

# Lernziele

# Lernziele

- Sie haben ein gutes Verständnis dafür, wie mit dem Computer Zahlen repräsentiert werden.

# Lernziele

- Sie haben ein gutes Verständnis dafür, wie mit dem Computer Zahlen repräsentiert werden.
- Sie können Ganzzahlen in *Binärdarstellung* bringen und damit rechnen.

# Lernziele

- Sie haben ein gutes Verständnis dafür, wie mit dem Computer Zahlen repräsentiert werden.
- Sie können Ganzzahlen in *Binärdarstellung* bringen und damit rechnen.
- Sie verstehen, wie der Wertebereich von Ganzzahlen zustande kommt.

# Lernziele

- Sie haben ein gutes Verständnis dafür, wie mit dem Computer Zahlen repräsentiert werden.
- Sie können Ganzzahlen in *Binärdarstellung* bringen und damit rechnen.
- Sie verstehen, wie der Wertebereich von Ganzzahlen zustande kommt.
- Sie können qualitativ die Repräsentation von Fließkommazahlen beschreiben.

# Lernziele

- Sie haben ein gutes Verständnis dafür, wie mit dem Computer Zahlen repräsentiert werden.
- Sie können Ganzzahlen in *Binärdarstellung* bringen und damit rechnen.
- Sie verstehen, wie der Wertebereich von Ganzzahlen zustande kommt.
- Sie können qualitativ die Repräsentation von Fließkommazahlen beschreiben.
- Sie kennen die drei *Fließkomma-Richtlinien*.

# Lernziele

- Sie haben ein gutes Verständnis dafür, wie mit dem Computer Zahlen repräsentiert werden.
- Sie können Ganzzahlen in *Binärdarstellung* bringen und damit rechnen.
- Sie verstehen, wie der Wertebereich von Ganzzahlen zustande kommt.
- Sie können qualitativ die Repräsentation von Fließkommazahlen beschreiben.
- Sie kennen die drei *Fliesskomma-Richtlinien*.
- Sie können mit Wahrheitswerten und *boolschen Ausdrücken* in Java umgehen.

## 4. Zahlendarstellungen

Wertebereich der Typen `int`, `float` und `double` Gemischte  
Ausdrücke und Konversionen; Löcher im Wertebereich;  
Fließkomma-Richtlinien;

# Binäre Zahlendarstellungen

Binäre Darstellung ("Bits" aus  $\{0, 1\}$ )

$$b_n b_{n-1} \dots b_1 b_0$$

entspricht der Zahl  $b_n \cdot 2^n + \dots + b_1 \cdot 2^1 + b_0 \cdot 2^0$

# Binäre Zahlendarstellungen

Binäre Darstellung ("Bits" aus  $\{0, 1\}$ )

$$b_n b_{n-1} \dots b_1 b_0$$

entspricht der Zahl  $b_n \cdot 2^n + \dots + b_1 \cdot 2 + b_0$

# Binäre Zahlendarstellungen

Binäre Darstellung ("Bits" aus  $\{0, 1\}$ )

$$b_n b_{n-1} \dots b_1 b_0$$

entspricht der Zahl  $b_n \cdot 2^n + \dots + b_1 \cdot 2 + b_0$

Beispiel: 101011

# Binäre Zahlendarstellungen

Binäre Darstellung ("Bits" aus  $\{0, 1\}$ )

$$b_n b_{n-1} \dots b_1 b_0$$

entspricht der Zahl  $b_n \cdot 2^n + \dots + b_1 \cdot 2 + b_0$

Beispiel: 101011 entspricht  $32+8+2+1$ .

# Binäre Zahlendarstellungen

Binäre Darstellung ("Bits" aus  $\{0, 1\}$ )

$$b_n b_{n-1} \dots b_1 b_0$$

entspricht der Zahl  $b_n \cdot 2^n + \dots + b_1 \cdot 2 + b_0$

Beispiel: 101011 entspricht 43.

# Binäre Zahlendarstellungen

Binäre Darstellung ("Bits" aus  $\{0, 1\}$ )

$$b_n b_{n-1} \dots b_1 b_0$$

entspricht der Zahl  $b_n \cdot 2^n + \dots + b_1 \cdot 2 + b_0$

Beispiel: **101011** entspricht 43.

*Niedrigstes Bit, Least Significant Bit (LSB)*

*Höchstes Bit, Most Significant Bit (MSB)*

# Binäre Zahlen: Zahlen der Computer?

## Wahrheit: Computer rechnen mit Binärzahlen.

NEUE ZÜRCHER ZEITUNG

# TECHNIK

Mittwoch, 30. August 1950 Blatt 13  
Mittagsausgabe Nr. 1796 (50)

### Das programmgesteuerte Rechengertät an der Eidgenössischen Technischen Hochschule in Zürich

Die Entwicklung programmgesteuerter Rechenmaschinen in den Vereinigten Staaten von Amerika sind in den Artikeln „Elektronische Rechenmaschinen“ (vgl. Nr. 2149 der „N. Z. Z.“ vom 13. Oktober 1948) und „Die neueste elektronische Rechenmaschine“ (vgl. Nr. 872 der „N. Z. Z.“ vom 26. April 1950) behandelt. Nachstehend soll von einem Gerät deutscher Herkunft — Zuse K-6, Neubibchen — die Rolle sein, welches im Juli dieses Jahres am Institut für angewandte Mathematik der Eidgenössischen Technischen Hochschule in Zürich, das unter der Leitung von Prof. Dr. F. Siefel steht, in Betrieb genommen wurde. Dient ist dieses Institut in der Lage, dem in der Schweiz immer stärker werdenden Bedarf nach einer leistungsfähigen Zentraltabelle für numerische Rechnungen wenigstens teilweise gerecht zu werden. Bereits sind einige mathematische Probleme behandelt worden, und die Erfüllung vieler anderer Aufgaben ist vorbereitet.

