GTI – ÜBUNG 4

BINÄR-, HEX- UND GLEITKOMMAZAHLEN-ARITHMETIK

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- Führen Sie die folgenden Berechnungen im angegebenen Zahlensystem aus, ohne die Zahlen ins Dezimalsystem umzuwandeln:
- Hinweis: Arbeiten Sie mit dem Zweierkomplement und benutzen Sie 10 Bits für die Binärzahlendarstellung

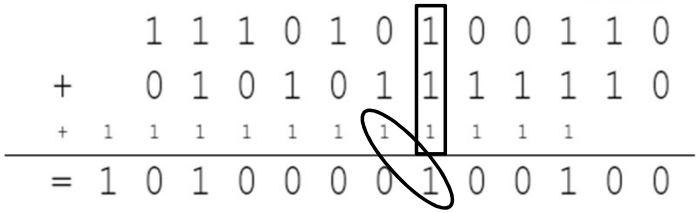
- $111010100110_2 + 010101111110_2$ (Addition im Dualsystem)
- $B674FC12_{16} + 2DA9D4B2_{16}$ (Addition im Hexadezimalsystem)
- $11101010_2 \cdot 1011_2$ (Multiplikation im Dualsystem)
- $11010010_2 - 10110101_2$ (Subtraktion im Dualsystem)
- $01110110_2 - 10011001_2$ (Subtraktion im Dualsystem)

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- Addition in fremden Systemen

Ziel: wir berechnen a + b = c im System d



Ansatz: wir addieren jeweils stellenweise

Hierbei gilt:

- Ist die Rechnung zu kompliziert, wandle a und b in das Dezimalsystem und c wieder zurück
- Hat das Ergebnis der Addition mehrere Stellen, schreibe nur die letzte Stelle nieder
- Der Rest e wandert als Übertrag zur Berechnung der nächsten Stelle
- Die Berechnung der nächsten Stelle lautet dann: a + b + e = c

Beispiel: $1_2 + 1_2 + 1_2 = 11_2$, d.h. 1_2 Übertrag 1_2 (siehe oben)

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- $111010100110_2 + 010101111110_2$ (Addition im Dualsystem)

Tipp: Rechenregeln erfüllen bekannte Dezimaladditionsregeln, d.h. z.B.
$1_2+1_2+1_2+1_2+1_2 = 5_{10} = 101_2$ (wobei hier 10 ein Übertrag ist)

Mechanismus der Addition nach Grundschuladdition:

```
      1 1 1 0 1 0 1 0 0 1 1 0
+     0 1 0 1 0 1 1 1 1 1 1 0
  + 1 1 1 1 1 1 1 1 1 1 1
-------------------------------
= 1 0 1 0 0 0 0 1 0 0 1 0 0
```

Bei jeder Stelle nach dem Tipp das Ergebnis bestimmen.

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- $B674FC12_{16} + 2DA9D4B2_{16}$ (Addition im Hexadezimalsystem)

Tipp: wer das Ergebnis einer Hexadezimaladdition a + b = c nicht direkt sieht, wandelt vorher a und b in das Dezimalsystem, c wieder zurück

Mechanismus der Addition nach Grundschuladdition:

```
   B 6 7 4 F C 1 2
 + 2 D A 9 D 4 B 2
 + 1 1   1 1
-----------------
 = E 4 1 E D 0 C 4
```

Bei jeder Stelle nach dem Tipp das Ergebnis bestimmen.

