

6.2 Kodierung von Zahlen

Neue Begriffe

- Festkommadarstellungen
 - Zahlendarstellung durch Betrag und Vorzeichen
 - Einer-/Zweierkomplement-Darstellung
- Gleitkommadarstellung
 - IEEE-754 Format

BB T I I 6.2/1

Zahlensysteme

Definition

Ein **Stellenwertsystem (Zahlensystem)** ist ein Tripel

$S = (b, Z, \delta)$ mit den folgenden Eigenschaften:

- $b \geq 2$ ist eine natürliche Zahl, die **Basis** des Stellenwertsystems.
- Z ist eine b - elementige Menge von Symbolen, den Ziffern.
- $\delta : Z \rightarrow \{0, 1, \dots, b-1\}$ ist eine Abbildung, die jeder Ziffer umkehrbar eindeutig eine natürliche Zahl zwischen 0 und $b-1$ zuordnet.

BB T I I 6.2/2

Zahlensysteme

Beispiele

- **Dualsystem** $b=2, \quad Z = \{0,1\}$
- **Oktalsystem** $b=8, \quad Z = \{0,1,2,3,4,5,6,7\}$
- **Dezimalsystem** $b=10, \quad Z = \{0,1,2,3,4,5,6,7,8,9\}$
- **Hexadezimalsystem:**
 $b=16 \quad Z = \{0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F\}$

BB T I I 6.2/3

Festkommazahlen

Definition:

Eine **Festkommazahl** ist

eine endliche Folge von Ziffern aus einem Zahlensystem zur Basis b mit Ziffernmenge Z .

Sie besteht aus $n+1$ Vorkommastellen ($n \geq 0$) und $k \geq 0$ Nachkommastellen.

Der Wert $\langle d \rangle$ einer nicht-negativen Festkommazahl

$$d = d_n d_{n-1} \dots d_1 d_0 d_{-1} \dots d_{-k}$$

mit $d_i \in Z$ ist gegeben durch

$$\langle d \rangle = \sum_{i=-k}^n b^i \cdot d(d_i)$$

BB T I I 6.2/4

Festkommazahlen

Schreibweise

Vorkomma- und Nachkommastellen werden zur Verdeutlichung durch ein Komma oder einen Punkt getrennt:

$$d = d_n d_{n-1} \dots d_1 d_0 . d_{-1} \dots d_{-k}$$

Um anzudeuten, welches Zahlensystem zugrunde liegt, wird gelegentlich die Basis als Index an die Ziffernfolge angehängt.

Bsp.: 0110_2

BB T I I

6.2/5

Festkommazahlen

Beispiele

Seien $n = 3$, $k = 0$, $d = 0110$

$b = 2$	$\langle d \rangle = 6$
$b = 8$	$\langle d \rangle = 72$
$b = 10$	$\langle d \rangle = 110$
$b = 16$	$\langle d \rangle = 272$

BB T I I

6.2/6

Negative Festkommazahlen

(Im folgenden $b = 2$.)

Bei der Darstellung negativer Festkommazahlen nimmt die höchstwertige Stelle d_n eine Sonderrolle ein:

- Ist $d_n = 0$, so handelt es sich um eine nichtnegative Zahl.

Bei der Darstellung negativer Zahlen gibt es folgende Alternativen:

■ Darstellung durch Betrag und Vorzeichen

$$[d_n, d_{n-1}, \dots, d_0, d_{-1}, \dots, d_{-k}]_{BV} := (-1)^{d_n} \sum_{i=-k, \dots, n-1} d_i 2^i$$

■ Einer-Komplement Darstellung

$$[d_n, d_{n-1}, \dots, d_0, d_{-1}, \dots, d_{-k}]_1 := \sum_{i=-k, \dots, n-1} d_i 2^i - d_n (2^n - 2^{-k})$$

■ Zweier-Komplement Darstellung

$$[d_n, d_{n-1}, \dots, d_0, d_{-1}, \dots, d_{-k}]_2 := \sum_{i=-k, \dots, n-1} d_i 2^i - d_n 2^n$$

Betrag und Vorzeichen

$$[d_n, d_{n-1}, \dots, d_0, d_{-1}, \dots, d_{-k}]_{BV} := (-1)^{d_n} \sum_{i=-k, \dots, n-1} d_i 2^i$$

Beispiel: $n = 2, k = 0$

a	000	001	010	011	100	101	110	111
$[a]_{BV}$	0	1	2	3	0	-1	-2	-3

Eigenschaften:

- Der Zahlenbereich ist symmetrisch:
 - Kleinste Zahl: $-(2^n - 2^{-k})$
 - Größte Zahl: $2^n - 2^{-k}$
- Man erhält zu a die inverse Zahl, indem man alle Bits komplementiert.
- Zwei Darstellungen für die Null (000 und 100 bei $n = 2, k = 0$).
- „Benachbarte Zahlen“ haben gleichen Abstand 2^{-k} .