#### Merkmale des Gerätes

Das Gerät ist ein Glied in dem progressiven Entwicklungprogramm des Ingenieurs Konrad Zuse; es wurde im Auftrag des Instituts für angewandte Mathematik der E. T. H. unter Berücksichtigung von dessen Wünschen und Ideen von Zuse als „Modell E 4“ konstruiert. Die ursprüngliche Entwicklung in Deutschland erfolgte in den Kriegsjahren und verlief völlig unabhängig von den Untersuchungen des Vereinigten Staates. Es ist überaus interessant festzustellen, wie für die meisten wichtigen funktionalen Probleme bedenkenswerte genau dieselbe Lösung gefunden wurde, wie aber andererseits gewissen Fragen sekundärer Wichtigkeit eine ganz unterschiedliche Bedeutung beigegeben wurde.

Eine kurze technische Charakterisierung lautet wie folgt: Elektronenmechanisches Gerät mit 2200 Relais, 21 Schrittschaltern und einem Speicher für 64 Zahlen, welcher mit neuartigen, mechanischen Schaltgliedern arbeitet; Vervollständigung des Dualsystems und der logarithmischen Darstellung; Multiplikationszeit 2,5 Sekunden; Programmsteuerung mit Hilfe zweier Lochstreifen, auf die willkürlich umgeschaltet werden kann; Eingabe von Zahlen durch eine Tastatur oder durch einen Lochstreifen; Abgabe der Resultate durch Lampenfeld, Lochstreifen oder Druckwerk.

#### Das duale Zahlensystem

Allgemein wird programmgesteuertes Rechengertät häufig das duale Zahlensystem zugrunde gelegt, welches nur die zwei Zahlensymbole 0 und 1 verwendet, während das bekannte Dezimalsystem

lesen wir eine Dezimalzahl von rechts nach links, so erhält sich das Gewicht von Stelle zu Stelle um den Faktor 10. Im Dualsystem ist nun einfach dieser Faktor 10 durch 2 zu ersetzen. Also bedeutet die (zunehmend duale) Zahl abelsch den Ausdruck:

$$a \cdot 2^3 + b \cdot 2^2 + c \cdot 2^1 + d \cdot 2^0 + e \cdot 2^{-1} + f \cdot 2^{-2} + g \cdot 2^{-3}$$

Die Zahl 1 wird in beiden Systemen gleich dargestellt. Im jedoch duale von dezimalen Zahlen deutlich zu trennen, schreiben wir die duale 1 als  $1_2$ . — Dagegen weicht schon die 2 ab, indem sie dual  $10_2$  lautet, denn dies bedeutet  $1 \cdot 2^1 + 0 \cdot 2^0 = 2$ . Wenn einer Zahl (ohne Stellen nach dem Komma) bereits eine Null zugefügt wird, so vergrößert sie sich um den Faktor 2 (und sieht, wie im Dualsystem, um den Faktor 10). Auf diese Weise kann aus  $10_2 = 2$  auf einfachste Weise gebildet werden:  $100_2 = 4$ ,  $1000_2 = 8$ ,  $10000_2 = 16$ , usw.

Die Dualzahl  $10101_2$  bedeutet also nach:

$$1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 21$$

Ganz analog sind etwaige Stellen nach dem Komma zu interpretieren; so wird  $1,011_2$  wie folgt interpretiert:

$$1 \cdot 2^0 + 0 \cdot 2^{-1} + 1 \cdot 2^{-2} + 1 \cdot 2^{-3} = 1 + \frac{1}{4} + \frac{1}{8} = 1,375$$

Der große Vorteil, der das Dualsystem für Rechenautomaten so geeignet macht, nämlich die Reduktion der Anzahl der verwendeten Symbole auf nur zwei, wird allerdings durch einen Nachteil erkauft: Es braucht mehr Stellen, um eine bestimmte Zahl darzustellen. Die einstellige Zahl 8

Änderung des Maßstabes durchgeführt werden können.

Die beschriebene Darstellung bringt eine gewisse Komplikation der Rechenoperationen mit sich. So müssen vor einer Addition die beiden Summanden zunächst so verschoben werden, daß ihre Kommata untereinander zu liegen kommen, was am Hand eines Beispiels erläutert werden soll. Damit der Leser nicht durch das ausgeführte duale Zahlensystem verirrt wird, ist das Beispiel im Dualsystem durchgeführt; doch wird daran erinnert, daß das Gerät in Wirklichkeit mit dualen Zahlen rechnet.