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Einschub: BCD-Addition

Pseudotetrade:

Binärdarstellung von nicht dezimalen Ziffern, also 1010 bis 1111 (10 bis 15)

Addition:

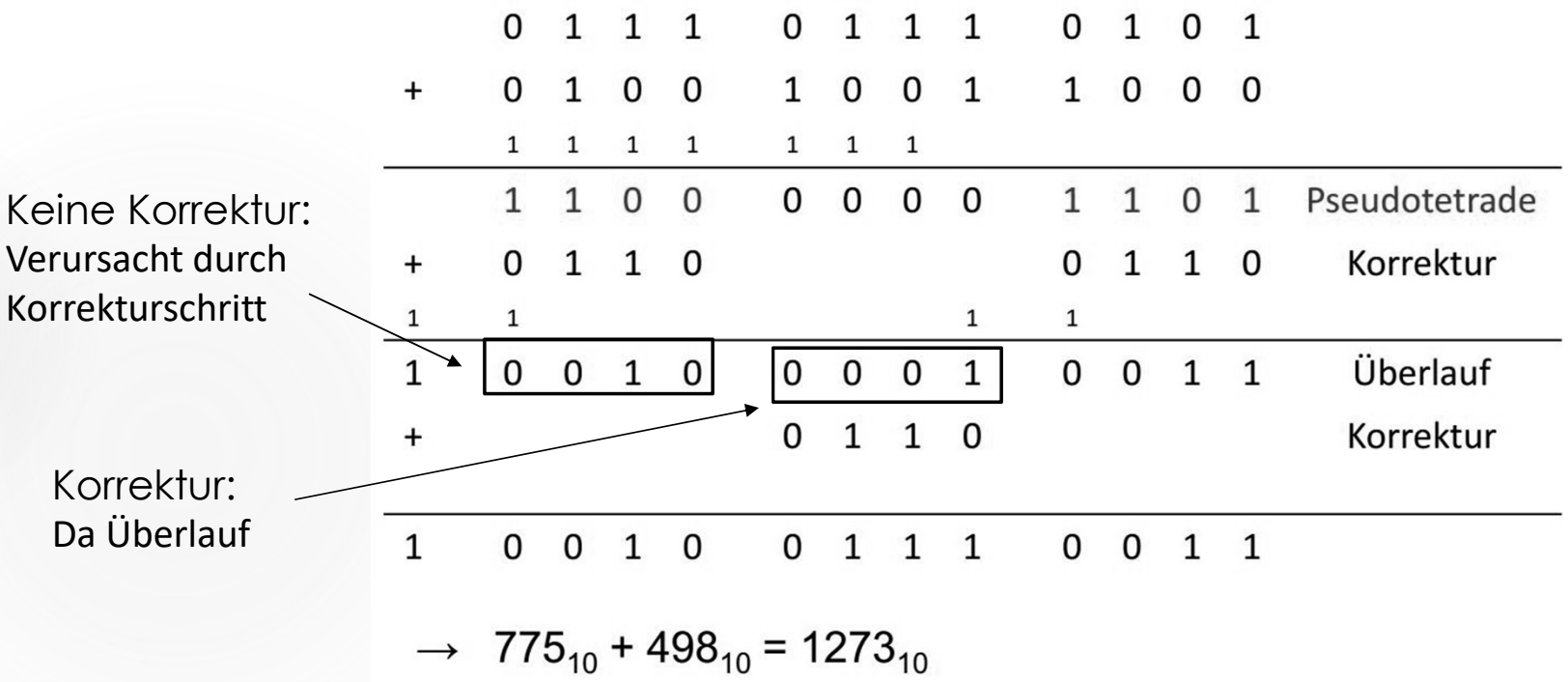
Addition funktioniert prinzipiell wie die Binäraddition, mit folgenden Abweichungen:

- Stellen (je 4 bit) werden getrennt addiert und der Übertrag in die nächste Stelle gezogen
- Pseudotetraden müssen korrigiert werden
- Überlauf muss korrigiert werden, wenn er nicht durch einen Korrekturschritt verursacht wurde
- Korrekturschritt ist die Addtion von 6 (0110)

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Einschub: BCD-Addition

Beispiel:



$\rightarrow \quad 775_{10} + 498_{10} = 1273_{10}$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

