 1, \dots, b-1\}.$

Beispiele:

 $b=10: \text{Dezimal-}, b=2: \text{Dual-/Bin\"ar-}, b=8: \text{Oktal-}, b=16: \text{Hexadezimal system } (\Sigma_{16}:=\{0,1,\ldots,9,A,B,C,D,E,F\}), b=10: \text{Dezimal-}, b=2: \text{Dual-/Bin\"ar-}, b=8: \text{Oktal-}, b=16: \text{Hexadezimal system } (\Sigma_{16}:=\{0,1,\ldots,9,A,B,C,D,E,F\}), b=10: \text{Dezimal-}, b=10: \text{Dezimal-},$

Eine feste "Wort"-Länge erreicht man durch führende Nullen: (01234)_b.

Wir definieren die Abbildungen:

div:
$$\mathbb{N} \times \mathbb{N} \to \mathbb{N}$$
, $(z, b) \mapsto z \operatorname{div} b := \max\{k \in \mathbb{N} : k \le z/b\}$
 $\operatorname{mod}: \mathbb{N} \times \mathbb{N} \to \Sigma_b$, $(z, b) \mapsto z \operatorname{mod} b := z - (z \operatorname{div} b) \cdot b$.

(ganzzahlige Division ohne Rest und Rest der ganzzahligen Division).

Es gilt f.a. $z \in \mathbb{N}$:

$$z = (z \operatorname{div} b) \cdot b + z \operatorname{mod} b.$$

Satz: Natürliche Zahlen können in jedem b-adischen Zahlensystem eindeutig dargestellt werden: Sei $b \in \mathbb{N}, b > 1$ gegeben. Dann gilt:

$$\forall z \in \mathbb{N} \quad \exists n \in \mathbb{N} \quad \exists ! z_i \in \Sigma_b, i = 0, \dots, n - 1: \qquad z = \sum_{i=0}^{n-1} z_i b^i. \tag{1}$$

Ziffernschreibweise: $z = (z_{n-1} \dots z_1 z_0)_b$.

Beweis:

1. Existenz (über vollst. Ind.):

Induktionsanfang: z < b hat Darstellung $z = (0 \dots 0z_0)$ mit $z_0 = z$.

Induktionsvoraussetzung: Die Aussage sei wahr für die Zahlen $0, \dots, z-1$, d.h. diese Zahlen haben eine Darstellung (1). Induktionsschluss: Zu zeigen ist, dass jetzt auch z (mit $z \ge b$) eine Darstellung (1) hat. Es gilt:

$$(z \operatorname{div} b) \cdot b < z$$

und damit wegen b > 1:

$$z' := z \operatorname{div} b < z$$
,

d.h. es existiert nach I.V. die Darstellung

$$z' := \sum_{i=0}^{n-1} z'_i b^i, z'_i \in \Sigma_b.$$

Außerdem:

$$z'b \leq z < b^n \Rightarrow z' < b^{n-1} \Rightarrow z'_{n-1} = 0 \Rightarrow z' = \sum_{i=0}^{n-2} z'_i b^i,$$

denn sonst

$$z' = \underbrace{z'_{n-1}b^{n-1}}_{\geq b^{n-1}, \text{wenn } z'_{n-1} \geq 1} + \sum_{i=0}^{n-2} \underbrace{z'_{i}b^{i}}_{\geq 0} \quad \geq \quad b^{n-1}.$$

Es gilt:

$$z \ = \ z'b + z \, \mathrm{mod} \, b \quad = \quad \left(\sum_{i=0}^{n-2} z_i'b^i\right)b + z \, \mathrm{mod} \, b \ = \ \sum_{i=0}^{n-2} z_i'b^{i+1} + z \, \mathrm{mod} \, b \ = \ \sum_{i=1}^{n-1} z_{i-1}' \, b^i + \underbrace{z \, \mathrm{mod} \, b}_{=:z_0} = \ \sum_{i=0}^{n-1} z_ib^i,$$

d.h. z hat Darstellung (1).