Es soll also etwa addiert werden:  $2,345678 \times 10^3 + 9,876543 \times 10^1$  (Man beachte, daß die gesamte Zahl stets zwischen 1 und 10 liegt, also das Komma nachher ersten Stelle ist). Nun müssen die beiden Summanden „ausgerichtet“ werden, d. h. die beiden Exponenten sind einander gleich zu machen, um eine einheitliche Exponent den Wert des größeren, also 2. Die Zahlen unten nun, richtig untereinander geschrieben und addiert, wie folgt:

$$\begin{array}{r} 2,345678 \times 10^3 \\ 0,987654 \times 10^2 \\ \hline 2,535504 \times 10^3 \end{array}$$

Es ist ersichtlich, daß bei der kleineren der beiden Zahlen rechts einige Stellen abgeschrieben werden mußten; denn wenn die Summanden siebenstellig gegeben waren, so soll auch das Resultat nicht mehr als sieben Stellen enthalten.



Abb. 2. Der Schalter bei der Festlegung eines Rechnerplatzes. Die Ableser für den Lochstreifen sind deutlich sichtbar.

Befehle können „bellend“ gegeben werden, d. h. ihre Ausführung wird von der Natur eines errechneten Resultates abhängig gemacht. Erst danach werden die anforderungsgemäßen Resultate

# Binäre Zahlen: Zahlen der Computer?

Klischee: Computer reden 0/1-Kauderwelsch.

## Biooino Bioiioio Bioiioio

01001110 01011010 01011010

Freitag, 8. Juni 2012 · Nr. 131 · 233. Jhg.

01001010 01010110 01001101

www.nzz.ch · Fr. 4.00 · € 3.50



01000010 01100101  
01110010 01101001

01100011 01101000 01110100 01100101

00100000 11111100 01100010  
01100101 01110010 00100000  
01101110 01100101 01110101 011-  
00101 01110011 00100000 010-  
01101 01100001 01110011

01120011 01100001

01100011 01100001 01100000 01100000 011-  
00001 01101110 00100000 01010011 01111-  
001 01110010 01100001 01100101 01101110  
0000101 00001010 00001101 00000010  
0000101 0101110 0100111 00101101 010-  
00010 0100001 0001110 01100010 01100001  
0110011 01101000 01110010 01100101 011-  
10010 00100000 01110110 0001111 0001101  
0010000 01010011 01100001 01100000 011-  
00001 01110101 01100000 01101100 01100-  
001 01101010 01110000 01010000

01100110 01100101

01100010 01101110 01100111 01100001 011-  
0000 01100001 01101100 01101000 010-  
0001 01101110 0000101 00100010 0000101  
00001000 00001100 01100001 01110101 011-  
10000 0000000 00000010 01100101 0111-  
0010 01101001 01100001 01100000 0110000

01100101 01110001 00100000 01001101 011-  
00001 01110001 01110011 01100000 01101-  
011 01100101 01110010 01010000 01110011  
01110000 01100001 01110010 01110000 011-  
0010 0000101 01100110 0110101 0110110  
01100100 01100101 01110110 0010110 001-  
00000 011000100 01101001 01100101 001-  
00000 01101000 01100001 01100110 01100001  
01100101 01110010 01110011 01100101  
01100111 00100000 01100101 01100001 011-  
00010 01101000 01110010 01100101 00000-  
000 01110000 01101110 01100111 01100000  
01100110 01100001 01101110 011110 0111-  
0000 01100101 00100000 10001011 01010100  
01100101 01110010 01110010 01100111 011-  
10000 0110101 01110011 01110010 01100000  
01100110 10110111 00100000 01100100 011-  
00001 01100110 11111100 01110010

00100000 01110110

01100101 01110010 01100000 01101110 0111-  
0000 01110111 01110010 01110010 01100100  
01101100 01100001 01100101 01100000 001-  
0110 0000101 00001010 0000101 000-  
0010 01000010 11111100 01110010 011001-  
11 00100000 01100000 01100101 01100111  
01100001 01100100 01101111 01100110 011-  
0010 00101100 00100000 00000010 0110000

01000110 01101100 11111100

01100011 01101000 01110100 01101100 01100001 01101110 01100011 01110001 00100101 00101100 01100101 01101110 01100100 0000-  
0000 01101000 01101110 00100000 00100000 01100001 01110100 01110010 01100001 01110011 00001011 00000100 00000100 00101100 01100101  
01100101 00100000 01100111 0110010 01100100 01100100 01100001 01100000 01101001 01100101 01100001

01201000

## Definition: *Wertebereich*

*Bei numerischen Typen gibt der Wertebereich an, welches Zahlenintervall abgedeckt werden kann.*

Buch, auf Seite 24

# Wertebereich des Typs int

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("Minimum int value is ");  
        Out.println(Integer.MIN_VALUE);  
        Out.print("Maximum int value is ");  
        Out.println(Integer.MAX_VALUE);  
    }  
}
```

# Wertebereich des Typs int

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("Minimum int value is ");  
        Out.println(Integer.MIN_VALUE);  
        Out.print("Maximum int value is ");  
        Out.println(Integer.MAX_VALUE);  
    }  
}
```

Minimum int value is -2147483648.

Maximum int value is 2147483647.