- Subtraktion im 2er-Komplement

Ziel: wir berechnen a - b = c im Binärsystem und nutzen das 2er-Komplement

Ansatz: wir bilden die Subtraktion auf die Addition ab

a + (-b) = c; -b bilden wir mit Hilfe des 2er-Komplements

Fallunterscheidung:

ist das erste Bit von c gesetzt → c positiv, Bit vorne abschneiden

ist das erste Bit nicht gesetzt → c negativ, noch 2er Komplement bilden

Aufgabe 1 – Bin- und Hex Arithmetik

- Subtraktion $11010010_2 - 10110101_2$

Ansatz: a + (-b) = c; -b bilden wir mit Hilfe des 2er-Komplements

Wir rechnen mit 10 Bits

	0011010010	Minuend
−	0010110101	Subtrahend
	1101001010	Subtrahend (B-1)
+	0000000001	
	1101001011	Subtrahend (B-2)
+	0011010010	Minuend
=	(1) 0000011101	Übertrag = 1
		Ergebnis positiv
	0000011101	Ergebnis

Bilden des 2er -Komplements

Addition a + (-b)

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- Subtraktion $01110110_2 - 10011001_2$

Ansatz: a + (-b) = c; -b bilden wir mit Hilfe des 2er-Komplements

	0001110110	Minuend
−	0010011001	Subtrahend
	1101100110	Subtrahend (B-1)
+	0000000001	
	1101100111	Subtrahend (B-2)
+	0001110110	Minuend
=	0|1111011101	Übertrag = 0
		Ergebnis negativ
	1111011101	Ergebnis
	0000100010	Betragsbildung
+	0000000001	zur Kontrolle
	0000100011	Betrag

Bilden des 2er -Komplements

Addition a + (-b)

Bilden des 2er -Komplements

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- Multiplikation in fremden Systemen

Ziel: wir berechnen $a \cdot b = c$ im System d

Ansatz: wir bilden die Multiplikation auf die Addition ab

d.h. wir bilden die Summe aus $a \cdot b_{STELLE1} + a \cdot b_{STELLE2} + \ldots + a \cdot b_{STELLEn}$

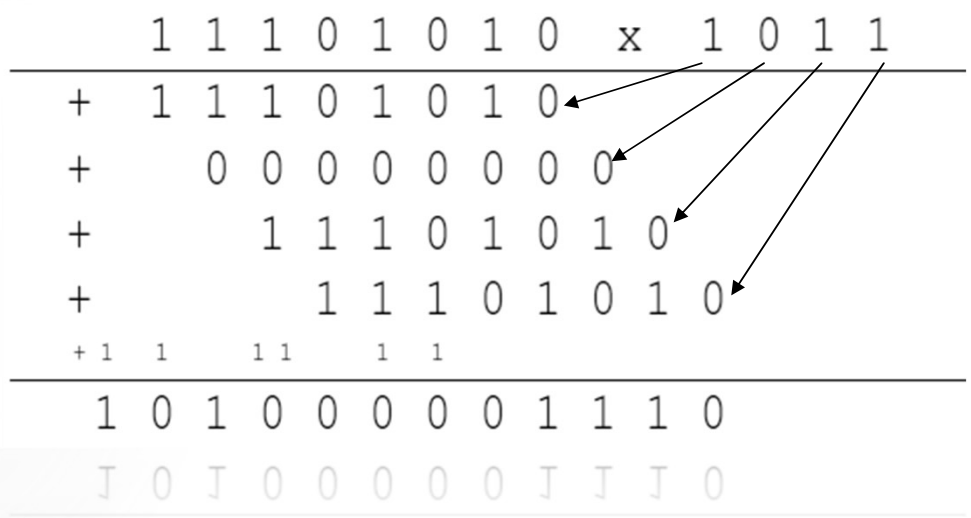
$$
\begin{array}{lll}
 & a_0\, a_1\, a_2 \ldots a_n & \\
\hline
 & a_0\, a_1\, a_2 \ldots a_n & * \quad b_n \\
+ & a_0\, a_1\, a_2 \ldots a_n & * \quad b_{n-1} \\
+ & \vdots & \vdots \\
+ & a_0\, a_1\, a_2 \ldots a_n & * \quad b_1 \\
\hline
 & c_0\, c_1 \ldots \quad c_n &
\end{array}
$$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- $11101010_2 * 1011_2$ (Multiplikation im Dualsystem

Mechanismus der Abbildung auf die Addition:



Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

- Prozessoren arbeiten mit einer endlichen Wortbreite, was den darstellbaren Zahlenbereich einschränkt. Welche Auswirkungen kann dies bei der Subtraktion und Addition haben?
- Können Sie sich vorstellen, wie diese Probleme in der Praxis gelöst werden?