Einer-Komplement

$$[d_n, d_{n-1}, \dots, d_0, d_{-1}, \dots, d_{-k}]_1 := \sum_{i=-k, \dots, n-1} d_i 2^i - d_n (2^n - 2^{-k})$$

Beispiel: $n = 2, k = 0$

a	000	001	010	011	100	101	110	111
$[a]_1$	0	1	2	3	-3	-2	-1	0

Eigenschaften:

- Der Zahlenbereich ist symmetrisch:
 - Kleinste Zahl: $-(2^n - 2^{-k})$
 - Größte Zahl: $2^n - 2^{-k}$
- Zwei Darstellungen für die Null (000 und 111 bei $n = 2, k = 0$).
- „Benachbarte Zahlen“ haben gleichen Abstand 2^{-k} .
- **Lemma:** Sei a eine Festkommazahl, a' die Festkommazahl, die aus a durch Komplementieren aller Bits ($0 \rightarrow 1, 1 \rightarrow 0$) hervorgeht. Dann gilt $[a']_1 = -[a]_1$.
- Man erhält zu a die inverse Zahl, indem man alle Bits komplementiert.

Zweier-Komplement

$$[d_n, d_{n-1}, \dots, d_0, d_{-1}, \dots, d_{-k}]_2 := \sum_{i=-k, \dots, n-1} d_i 2^i - d_n 2^n$$

Beispiel: $n = 2, k = 0$

a	000	001	010	011	100	101	110	111
$[a]_2$	0	1	2	3	-4	-3	-2	-1

Eigenschaften:

- Der Zahlenbereich ist unsymmetrisch:
 - Kleinste Zahl: -2^n
 - Größte Zahl: $2^n - 2^{-k}$
- Die Zahlendarstellung ist eindeutig.
- „Benachbarte Zahlen“ haben gleichen Abstand 2^{-k} .
- **Lemma:** Sei a eine Festkommazahl, a' die Festkommazahl, die aus a durch Komplementieren aller Bits ($0 \rightarrow 1, 1 \rightarrow 0$) hervorgeht. Dann gilt $[a']_2 + 2^{-k} = -[a]_2$.
- Man erhält zu a die inverse Zahl, indem man alle Bits komplementiert und an der niederwertigsten Stelle 2^{-k} addiert.

Vorteil des Zweierkomplements:

Schaltkreise zur Realisierung der Addition / Subtraktion zweier vorzeichenbehafteter Festkommazahlen werden einfach. (→ später)

Probleme bei Festkommazahlen

Betrachte die Menge aller Zahlen, die eine Zweierkomplement-Darstellung mit n Vorkommastellen und k Nachkommastellen haben.

- keine ganz großen bzw. kleinen Zahlen darstellbar !

Zahlen mit größtem Absolutbetrag: -2^n und 2^{n-2^k}

Zahlen mit kleinstem Absolutbetrag: -2^{-k} und 2^{-k}

- Operationen sind nicht abgeschlossen !

$2^{n-1} + 2^{n-1}$ ist nicht darstellbar, obwohl die Operanden darstellbar sind.

- Assoziativgesetz und Distributivgesetz gelten nicht, da bei Anwendung der Gesetze evtl. der darstellbare Zahlenbereich verlassen wird!

Bsp.:

$$(2^{n-1} + 2^{n-1}) - 2^{n-1} \rightarrow \leftarrow 2^{n-1} + (2^{n-1} - 2^{n-1})$$

BB TII 6.2/11

Gleitkommadarstellung

Position des Kommas liegt nicht fest !

- ☛ Abdeckung eines größeren Zahlenbereichs bei gegebener Stellenanzahl

- Gleitkommadarstellung einfacher Genauigkeit: $(-1)^S \cdot M \cdot 2^E$

31	30	29	28	27	26	25	24	23	22	21	...	3	2	1	0
S	Exponent E								Mantisse M						

Gleitkommadarstellung doppelter Genauigkeit: $(-1)^S \cdot M \cdot 2^E$

63	62	61	...	53	52	51	50	...	3	2	1	0
S	Exponent E								Mantisse M			

Es bleibt noch festzulegen, wie die Mantissenbits bzw. Exponentenbits als Zahlen M bzw. E interpretiert werden sollen.

BB TII 6.2/12

Normalisierte Gleitkommadarstellungen

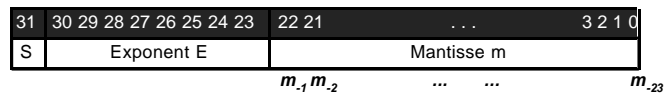
Beobachtung

Gleitkommadarstellung einer Zahl ist nicht eindeutig !
 $0.111 \cdot 2^3 = 0.0111 \cdot 2^4$

Definition

Eine Gleitkommazahl (S,M,E) heißt **normalisiert**, wenn $1 \leq M < 2$
d.h. wenn M von der Form $1.m_{-1} \dots m_{-k}$ ist.

Die 1 vor dem Komma braucht nicht abgespeichert zu werden (→ „hidden bit“)

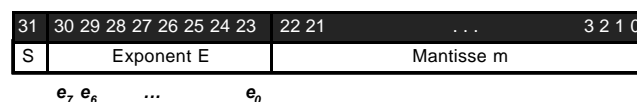


Für eine **normalisierte** Gleitkommazahl ergibt sich der Mantissenwert M als $M = 1 + \sum_{i=-1, \dots, -k} m_i 2^i$.