# Wertebereich des Typs int

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("Minimum int value is ");  
        Out.println(Integer.MIN_VALUE);  
        Out.print("Maximum int value is ");  
        Out.println(Integer.MAX_VALUE);  
    }  
}
```

```
Minimum int value is -2147483648.  
Maximum int value is 2147483647.
```

Woher kommen diese Zahlen?

# Wertebereich des Typs `int`

Repräsentation mit 32 Bits. Wertebereich

$$\{-2^{31}, \dots, -1, 0, 1, \dots, 2^{31} - 2, 2^{31} - 1\}$$

# Wertebereich des Typs `int`

Repräsentation mit 32 Bits. Wertebereich

$$\{-2^{31}, \dots, -1, 0, 1, \dots, 2^{31} - 2, 2^{31} - 1\}$$

Woher kommt gerade diese Aufteilung?

# Rechnen mit Binärzahlen (4 Stellen)

Einfache Addition

$$\begin{array}{r} 2 \\ +3 \\ \hline 5 \end{array}$$

$$\begin{array}{r} 0010 \\ +0011 \\ \hline 0101 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

Einfache Subtraktion

$$\begin{array}{r} 5 \\ -3 \\ \hline 2 \end{array}$$

$$\begin{array}{r} 0101 \\ -0011 \\ \hline 0010 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

Addition mit Überlauf

$$\begin{array}{r} 7 \\ +9 \\ \hline 16 \end{array}$$

$$\begin{array}{r} 0111 \\ +1001 \\ \hline (1)0000 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

Negative Zahlen?

$$\begin{array}{r} 5 \\ +(-5) \\ \hline 0 \end{array}$$

$$\begin{array}{r} 0101 \\ \quad ??? \\ \hline (1)0000 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

Einfacher: -1

$$\begin{array}{r} 1 \\ +(-1) \\ \hline 0 \end{array}$$

$$\begin{array}{r} 0001 \\ 1111 \\ \hline (1)0000 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

Nutzen das aus:

$$\begin{array}{r} 3 \\ +? \\ \hline -1 \end{array}$$

$$\begin{array}{r} 0011 \\ +???? \\ \hline 1111 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

Invertieren!

$$\begin{array}{r} 3 \\ +(-4) \\ \hline -1 \end{array}$$

$$\begin{array}{r} 0011 \\ +1100 \\ \hline 1111 \hat{=} 2^B - 1 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

$$\begin{array}{r} a \\ +(-a - 1) \\ \hline -1 \end{array}$$

$$\begin{array}{r} a \\ \bar{a} \\ \hline 1111 \hat{=} 2^B - 1 \end{array}$$

# Rechnen mit Binärzahlen (4 Stellen)

- Negation: Inversion und Addition von 1

$$-a \hat{=} \bar{a} + 1$$

# Rechnen mit Binärzahlen (4 Stellen)

- Wrap-around Semantik (Rechnen modulo  $2^B$ )

$$-a \hat{=} 2^B - a$$

# Warum das funktioniert

Modulo-Arithmetik: Rechnen im Kreis<sup>3</sup>



$11 \equiv 23 \equiv -1 \equiv \dots \pmod{12}$

$4 \equiv 16 \equiv \dots \pmod{12}$

$3 \equiv 15 \equiv \dots \pmod{12}$

<sup>3</sup>Die Arithmetik funktioniert auch mit Dezimalzahlen (und auch für die Multiplikation)

# Negative Zahlen (3 Stellen)

	$a$	$-a$
0	000	
1	001	
2	010	
3	011	
4	100	
5	101	
6	110	
7	111	

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001		
2	010		
3	011		
4	100		
5	101		
6	110		
7	111		

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001	<b>111</b>	-1
2	010		
3	011		
4	100		
5	101		
6	110		
7	<b>111</b>		

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001	111	-1
2	010	<b>110</b>	-2
3	011		
4	100		
5	101		
6	<b>110</b>		
7	111		

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001	111	-1
2	010	110	-2
3	011	<b>101</b>	-3
4	100		
5	<b>101</b>		
6	110		
7	111		

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001	111	-1
2	010	110	-2
3	011	101	-3
4	<b>100</b>	<b>100</b>	-4
5	101		
6	110		
7	111		

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001	111	-1
2	010	110	-2
3	011	101	-3
4	100	100	-4
5	101		
6	110		
7	111		

# Negative Zahlen (3 Stellen)

	$a$	$-a$	
0	000	000	0
1	001	111	-1
2	010	110	-2
3	011	101	-3
4	100	100	-4
5	101		
6	110		
7	111		

Das höchste Bit entscheidet über das Vorzeichen.

# Überlauf und Unterlauf

- Arithmetische Operationen (+, -, \*) können aus dem Wertebereich herausführen.
- Ergebnisse können inkorrekt sein.

`power8`:  $15^8 = -1732076671$

`power20`:  $3^{20} = -808182895$

- Es gibt *keine Fehlermeldung!*

## Definition: *Fliesskommazahlen*

*Fliesskommazahlen stellen Zahlen aus  $\mathbb{R}$  dar mit einer festen Anzahl signifikanter Stellen, multipliziert mit einer Zehnerpotenz. (Basis 10).*

Buch, auf Seite 67

# „Richtig Rechnen“