Hinweis: Was passiert beim Verlassen des Zahlenbereichs

Aufgabe 1 – Bin- und Hex Arithmetik

- Auswirkungen bei der Subtraktion und Addition mit endlicher Wortbreite

Im Folgenden sei eine Wortbreite von 8 bit angenommen.

Problem 1: Addition vorzeichenloser Zahlen ($[0,255]$):

$128 + 128 = 0$!

```
  10000000
+ 10000000
-----------
 100000000
```

Lösung: Carry Flag im Prozessor (Übertrag von der n-ten in die (n + 1)-te Stelle)

Programmierer kann Flag explizit überprüfen und verarbeiten

Dadurch lassen sich auch breitere Worte per Software emulieren (z. B. 64-Bit-Worte auf einem 32-Bit-Prozessor).

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 1 – Bin- und Hex Arithmetik

- Auswirkungen bei der Subtraktion und Addition mit endlicher Wortbreite

Problem 2: Addition im 2er-Komplement ($[-128, 127]$):

$127 + 3 = -126!$ (Überlauf)

```
  0 1111111
  0 0000011 +
 ---------
  1 0000010 =
 01 111111
```

Überlauf nur bei: Addition von Zahlen gleiches Vorzeichens

Subtraktion von Zahlen unterschiedliches Vorzeichens

Lösung: Overflow-Flag im Prozessor (zeigt Überlauf an)

Berechnung: $OF = c_{n-1} \oplus c_n$

Im Beispiel: $0 \oplus 1 = 1$ (Overflow gesetzt)

Dadurch lassen sich auch breitere Worte per Software emulieren.

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Zurück zu Gleitkommazahlen

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

- Gegeben seien folgende Zahlen im IEEE Standard 754 (einfache Genauigkeit):

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_2 = 0\ 1001\ 1001\ 1111\ 1111\ 0000\ 0000\ 0000\ 000_2$

- Addieren Sie die beiden Zahlen.

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Addition im IEEE-Format:

1: Transformation

Rechtsschieben der kleineren Zahl auf den Exponenten der Größeren.

2: Addition der Mantissen

Falls Ergebnis < 0: setze Vorzeichenbit und bilde Zweierkomplement.

3: Normalisierung

A, Falls Ergebnis ≥ 2: Rechtsschieben des Ergebnisses um eins (ggf. runden) und Inkrementierung des Exponenten.

B, Falls Ergebnis < 1: Linksschieben des Ergebnisses um eins und Dekrementierung des Exponenten.

C, Wiederhole A bzw. B bis Ergebnis = 0 oder $1 \leq$ Ergebnis < 2

4: Behandlung von Sonderfällen (Überlauf, Unterlauf, Null)

Aufgabe 2 – Addition & Subtraktion

- Addition von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_2 = 0\ 1001\ 1001\ 1111\ 1111\ 0000\ 0000\ 0000\ 000_2$

1: Transformation

x_1 und x_2 besitzt den gleichen Exponenten, weswegen ein Verschieben entfällt.

