Blatt 5 Sebastian Maschke

Sonntag, 24. Mai 2020 20:40

Aufgabe 16 Zahlendarstellung: Basiswechsel

a)

	Dezimal	Dual	Oktal	Hex
1)	30.125	1 1110. 001	36. 8	1E. 2
2)	27.484375	11011.011111	33. 37	17. 61
3)	61. 265625	111 101.010 001$_2$	75. 21	3D. 41
4)	10. 734 375	1010. 1100 1101	12. 532	A. BC

1) • $W_{10} = 30.125 \curvearrowright$ Dual

mit Euklidischem Algorithmus: $\boxed{W_{10} = \sum_{i=-m}^{n-1} b_i \cdot 2^i, \quad b_i \in \{0,1\}}$

-4	= 0.0625
-3	= 0.125
-2	= 0.25
-1	= 0.5
0	= 1
1	= 2
2	= 4
3	= 8
4	= 16
5	= 32

$2^{n-1} \leq W_{10} < 2^n \rightarrow n = 5$

- 30.125 div 2^4 = 1 Rest 14.125
- 14.125 div 2^3 = 1 Rest 6.125
- 4.125 div 2^2 = 1 Rest 2.125
- 2.125 div 2^1 = 1 R 0.125
- 0.125 div 2^0 = 0 R 0.125
- 0.125 div 2^{-1} = 0 R 0.125
- 0.125 div 2^{-2} = 0 R 0.125
- 0.125 div 2^{-3} = 1 R 0

$\hookrightarrow 30.125_{10} = 1\,1110.001_2$

$W_{10} = 30.125 \curvearrowright$ oktal

-2	0.015625
-1	0.125
0	0
1	8
2	64
3	512

• erneut Euklidischer Algorithmus:

$8^{n-1} \leq 30.125 \leq 8^n \rightarrow n = 2$

- 30.125 div 8^1 = 3 Rest 6.125
- 6.125 div 8^0 = 6 Rest 0.125
- 0.125 div 8^{-1} = 1 Rest 0

$\hookrightarrow 30.125_{10} = 36.1_8$

$W_{10} = 30.125 \curvearrowright$ Hex

-1	0.0625
0	1
1	16
2	256

• – " –

$16^{n-1} \leq 30.125 < 16^n \Rightarrow n = 2$

• $30.125 \text{ div } 16^1 = 1 \text{ R } 14.125$

• $14.125 \text{ div } 16^0 = 14 \text{ R } 0.125$

• $0.125 \text{ div } 16^{-1} = 2 \text{ R } 0$

$\hookrightarrow 30.125 = 1E.2$

2) • $W_2 = 1\,1011.01\,1111_2 \curvearrowright$ Dezimal

Vorkomma: $2^4 + 2^3 + 2^1 + 2^0 = 16+8+2+1 = 27$

Nachkomma: $2^{-2} + 2^{-3} + 2^{-4} + 2^{-5} + 2^{-6} = 0.484375$

$\hookrightarrow W_{10} = 27.484375$

• $W_2 \curvearrowright$ Oktal → drei-Bit zusammenfassen

$W_2 = 1\,1011.01\,1111 = 33.37_8$

• $W_2 \curvearrowright$ Hex → vier-Bit zusammenfassen

$W_2 = 1\,1011.0111\,1 = 17.61_{16}$

3) • $W_8 = 75.21 \curvearrowright$ Hex

Umweg über Dual

• $W_8 = 75.21_8 \curvearrowright$ Dual

löse dreier-Pakete auf

$W_8 = 75.21_8 = 111\,101.010\,001_2$

• $W_2 = 11\,1101.0100\,01_2 \curvearrowright$ Hex

$W_2 = 11\,1101.0100\,01_2 = 3D.41_{16}$

• $W_8 \curvearrowright$ Dezimal

$W_8 = 75.21_8 = 61.265625_{10}$

$\cdot\ W_8 \rightsquigarrow$ Dezimal

$W_8 = 75.21_8 = 61.265625_{10}$

4) $\cdot\ W_{16} = A.BC_{16} \rightsquigarrow$ Dual

$W_2 = 1010.\,1100\,1101_2$

$\cdot\ W_2 \rightsquigarrow$ oktal

$W_8 = 12.532_8$

$\cdot\ W_{16} \rightsquigarrow$ Dezimal

$W_{10} = 10 \cdot 16^0 + 11 \cdot 16^{-1} + 12 \cdot 16^{-2}$

$= 10.734375$

b) · von Dezimal nach Dual,- Oktal- und Hexadezimalsystem benötigt man einen Algorithmus z.B. euklidischer Algorithmus.

· Unter den Stellenwertsystemen Dual,- Oktal- und Hexadezimal kann man über Dualzahlen-Pakete günstig umrechnen.