```
public class Main {  
  
    public static void main(String[] args) {  
        Out.print("Celsius: ");  
        int celsius = In.readInt();  
        int fahrenheit = 9 * celsius / 5 + 32;  
        Out.print(celsius + " degrees Celsius are ");  
        Out.println(fahrenheit + " degrees Fahrenheit");  
    }  
}
```

28 degrees Celsius are 82 degrees Fahrenheit.

# „Richtig Rechnen“

```
public class Main {  
  
    public static void main(String[] args) {  
        Out.print("Celsius: ");  
        int celsius = In.readInt();  
        int fahrenheit = 9 * celsius / 5 + 32;  
        Out.print(celsius + " degrees Celsius are ");  
        Out.println(fahrenheit + " degrees Fahrenheit");  
    }  
}
```

28 degrees Celsius are 82 degrees Fahrenheit.

↑  
richtig wäre 82.4

# „Richtig Rechnen“

```
public class Main {  
  
    public static void main(String[] args) {  
        Out.print("Celsius: ");  
        float celsius = In.readInt();  
        float fahrenheit = 9 * celsius / 5 + 32;  
        Out.print(celsius + " degrees Celsius are ");  
        Out.println(fahrenheit + " degrees Fahrenheit");  
    }  
}
```

28 degrees Celsius are 82.4 degrees Fahrenheit.

# Typen `float` und `double`

- sind die fundamentalen Typen für Fließkommazahlen
- approximieren den Körper der reellen Zahlen  $(\mathbb{R}, +, \times)$  in der Mathematik
- haben grossen Wertebereich, ausreichend für viele Anwendungen (`double` hat mehr Stellen als `float`)
- sind auf vielen Rechnern sehr schnell

# Typen `float` und `double`

- sind die fundamentalen Typen für Fließkommazahlen
- approximieren den Körper der reellen Zahlen  $(\mathbb{R}, +, \times)$  in der Mathematik
- haben grossen Wertebereich, ausreichend für viele Anwendungen (`double` hat mehr Stellen als `float`)
- sind auf vielen Rechnern sehr schnell

# Typen `float` und `double`

- sind die fundamentalen Typen für Fließkommazahlen
- approximieren den Körper der reellen Zahlen  $(\mathbb{R}, +, \times)$  in der Mathematik
- haben grossen Wertebereich, ausreichend für viele Anwendungen (`double` hat mehr Stellen als `float`)
- sind auf vielen Rechnern sehr schnell

# Typen `float` und `double`

- sind die fundamentalen Typen für Fließkommazahlen
- approximieren den Körper der reellen Zahlen  $(\mathbb{R}, +, \times)$  in der Mathematik
- haben grossen Wertebereich, ausreichend für viele Anwendungen (`double` hat mehr Stellen als `float`)
- sind auf vielen Rechnern sehr schnell

# Fixkommazahlen

- feste Anzahl Vorkommastellen (z.B. 7)
- feste Anzahl Nachkommastellen (z.B. 3)

# Fixkommazahlen

- feste Anzahl Vorkommastellen (z.B. 7)
- feste Anzahl Nachkommastellen (z.B. 3)

$$82.4 = 0000082.400$$

# Fixkommazahlen

- feste Anzahl Vorkommastellen (z.B. 7)
- feste Anzahl Nachkommastellen (z.B. 3)

$$82.4 = 0000082.400$$

## Nachteile

- Wertebereich wird *noch* kleiner als bei ganzen Zahlen.

# Fixkommazahlen

- feste Anzahl Vorkommastellen (z.B. 7)
- feste Anzahl Nachkommastellen (z.B. 3)

0.0824 = 0000000.082 ← dritte Stelle abgeschnitten

## Nachteile

- Repräsentierbarkeit hängt von der Stelle des Kommas ab.

# Fliesskommazahlen

- feste Anzahl signifikante Stellen (z.B. 10)
- plus Position des Kommas

$$82.4 = 824 \cdot 10^{-1}$$

$$0.0824 = 824 \cdot 10^{-4}$$

- Zahl ist  $\text{Signifikand} \times 10^{\text{Exponent}}$

# Fliesskommazahlen

- feste Anzahl signifikante Stellen (z.B. 10)
- plus Position des Kommas

$$82.4 = 824 \cdot 10^{-1}$$

$$0.0824 = 824 \cdot 10^{-4}$$

- Zahl ist  $\text{Signifikand} \times 10^{\text{Exponent}}$

# Fliesskommazahlen

- feste Anzahl signifikante Stellen (z.B. 10)
- plus Position des Kommas

$$82.4 = 824 \cdot 10^{-1}$$

$$0.0824 = 824 \cdot 10^{-4}$$

- Zahl ist  $\text{Signifikand} \times 10^{\text{Exponent}}$

# Wertebereich

Ganzzahlige Typen:

- Über- und Unterlauf häufig, aber ...

# Wertebereich

Ganzzahlige Typen:

- Über- und Unterlauf häufig, aber ...
- Wertebereich ist zusammenhängend (keine „Löcher“):  $\mathbb{Z}$  ist „diskret“.

# Wertebereich

Fliesskommatypen:

- Über- und Unterlauf selten, aber ...

# Wertebereich

Fließkommatypen:

- Über- und Unterlauf selten, aber ...
- es gibt Löcher:  $\mathbb{R}$  ist „kontinuierlich“.

# Löcher im Wertebereich

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("First number =? ");  
        float n1 = In.readFloat();  
  
        Out.print("Second number =? ");  
        float n2 = In.readFloat();  
  
        Out.print("Their difference =? ");  
        float d = In.readFloat();  
  
        Out.print("computed difference - input difference = ");  
        Out.println(n1-n2-d);  
    }  
}
```