2. Eindeutigkeit:

Annahme: $\bar{z} = (z_{n-1} \dots z_1 z_0)_b = (\bar{z}_{n-1} \dots \bar{z}_1 \bar{z}_0)_b$.

Sei $l:=\max\{i=0,\ldots,n-1:z_i\neq \bar{z}_i\}$, also $z_i=\bar{z}_i$ f.a. $i=l+1,\ldots,n-1$. Daher gilt

$$\sum_{i=l+1}^{n-1} z_i b^i = \sum_{i=l+1}^{n-1} \bar{z}_i b^i,$$

also auch

$$\sum_{i=0}^{l} z_i b^i = \sum_{i=0}^{l} \bar{z}_i b^i,$$

d.h.

$$\sum_{i=0}^{l-1} z_i b^i + z_l b^l = \sum_{i=0}^{l-1} \bar{z}_i b^i + \bar{z}_l b^l.$$

O.B.d.A. sei $z_l > \bar{z}_l$. Dann folgt

$$\sum_{i=0}^{l-1} (\bar{z}_i - z_i)b^i = \underbrace{(z_l - \bar{z}_l)}_{\geq 1} b^l \geq b^l.$$

Andererseits

$$\sum_{i=0}^{l-1} (\bar{z}_i - z_i) b^i \leq \sum_{i=0}^{l-1} \underbrace{|\bar{z}_i - z_i|}_{\leq b-1} b^i \leq (b-1) \sum_{i=0}^{l-1} b^i = (b-1) \frac{b^l - 1}{(b-1)} = b^l - 1 < b^l$$

im Widerspruch zu oben.

Darstellung ganzer Zahlen auf dem Rechner

Ganze Zahlen (engl.: integer numbers, integers).

Auf dem Rechner: Zweiersystem, b = 2, (1 Stelle = 1 Bit).

Möglichkeit: "Z = N + Vorzeichen":

$$z = (-1)^{\nu} \sum_{i=0}^{n-2} z_i 2^i, \quad \nu, z_i \in \Sigma_2 = \{0, 1\}, i = 0, \dots, n-2$$

mit $z_{n-1} = \nu$ (Vorzeichenbit) benutzt man n Bits.

2. Möglichkeit: $z \ge 0$: mit z_i , $i = 0, \ldots, n-2$ wie oben, d.h. mit $z_{n-1} = 0$.

z<0im Zweierkomplement: $\tilde{z}:=|z|>0$ hat die Darstellung

$$\tilde{z} = \sum_{i=0}^{n-2} \tilde{z}_i 2^i, \quad \text{gespeichert als Bitmuster} \quad (0 \ \tilde{z}_{n-2} \cdots \tilde{z}_0)$$

Als Zweierkomplement von \tilde{z} bezeichnet man die Zahl

$$\tilde{\bar{z}} := \sum_{i=0}^{n-2} (1-\tilde{z}_i) 2^i - 2^{n-1}, \quad \text{gespeichert als} \quad (1, \, (1-\tilde{z}_{n-2}), \cdots, (1-\tilde{z}_0)), \text{ erhalten durch Bitumkehr von } \tilde{z}.$$

Die führende 1 in dem Bitmuster $(1, (1-\tilde{z}_{n-2}), \cdots, (1-\tilde{z}_0))$ rechts wird in der binären Darstellung links sozusagen mit Minuszeichen versehen und entspricht dem Summanden -2^{n-1} .

Es gilt (mit der Summenformel für die geom. Reihe)

$$\tilde{z} + \tilde{\bar{z}} = \sum_{i=0}^{n-2} 2^i - 2^{n-1} = \frac{1 - 2^{n-1}}{1 - 2} - 2^{n-1} = -1,$$

also

$$\bar{\tilde{z}} + 1 = -\tilde{z} = -|z| = z.$$

Die Darstellung von z < 0 ergibt sich also durch:

- 1. Berechnen der Darstellung von |z|
- 2. Bilden des Zweierkomplements dieser Darstellung
- 3. Addieren von 1.