► Die Zahl 0 muss als Spezialfall behandelt werden !

BB TII 6.2/13

Gleitkommadarstellung - IEEE 754 Standard



- Gemäß IEEE 754-Standard werden die Exponentenbits als vorzeichenlose Zahl interpretiert.
- Um auch negative Exponenten darstellen zu können, wird von der Interpretation als vorzeichenlose Zahl eine Konstante, der sogenannte **Bias**, subtrahiert.
- Bei n Exponentenbits wird der Bias gewählt als **BIAS = $2^{n-1} - 1$** , also bei einfacher Genauigkeit BIAS = 127, bei doppelter Genauigkeit BIAS = 1023.
- Bei n Exponentenbits ergibt sich also für E:

$$E = \sum_{i=0, \dots, n-1} e_i 2^i - \text{BIAS}$$

BB TII 6.2/14

Sonderfälle IEEE 754 Standard

- **Der Exponent 0 spielt beim IEEE 754-Standard eine Sonderrolle:**
Sind alle Exponentenbits 0, so wird **ausnahmsweise** das „hidden bit“ der Mantissendarstellung weggelassen, so dass die Zahl

dargestellt wird.

$$(-1)^s \cdot 2^{-126} \cdot \sum_{i=1}^k m_i 2^i$$

- Auf diese Weise können „**denormalisierte Zahlen**“ dargestellt werden, die kleiner als die kleinste darstellbare normalisierte Zahl sind.
- Auf diese Weise kann die **Null** dargestellt werden: Sämtliche Mantissenbits und Exponentenbits sind 0.
- **Der Exponent $2^n - 1$ spielt ebenfalls eine Sonderrolle:**
Sind alle Exponentenbits 1 und alle Mantissenbits 0, so wird der Wert ∞ dargestellt.

BB T I I 6.2/15

Darstellbare normalisierte Gleitkommazahlen

	single precision	double precision
Vorzeichenstellen	1	1
Exponentenstellen	8	11
Mantissenstellen (ohne hidden Bit)	23	52
Bitstellen insgesamt	32	64
Bias	127	1023
Exponentenbereich	-126 bis 127	-1022 bis 1023
Darstellbare normalisierte Zahl mit kleinstem Absolutbetrag	2^{-126}	2^{-1022}
Darstellbare normalisierte Zahl mit größtem Absolutbetrag	$(1-2^{-24}) 2^{128}$	$(1-2^{-53}) 2^{1024}$
Darstellbare denormalisierte Zahl mit kleinstem Absolutbetrag	2^{-149}	2^{-1074}
Darstellbare denormalisierte Zahl mit größtem Absolutbetrag	$(1-2^{-23}) 2^{-126}$	$(1-2^{-52}) 2^{-1022}$

BB T I I 6.2/16

IEEE 754 Standard - Eigenschaften

- Eindeutige Zahlendarstellung, falls auf normalisierte Darstellungen beschränkt
- Nicht alle Zahlen zwischen der kleinsten und größten darstellbaren Zahl sind darstellbar.
- Je näher bei der Null, desto dichter liegen die darstellbaren Zahlen.
- Arithmetische Operationen sind nicht abgeschlossen!
- Assoziativgesetz und Distributivgesetz gelten nicht, da bei Anwendung der Gesetze evtl. der darstellbare Zahlenbereich verlassen wird!

BB TII 6.2/17

Prinzipielle Arbeitsweise: Addition von Gleitkommazahlen

Rechenvorschrift

- Angleichung des kleineren an den grösseren Exponenten
- Addition der Mantissen
- Normalisierung, Rundung (falls erforderlich)

Beispiel

$$\begin{aligned}+(1.000)_2 \cdot 2^{-1} + -(1.110)_2 \cdot 2^{-2} &= +(1.000)_2 \cdot 2^{-1} + -(0.111)_2 \cdot 2^{-1} \\ &= +(0.001)_2 \cdot 2^{-1} \\ &= +(1.000)_2 \cdot 2^{-4}\end{aligned}$$

BB TII 6.2/18

Prinzipielle Arbeitsweise: Multiplikation von Gleitkommazahlen

Rechenvorschrift

- Multipliziere die Vorzeichen
- Multipliziere die beiden Mantissen
- Addiere die beiden Exponenten und subtrahiere (einmal) den Bias-Wert
- Normalisierung, Rundung (falls erforderlich)

Beispiel

$$+(1.000)_2 \cdot 2^{-1+\text{BIAS}} \times -(1.110)_2 \cdot 2^{-2+\text{BIAS}}$$

Multiplikation der Vorzeichen: $0 \oplus 1 = 1$

Multiplikation der Mantissen: $(1.000)_2 \times (1.110)_2 = (1.110)_2$

Addition der Exponenten: $(-1+\text{BIAS}) + (-2+\text{BIAS}) - \text{BIAS} = (-3+\text{BIAS})$

Resultat: $-(1.110)_2 \cdot 2^{-3+\text{BIAS}}$