# Löcher im Wertebereich

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("First number =? ");      Eingabe 1.5  
        float n1 = In.readFloat();  
  
        Out.print("Second number =? ");    Eingabe 1.0  
        float n2 = In.readFloat();  
  
        Out.print("Their difference =? ");  Eingabe 0.5  
        float d = In.readFloat();  
  
        Out.print("computed difference – input difference = ");  
        Out.println(n1–n2–d);  
    }  
}
```

# Löcher im Wertebereich

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("First number =? ");      Eingabe 1.5  
        float n1 = In.readFloat();  
  
        Out.print("Second number =? ");     Eingabe 1.0  
        float n2 = In.readFloat();  
  
        Out.print("Their difference =? ");  Eingabe 0.5  
        float d = In.readFloat();  
  
        Out.print("computed difference – input difference = ");  
        Out.println(n1–n2–d);  
    }                                       Ausgabe 0  
}
```

# Löcher im Wertebereich

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("First number =? ");      Eingabe 1.1  
        float n1 = In.readFloat();  
  
        Out.print("Second number =? ");    Eingabe 1.0  
        float n2 = In.readFloat();  
  
        Out.print("Their difference =? ");  Eingabe 0.1  
        float d = In.readFloat();  
  
        Out.print("computed difference – input difference = ");  
        Out.println(n1-n2-d);  
    }  
}
```

# Löcher im Wertebereich

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("First number =? ");  
        float n1 = In.readFloat();  
  
        Out.print("Second number =? ");  
        float n2 = In.readFloat();  
  
        Out.print("Their difference =? ");  
        float d = In.readFloat();  
  
        Out.print("computed difference – input difference = ");  
        Out.println(n1–n2–d);  
    }  
}
```

Eingabe 1.1

Eingabe 1.0

Eingabe 0.1

Ausgabe 2.2351742E-8

# Löcher im Wertebereich

```
public class Main {  
    public static void main(String[] args) {  
        Out.print("First number =? ");  
        float n1 = In.readFloat();  
  
        Out.print("Second number =? ");  
        float n2 = In.readFloat();  
  
        Out.print("Their difference =? ");  
        float d = In.readFloat();  
  
        Out.print("computed difference – input difference = ");  
        Out.println(n1–n2–d);  
    }  
}
```

Eingabe 1.1

Eingabe 1.0

Eingabe 0.1

Ausgabe 2.2351742E-8

Ja was ist denn hier los?

## Regel 1

Teste keine gerundeten Fließkommazahlen auf Gleichheit!

```
for (float i = 0.1; i != 1.0; i += 0.1){  
    Out.println(i);  
}
```

Mehr dazu nächstes mal!

## Regel 1

Teste keine gerundeten Fließkommazahlen auf Gleichheit!

```
for (float i = 0.1; i != 1.0; i += 0.1){  
    Out.println(i);  
}
```

Mehr dazu nächstes mal!

## Regel 1

Teste keine gerundeten Fließkommazahlen auf Gleichheit!

```
for (float i = 0.1; i != 1.0; i += 0.1){  
    Out.println(i);  
}
```

Endlosschleife, weil i niemals exakt 1 ist!

Mehr dazu nächstes mal!

### Regel 2

Addiere keine zwei Zahlen sehr unterschiedlicher Grösse!

### Regel 2

Addiere keine zwei Zahlen sehr unterschiedlicher Grösse!

$$\begin{array}{r} 1.000 \cdot 10^5 \\ +1.000 \cdot 10^0 \end{array}$$

### Regel 2

Addiere keine zwei Zahlen sehr unterschiedlicher Grösse!

$$\begin{aligned} & 1.000 \cdot 10^5 \\ & + 1.000 \cdot 10^0 \\ & = 1.00001 \cdot 10^5 \end{aligned}$$

### Regel 2

Addiere keine zwei Zahlen sehr unterschiedlicher Grösse!

$$\begin{aligned} & 1.000 \cdot 10^5 \\ & + 1.000 \cdot 10^0 \\ & = 1.00001 \cdot 10^5 \\ & \text{"="} 1.000 \cdot 10^5 \quad (\text{Rundung auf 4 Stellen}) \end{aligned}$$

### Regel 2

Addiere keine zwei Zahlen sehr unterschiedlicher Grösse!

$$\begin{aligned} & 1.000 \cdot 10^5 \\ & + 1.000 \cdot 10^0 \\ & = 1.00001 \cdot 10^5 \\ & \text{"="} 1.000 \cdot 10^5 \quad (\text{Rundung auf 4 Stellen}) \end{aligned}$$

Addition von 1 hat keinen Effekt!

## Regel 3

Subtrahiere keine zwei Zahlen sehr ähnlicher Grösse!