Aufgabe 18

Zu einer Basis $b \in \mathbb{N}$, $b \geq 2$ lässt sich eine positive reele Zahl $x \in [b^{-n}, b^{n-1}]$ eindeutig als

$$x = b^e \cdot f = b^{t \cdot e(v)} \sum_{j=1}^{\infty} d_j b^{-j} \quad \text{mit } e(v) = \sum_{j=0}^{n-1} v_j b_j$$

und $1 \leq d_1 \leq b-1$, $0 \leq d_j \leq b-1$ für $j = 2, 3, \dots$ und $d_j < b-1$

Exponent e ist eine n-stellige natürliche Zahl $e(v)$ zur Basis b mit den Ziffern $v = (v_{n-1}, \dots, v_0)$ und Vorzeichen $t \in \{-1, 1\}$.

In dieser Aufgabe beschränken wir uns auf positive reelle Zahlen x, die sich zur gegebenen Basis b mit endlicher Mantissenlänge m darstellen lassen.

a) ✓

b) ✓

$$\begin{array}{cccc} b & m & n & x \\ | & | & / & | \end{array}$$

c) $[d, v, t] = flp(2, 3, 3, 0.0625)$

$$2^{n-1} \leq 0.0625 < 2^{n} \longrightarrow n = -5$$

$0.0625 \text{ div } 2^{-4} = 1 \quad R \quad 0$

$\hookrightarrow 0.0625_{10} = 0.0001_2$

$x = 2^{1 \cdot 3} \ldots$

Aufgabe 19

a) z.z. $x_{min} = b^{v-1}$ und $x_{max} = (1 - b^{-m}) \cdot b^{R}$

$x_{min} = 0.1000\ldots0 \cdot b^{(v)} = b^{v-1}$ da $v \in \min\{e\}$

$x_{max} = 0.aa\ldots a \cdot b^{R} = (1 - b^{-m}) \cdot b^{R}$ wobei $a = b-1$

b) $x \in [x_{min}, x_{max}]$ $\quad x = \left(\sum_{j=1}^{\infty} d_j b^{-j}\right) b^{e}$ mit $v \leq e \leq R,\ e \in \mathbb{Z}$

z.z.: $|fl(x) - x| \leq \frac{b^{-m}}{2} \cdot b^{e}$

$fl(x)$ entsteht durch Reduktionsabbildung $\pm \begin{cases} \left(\sum_{j=1}^{m} d_j b^{-j}\right) \cdot b^{e} & \text{falls } d_{m+1} < \frac{b}{2} \\ \left(\sum_{j=1}^{m} d_j b^{-j} + b^{-m}\right) \cdot b^{e} & \text{falls } d_{m+1} \geq \frac{b}{2} \end{cases}$

$\left| \left(\sum_{j=1}^{m} d_j b^{-j} + b^{-m}\right) \cdot b^{e} - \left(\sum_{j=1}^{\infty} d_j b^{-j}\right) \cdot b^{e} \right|$

Aufgabe 20:

$x = \left(\frac{3}{5}\right)_{10}, \quad y = \left(\frac{4}{7}\right)_{10}$

a) $x = \pm \left(\sum_{j=1}^{\infty} d_j b^{-j}\right) \cdot b^{e}; \quad b = 2$

- $\frac{3}{5} \cdot 2 = \frac{6}{5} \rightarrow 1$
  $\frac{1}{5} \cdot 2 = \frac{2}{5} \rightarrow 0$
  $\frac{2}{5} \cdot 2 = \frac{4}{5} \rightarrow 0$
  $\frac{4}{5} \cdot 2 = \frac{8}{5} \rightarrow 1$
  $\frac{3}{5} \cdot 2 = \frac{6}{5} \rightarrow 1$
  $\frac{1}{5}$

$m = 1001 \ 1001 \ldots$

$e = 0, \quad t = -1$

$\rightarrow x = 2^{-0} \cdot \left(1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 0 \cdot 2^{-3} + 1 \cdot 2^{-4} + 1 \cdot 2^{-5} + 0 \cdot 2^{-6} + 0 \cdot 2^{-7} + 1 \cdot 2^{-8} \ldots\right)$

- $\frac{4}{7} \cdot 2 = \frac{8}{7} \to 1$

$\frac{1}{7} \cdot 2 = \frac{2}{7} \to 0$

$\frac{2}{7} \cdot 2 = \frac{4}{7} \to 0$

$\vdots$

$m = 100\ 100 \ldots$

$e = 0 \quad t = -1$

$x = 2^{-0} \left(1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 0 \cdot 2^{-3} + 1 \cdot 2^{-4} + 0 \cdot 2^{-5} + 0 \cdot 2^{-6} + \ldots \right)$

- $x - y = \left(\frac{3}{5}\right)_{10} - \left(\frac{4}{7}\right)_{10} = \frac{1}{35}$