E = 1001 1001

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

- Addition von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_2 = 0\ 1001\ 1001\ 1111\ 1111\ 0000\ 0000\ 0000\ 000_2$

2: Addition der Mantissen

	m_1:	$1{,}0000\ 1001\ 1001_2$
+	m_2:	$1{,}1111\ 1111\ 0000_2$
	$m_{ges:}$	$11{,}0000\ 1000\ 1001_2$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

- Addition von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_2 = 0\ 1001\ 1001\ 1111\ 1111\ 0000\ 0000\ 0000\ 000_2$

3: Normalisierung

$11{,}0000\ 1000\ 1001_2 \rightarrow 1{,}10000\ 1000\ 1001_2 \cdot 2^1$

Exponent anpassen

$2^1 \cdot 2^E = 2^{E+1}$

E = 1001 1001 + 1 = 1001 1010

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

- Addition von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_2 = 0\ 1001\ 1001\ 1111\ 1111\ 0000\ 0000\ 0000\ 000_2$

4: Zusammenfassung

VB:	0 (Addition zweier positiver Zahlen)
e_1:	$1001\ 1010_2$
$m_{ges:}$	$1{,}1000\ 0100\ 0100\ 1000\ 0000\ 000_2$

Ergebnis x_3: $0\ 1001\ 1010\ \ 1000\ 0100\ 0100\ 1000\ 0000\ 000_2$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Subtraktion im IEEE-Format (Abbildung auf die Addition a + (-b) = c):

1: Transformation

Rechtsschieben der kleineren Zahl auf den Exponenten der Größeren.

- 2: Zweierkomplement von b bilden

3: Addition der Mantissen

Falls Ergebnis < 0: setze Vorzeichenbit und bilde Zweierkomplement.

4: Normalisierung

A, Falls Ergebnis ≥ 2: Rechtsschieben des Ergebnisses um eins (ggf. runden) und Inkrementierung des Exponenten.

B, Falls Ergebnis < 1: Linksschieben des Ergebnisses um eins und Dekrementierung des Exponenten.

C, Wiederhole A bzw. B bis Ergebnis = 0 oder $1 \leq$ Ergebnis < 2

5: Behandlung von Sonderfällen (Überlauf, Unterlauf, Null)

Aufgabe 2 – Addition & Subtraktion

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

- „Subtraktion“ $x_1 + x_3$ von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_3 = 1\ 1001\ 0100\ 0011\ 0110\ 0000\ 0000\ 0000\ 000_2$

1: Transformation

Die Exponenten sind unterschiedlich, also müssen wir den kleineren Exponenten auf den größeren schieben.

$$1001\ 1001_2 - 1001\ 0100_2 = 101_2\ (= 5_{10})$$

Wir müssen also m_3 5 Stellen nach rechts schieben (2^5 ist positiv!)

m_3 (alt) = $1{,}0011\ 0110\ 0000\ 0000\ 0000\ 000_2$

m_3 (neu) = $0{,}0000\ 1001\ 1011\ 0000\ 0000\ 000_2$

Aufgabe 2 – Addition & Subtraktion

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

- „Subtraktion“ $x_1 + x_3$ von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_3 = 1\ 1001\ 0100\ 0011\ 0110\ 0000\ 0000\ 0000\ 000_2$

2: Zweierkomplement der Mantisse von x_3 (VB = 1)

Wir erweitern die Mantissen von x_1 und x_3 um ein führendes Vorzeichenbit, um das Vorzeichen des Ergebnisses der Addition überprüfen zu können.

m_3:	$00,\ 0000\ 1001\ 1011_2$
Einerkomplement:	$11,\ 1111\ 0110\ 0100_2$
Zweierkomplement:	$11,\ 1111\ 0110\ 0101_2$

Erinnerung
Erste 1 von rechts suchen. Linken Rest invertieren

Aufgabe 2 – Addition & Subtraktion

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

- „Subtraktion“ $x_1 + x_3$ von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_3 = 1\ 1001\ 0100\ 0011\ 0110\ 0000\ 0000\ 0000\ 000_2$

3: Addition der Mantissen

	m_1:	$01, 0000\ 1001\ 1001_2$
+	$m_{3,\,neu}$:	$11, 1111\ 0110\ 0101_2$
	$m_{ges:}$	$00, 1111\ 1111\ 1110_2$

Diesmal ist das Vorzeichenbit nicht gesetzt (deshalb auch Erweiterung um führende 0), also müssen wir vom Ergebnis nicht mehr das Zweierkomplement bilden.