Auslöschungsproblematik (ohne weitere Erklärung).

# 5. Wahrheitswerte

Boolesche Funktionen; der Typ `boolean`; logische und relationale Operatoren; Kurzschlussauswertung

# Wo wollen wir hin?

```
int a = In.readInt();  
if (a % 2 == 0){  
    Out.println("even");  
} else {  
    Out.println("odd");  
}
```

# Wo wollen wir hin?

```
int a = In.readInt();  
if (a % 2 == 0){  
    Out.println("even");  
} else {  
    Out.println("odd");  
}
```

Verhalten hängt ab vom Wert eines *Booleschen Ausdrucks*

# Wo wollen wir hin?

```
int a = In.readInt();  
if (a % 2 == 0){  
    Out.println("even");  
} else {  
    Out.println("odd");  
}
```

Verhalten hängt ab vom Wert eines *Booleschen Ausdrucks*

# Boolesche Werte in der Mathematik

Boolesche Ausdrücke können zwei mögliche Werte annehmen:

*F* oder *T*

# Boolesche Werte in der Mathematik

Boolesche Ausdrücke können zwei mögliche Werte annehmen:

$F$  oder  $T$

- $F$  entspricht „*falsch*“
- $T$  entspricht „*wahr*“

# Der Typ `boolean` in Java

- Repräsentiert *Wahrheitswerte*

# Der Typ `boolean` in Java

- Repräsentiert *Wahrheitswerte*
- Literale `false` und `true`

# Der Typ `boolean` in Java

- Repräsentiert *Wahrheitswerte*
- Literale `false` und `true`
- Wertebereich {*false*, *true*}

```
boolean b = true; // Variable mit Wert true (wahr)
```

# Relationale Operatoren

$a < b$  (kleiner als)

Zahlentyp  $\times$  Zahlentyp  $\rightarrow$  boolean

# Relationale Operatoren

$a < b$  (kleiner als)

```
boolean b = (1 < 3); // b =
```

# Relationale Operatoren

$a < b$  (kleiner als)

```
boolean b = (1 < 3); // b = true (wahr)
```

# Relationale Operatoren

`a >= b` (grösser gleich)

```
int a = 0;  
boolean b = (a >= 3); // b =
```

# Relationale Operatoren

`a >= b` (grösser gleich)

```
int a = 0;  
boolean b = (a >= 3); // b = false (falsch)
```

# Relationale Operatoren

a == b (gleich)

```
int a = 4;  
boolean b = (a % 3 == 1); // b =
```

# Relationale Operatoren

a == b (gleich)

```
int a = 4;  
boolean b = (a % 3 == 1); // b = true (wahr)
```

# Relationale Operatoren

`a != b` (ungleich)

```
int a = 1;  
boolean b = (a != 2*a-1); // b =
```

# Relationale Operatoren

`a != b` (ungleich)

```
int a = 1;  
boolean b = (a != 2*a-1); // b = false (falsch)
```

# Boolesche Funktionen in der Mathematik

- Boolesche Funktion

$$f : \{F, T\}^2 \rightarrow \{F, T\}$$

- $F$  entspricht „falsch“.
- $T$  entspricht „wahr“.

- “Logisches Und”

$$f : \{F, T\}^2 \rightarrow \{F, T\}$$

- $F$  entspricht „falsch”.
- $T$  entspricht „wahr”.

$x$	$y$	$\text{AND}(x, y)$
$F$	$F$	$F$
$F$	$T$	$F$
$T$	$F$	$F$
$T$	$T$	$T$

# Logischer Operator &&

`a && b` (logisches Und)

`boolean × boolean → boolean`

# Logischer Operator &&

a && b      (logisches Und)

```
int n = -1;  
int p = 3;  
boolean b = (n < 0) && (0 < p); //
```

# Logischer Operator &&

a && b      (logisches Und)

```
int n = -1;  
int p = 3;  
boolean b = (n < 0) && (0 < p); // b = true (wahr)
```

- “Logisches Oder”

$$f : \{F, T\}^2 \rightarrow \{F, T\}$$

- $F$  entspricht „falsch“.
- $T$  entspricht „wahr“.

$x$	$y$	$\text{OR}(x, y)$
$F$	$F$	$F$
$F$	$T$	$T$
$T$	$F$	$T$
$T$	$T$	$T$

# Logischer Operator ||

`a || b` (logisches Oder)

`boolean × boolean → boolean`

# Logischer Operator ||

a || b      (logisches Oder)

```
int n = 1;  
int p = 0;  
boolean b = (n < 0) || (0 < p); //
```

# Logischer Operator ||

a || b      (logisches Oder)

```
int n = 1;  
int p = 0;  
boolean b = (n < 0) || (0 < p); // b = false (falsch)
```

- “Logisches Nicht”

$$f : \{F, T\} \rightarrow \{F, T\}$$

- $F$  entspricht „falsch“.
- $T$  entspricht „wahr“.