$\frac{1}{35} \cdot 2 = \frac{2}{35} \to 0$

$\frac{2}{35} \cdot 2 = \frac{4}{35} \to 0$

$\frac{4}{35} \cdot 2 = \frac{8}{35} \to 0$

$\frac{8}{35} \cdot 2 = \frac{16}{35} \to 0$

$\frac{16}{35} \cdot 2 = \frac{32}{35} \to 0$

$\frac{32}{35} \cdot 2 = \frac{64}{35} \to 1$

$\frac{29}{35} \cdot 2 = \frac{58}{35} \to 1$

$\frac{23}{35} \cdot 2 = \frac{46}{35} \to 1$

$\frac{11}{35} \cdot 2 = \frac{22}{35} \to 0$

$\frac{22}{35} \cdot 2 = \frac{44}{35} \to 1$

$\frac{9}{35} \cdot 2 = \frac{18}{35} \to 0$

$0.00000111010\ldots$

b) In $M(\overset{b}{2}, \overset{m}{5}, \overset{r}{-7}, \overset{R}{8})$: $\quad \left(eps = \frac{b^{1-m}}{2} = \frac{1}{32} \right)$

$\tilde{x} = 0.10011 \cdot 2^{0}$

$\tilde{y} = 0.10010 \cdot 2^{0}$

$\tilde{x} - \tilde{y} = \begin{array}{r} 0.10011 \\ -\ 0.10010 \end{array} = 0.00001 \Rightarrow 0.10000 \cdot 2^{-4}$

[Einerkomplement aus $\tilde{y}$: $0.10010 \to 0.01101$

+1 addieren : $\tilde{y}$: 0.01110

$\tilde{x}$ und $\tilde{y}$ addieren : $\begin{array}{r} 0.10011 \\ +\ 0.01110 \\ \hline 1.00001 \end{array}$]

In $M(2,3,-7,8)$: $\left(eps = \frac{1}{8}\right)$

$\tilde{x} = 0.100 \cdot 2^{0}$

$\tilde{y} = 0.100 \cdot 2^{0}$

$\tilde{x} - \tilde{y} = \begin{array}{r} 0.100 \\ -\,0.100 \end{array} = 0.000 \cdot 10^{0}$ Auslöschung

c) • In $M(2,5,-7,8)$

$\tilde{x} \ominus \tilde{y} = 0.10000 \cdot 2^{-4}$

$\underbrace{(\tilde{x} \ominus \tilde{y})}_{fl(x)} - \underbrace{(x-y)}_{x} = 0.1 \cdot 2^{-4} - \frac{1}{35} = -\frac{5}{224}$

$\left| \frac{fl(x)-x}{x} \right| = \frac{25}{32}$ rel. Fehler in Ausgabe

$\max\left\{ \frac{|\tilde{x}-x|}{|x|}, \frac{|\tilde{y}-y|}{|y|} \right\} = \max\left\{ \frac{\left|\frac{19}{32} - \frac{3}{5}\right|}{\left|\frac{3}{5}\right|}, \frac{\left|\frac{9}{16} - \frac{4}{7}\right|}{\left|\frac{4}{7}\right|} \right\}$

$\delta_x \qquad \delta_y$

$= \max\left\{ \frac{1}{96}, \frac{1}{64} \right\}$ rel. Fehler in Eingabe

$\text{Verstärkung} = \frac{\text{rel. Fehler Ausgabe}}{\text{rel. Fehler Eingabe}} = \frac{\frac{25}{32}}{\frac{1}{64}} = 50$

$\tilde{x} \ominus \tilde{y} = 0.10000 \cdot 2^{-4}$

$x - y = \frac{1}{35} = 0.00000111010 = 0.111010\ldots \cdot 2^{-5}$

Es stimmen nur die ersten Stellen überein

• In $M(2,3,-7,8)$

$\tilde{x} \ominus \tilde{y} = 0$

$(\tilde{x} \ominus \tilde{y}) - (x-y) = -\frac{1}{35}$

$\left| \frac{fl(x)-x}{x} \right| = 1$ rel. Fehler Ausgabe

$$\max\left\{ \underbrace{\frac{|\tilde{x}-x|}{|x|}}_{\delta_x},\ \underbrace{\frac{|\tilde{y}-y|}{|y|}}_{\delta_y} \right\} = \max\left\{ \frac{\left|\frac{1}{2}-\frac{3}{5}\right|}{\left|\frac{3}{5}\right|},\ \frac{\left|\frac{1}{2}-\frac{4}{7}\right|}{\left|\frac{4}{7}\right|} \right\}$$

$$= \max\left\{ \frac{1}{6},\ \frac{1}{8} \right\}$$ rel. Fehler in Eingabe

$$\text{Verstärkung} = \frac{\text{rel. Fehler Ausgabe}}{\text{rel. Fehler Eingabe}} = \frac{1}{\frac{1}{6}} = 6$$