Aufgabe 2 – Addition & Subtraktion

- „Subtraktion“ $x_1 + x_3$ von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_3 = 1\ 1001\ 0100\ 0011\ 0110\ 0000\ 0000\ 0000\ 000_2$

4: Normalisierung

Wir müssen unsere Zahl 1 nach links schieben, um auf das Format 1,… zu kommen.

$0,\ 1111\ 1111\ 1110_2$

$\rightarrow 1,111\ 1111\ 1110_2 \cdot 2^{-1}$

Exponent anpassen

$2^{-1} \cdot 2^E = 2^{E-1}$

E = 1001 1001 - 1 = 1001 1000

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 2 – Addition & Subtraktion

- „Subtraktion“ $x_1 + x_3$ von

  $x_1 = 0\ 1001\ 1001\ 0000\ 1001\ 1001\ 0000\ 0000\ 000_2$

  $x_3 = 1\ 1001\ 0100\ 0011\ 0110\ 0000\ 0000\ 0000\ 000_2$

5: Zusammenfassung

VB:	0 (da Ergebnis positiv; vgl. vorheriger Schritt)
e_1:	$1001\ 1000_2$
$m_{ges\ (2er):}$	$01,\ 1111\ 1111\ 1100\ 0000\ 0000\ 000_2$

Ergebnis x_3: $0\ 1001\ 1000\ 1111\ 1111\ 1100\ 0000\ 0000\ 000_2$

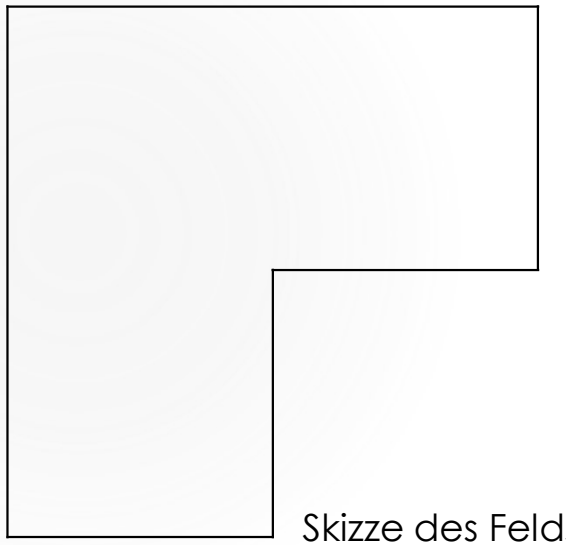
Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Denkpause

Aufgabe:

Ein Bauer möchte sein Feld in vier gleich große, identisch geformte Teilstücke aufteilen, um sie seinen vier Kindern zu vermachen.

Wie muss aufgeteilt werden?



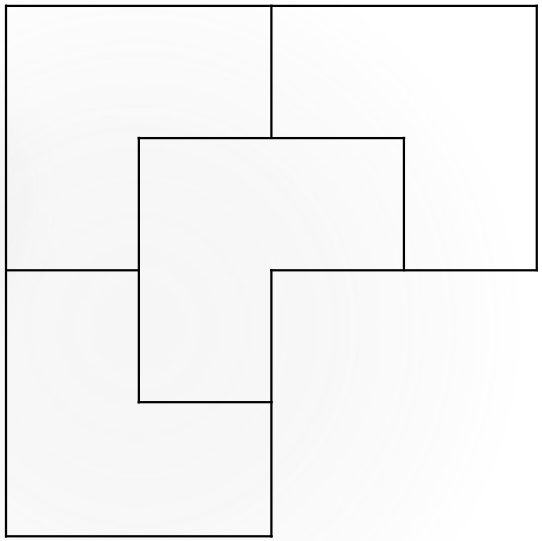
Skizze des Felds

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Denkpause

Lösung:

Folgende Aufteilung erfüllt alle Kriterien.



Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

- Gegeben seien folgende Gleitkommazahlen:

  $x_1 = 0\ 1001\ 000\ 1001\ 1011_2$

  $x_2 = 1\ 1001\ 010\ 1110\ 1000_2$

- Für die beiden Zahlen gilt:
  - Vorzeichen (V): 1 Bit breit (1: negativ)
  - Exponent (E): 7 Bit breit
  - Mantisse (M): 8 Bit breit (1,M wie beim IEEE Format üblich)
- Anordnung: VEM
- Berechnen Sie $x_1 \cdot x_2$.

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

- IEEE-Format im generellen Fall

Bisher: wir befinden uns auf einen 32 bit System, deshalb wird ein float mit 32 bit dargestellt

Es gibt nach IEEE aber auch Darstellungsmöglichkeiten für z.B. 8 bit, 16 bit und 64 bit Gleitkommazahlen.

Je größer die Anzahl der Bit, desto größer ist die Genauigkeit und die Auflösung der Gleitkommazahl.

Hier haben wir ein Beispiel für die 16 bit Darstellung:

$$0\ 1001\ 000\ 1001\ 1011_2$$

Die Berechnung funktioniert analog in allen Fällen. Achtung beim BIAS!

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Multiplikation im IEEE-Format

1: Multiplikation der Mantissen

Mantissen beide im Format 1,...

2: Addition der Biased Exponenten

vgl. hier Potenzrechenregeln $2^a \cdot 2^b = 2^{a+b}$

der Bias kommt als Hilfsgröße in beiden Faktoren vor, deshalb einmal den Bias (hier 63) abziehen.

3: Normalisierung der Mantisse

d.h. auf Form 1,MANTISSE bringen (ggf. Exponenten verschieben)

4: Vorzeichen getrennt behandeln

Aufgabe 3 – Multiplikation

- Multiplikation folgender Gleitkommazahlen:

  $x_1 = 0\ 1001\ 000\ 1001\ 1011_2$

  $x_2 = 1\ 1001\ 010\ 1110\ 1000_2$

1: Multiplikation der Mantissen

Wir müssen beiden Mantissen um die führende 1 des Formats erweitern, um die Korrektheit der Multiplikation zu gewährleisten.

$1{,}1001\ 1011_2 \cdot 1{,}1110\ 1000_2$ (vgl. letzte Übung für ausführlichen Rechenweg)

$= 11{,}0000111101111000_2$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

- Multiplikation folgender Gleitkommazahlen:

  $x_1 = 0\ 1001\ 000\ 1001\ 1011_2$

  $x_2 = 1\ 1001\ 010\ 1110\ 1000_2$

2: Addition der Biased Exponenten

$BIAS = 2^{n-1} - 1 = 2^6 - 1 = 63_{10} = 111111_2$

$1001\ 000_2 + 1001\ 010_2 - BIAS$ (vgl. letzte Übung für ausführlichen Rechenweg)

$= 10010\ 010_2 - 111111_2$ (BIAS einsetzen)

$= 1010011_2$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

- Multiplikation folgender Gleitkommazahlen:

  $x_1 = 0\ 1001\ 000\ 1001\ 1011_2$

  $x_2 = 1\ 1001\ 010\ 1110\ 1000_2$

3: Normalisierung der Mantisse

Mantisse: $11{,}0000111101111000_2$ (Zielformat 1,… + 8 NKS)

$1{,}10000111 \cdot 2^1$

Exponent: 1010011_2

Wir verrechnen jetzt noch den Shift-Faktor der Mantisse mit dem Exponenten.

$1010011_2 + 1_{10} = 1010011_2 + 1_2 = 1010100_2$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

- Multiplikation folgender Gleitkommazahlen:

  $x_1 = 0\ 1001\ 000\ 1001\ 1011_2$

  $x_2 = 1\ 1001\ 010\ 1110\ 1000_2$

4: Vorzeichen getrennt behandeln

0 = +, 1 = - - · + = - Vorzeichen folglich 1 (auch XOR möglich)

5: Zusammensetzen

VB: 1 Exponent: 1010100_2 Mantisse: $1{,}1000\ 0111_2$

Ergebnis: $1\ 1010100\ 1000\ 0111_2$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 3 – Multiplikation

- Multiplikation folgender Gleitkommazahlen:

  $x_1 = 0\ 1001\ 000\ 1001\ 1011_2$

  $x_2 = 1\ 1001\ 010\ 1110\ 1000_2$

Grad des Fehlers

Unser Ergebnis weicht von dem wahren Ergebnis, wenn man x_1 und x_2 in Dezimalzahlen wandelt und diese verrechnet, ab. Dies hängt mit dem Verlust der Genauigkeit der IEEE-Darstellung zusammen. Denn nicht alle Zahlen lassen sich durch gerade Mal 16 Bit wiedergeben!

Wahres Ergebnis: $822 \cdot -3904 = -3\ 209\ 088$

Unser Ergebnis: $1\ 1010100\ 1000\ 0111_2 = -2^{21} \cdot 1\ 135/256 = -3\ 203\ 072$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 4 - Assoziativität

- Die Operationen Addition und Multiplikation auf Operanden in einer Fließkommadarstellung sind normalerweise nicht assoziativ.
- Experimentieren Sie mit einem PC-Tabellenkalkulationsprogramm um dieses zu belegen. Bestimmen Sie dadurch auch die Anzahl der Bits, die zur Speicherung der Mantisse verwendet werden.
- (Beispiel: LibreOffice und Excel liefern bei der Berechnung von: $10^{20} + 17 - 10 - 10^{20} + 130$ als Ergebnis 130 - dies übrigens ohne irgendeine Warnung.)

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 4 - Assoziativität

- Die folgenden Berechnungen sind mit Excel 2013 vorgenommen worden:

	Formel	Ergebnis
A	$10^{20} - 10^{20} + 17 - 10 + 130$	$= 137$
B	$10^{15} - 10^{15} + 17 - 10 + 130$	$= 137$
C	$10^{16} - 10^{16} + 17 - 10 + 130$	$= 137$
D	$10^{20} + 17 - 10 - 10^{20} + 130$	$= 130$
E	$10^{15} + 17 - 10 - 10^{15} + 130$	$= 137$
F	$10^{16} + 17 - 10 - 10^{16} + 130$	$= 136$
G	$10^{20} + 17 + 130 - 10^{20} - 10$	$= -10$
H	$10^{15} + 17 + 130 - 10^{15} - 10$	$= 137$
I	$10^{16} + 17 + 130 - 10^{16} - 10$	$= 136$

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

Aufgabe 4 - Assoziativität

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck

D: $10^{20} + 17 - 10 - 10^{20} + 130 = 130$

Berechnung von D ist falsch, da die Zahl 10^{20} zwar darstellbar ist, aber auf Grund der Inkontinuität der Gleitkommazahlen $10^{20} + 17$ nicht mehr (auf 10^{20} gerundet).

F: $10^{16} + 17 - 10 - 10^{16} + 130 = 136$

Berechnung von F (bzw. auch I) ist falsch, da bei der Darstellung von 10^{16} die niederwertigste 1 von $17_{10} = 10001_2$ bei der Addition auf die 54te Stelle der Mantisse fallen würde (IEEE 754 – 64 bit hat 53 bit für die Mantisse) und somit abgeschnitten wird.

Somit ist klar, welche Genauigkeit (nämlich 64 bit) verwendet wurde.

Vielen Dank für eure geschätzte Aufmerksamkeit!

"En un mot, l'homme doit se créer sa propre essence; c'est en se jetant dans le monde, en y souffrant, en y luttant qu'il se définit peu à peu; et la définition demeure toujours ouverte"

A propos de l'existentialisme - Mise au point. Action, no. 17, 29 décembre 1944.

Friedrich-Alexander Universität Erlangen-Nürnberg
Jan Spieck