$x$	NOT( $x$ )
$F$	$T$
$T$	$F$

# Logischer Operator !

`!b` (logisches Nicht)

`boolean → boolean`

# Logischer Operator !

!b      (logisches Nicht)

```
int n = 1;  
boolean b = !(n < 0); //
```

# Logischer Operator !

!b      (logisches Nicht)

```
int n = 1;  
boolean b = !(n < 0); // b = true (wahr)
```

# Präzedenzen

`!b && a`

# Präzedenzen

`!b && a`  
⇕  
`(!b) && a`

# Präzedenzen

a && b || c && d

# Präzedenzen

a && b || c && d  
⇕  
(a && b) || (c && d)

# Präzedenzen

a || b && c || d

# Präzedenzen

$a \ || \ b \ \&\& \ c \ || \ d$   
 $\Updownarrow$   
 $a \ || \ (b \ \&\& \ c) \ || \ d$

# Präzedenzen

`7 + x < y && y != 3 * z || ! b`

# Präzedenzen

*Der unäre logische* Operator !

bindet stärker als

```
7 + x < y && y != 3 * z || (!b)
```

# Präzedenzen

*Der unäre logische* Operator !

bindet stärker als

*binäre arithmetische* Operatoren. Diese

binden stärker als

```
(7 + x) < y && y != (3 * z) || (!b)
```

# Präzedenzen

*Der unäre logische* Operator !

bindet stärker als

*binäre arithmetische* Operatoren. Diese

binden stärker als

*relationale* Operatoren,

und diese binden stärker als

```
((7 + x) < y) && (y != (3 * z)) || (!b)
```

# Präzedenzen

*Der unäre logische* Operator !

bindet stärker als

*binäre arithmetische* Operatoren. Diese

binden stärker als

*relationale* Operatoren,

und diese binden stärker als

*binäre logische* Operatoren.

```
((7 + x) < y) && (y != (3 * z)) || (!b)
```

Einige Klammern auf den vorher gezeigten Folien waren unnötig.

# DeMorgansche Regeln

- $!(a \ \&\& \ b) == (!a \ || \ !b)$

# DeMorgansche Regeln

■  $!(a \ \&\& \ b) == (!a \ || \ !b)$

! (reich *und* schön) == (arm *oder* hässlich)

# DeMorgansche Regeln

■  $!(a \ \&\& \ b) == (!a \ || \ !b)$

■  $!(a \ || \ b) == (!a \ \&\& \ !b)$

! (reich *und* schön) == (arm *oder* hässlich)

# Anwendung: Entweder ... Oder (XOR)

`(x || y) && !(x && y)`

# Anwendung: Entweder ... Oder (XOR)

`(x || y) && !(x && y)` x oder y, und nicht beide

# Anwendung: Entweder ... Oder (XOR)

`(x || y) && !(x && y)` x oder y, und nicht beide

`(x || y) && (!x || !y)`

# Anwendung: Entweder ... Oder (XOR)

`(x || y) && !(x && y)` x oder y, und nicht beide

`(x || y) && (!x || !y)` x oder y, und eines nicht

# Anwendung: Entweder ... Oder (XOR)

`(x || y) && !(x && y)` x oder y, und nicht beide

`(x || y) && (!x || !y)` x oder y, und eines nicht

`!(!x && !y) && !(x && y)`

# Anwendung: Entweder ... Oder (XOR)

$(x \ || \ y) \ \ \ \ \ \&\& \ ! (x \ \&\& \ y)$      $x$  oder  $y$ , und nicht beide

$(x \ || \ y) \ \ \ \ \ \&\& \ (!x \ || \ !y)$      $x$  oder  $y$ , und eines nicht

$!(\ !x \ \&\& \ !y) \ \ \ \ \ \&\& \ ! (x \ \&\& \ y)$     nicht keines, und nicht beide

# Anwendung: Entweder ... Oder (XOR)

`(x || y) && !(x && y)` x oder y, und nicht beide

`(x || y) && (!x || !y)` x oder y, und eines nicht

`!(!x && !y) && !(x && y)` nicht keines, und nicht beide

`!(!x && !y || x && y)`

# Anwendung: Entweder ... Oder (XOR)

$(x \ || \ y) \ \ \ \ \ \&\& \ ! (x \ \&\& \ y)$      $x$  oder  $y$ , und nicht beide

$(x \ || \ y) \ \ \ \ \ \&\& \ (!x \ || \ !y)$      $x$  oder  $y$ , und eines nicht

$!(\ !x \ \&\& \ !y) \ \ \ \ \ \&\& \ ! (x \ \&\& \ y)$     nicht keines, und nicht beide

$!(\ !x \ \&\& \ !y \ || \ x \ \&\& \ y)$     nicht: keines oder beide

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden* *zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 6  $\Rightarrow$

```
x != 0 && z / x > y
```

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden* *zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 6  $\Rightarrow$

```
true && z / x > y
```

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden* *zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 6  $\Rightarrow$

```
true && z / x > y
```

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 0  $\Rightarrow$

```
x != 0 && z / x > y
```

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 0  $\Rightarrow$

```
false && z / x > y
```

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden* *zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 0  $\Rightarrow$

```
false (falsch)
```

# Kurzschlussauswertung

- Logische Operatoren `&&` und `||` werten den *linken Operanden* *zuerst* aus.
- Falls das Ergebnis dann schon feststeht, wird der rechte Operand *nicht mehr* ausgewertet.

x hat Wert 0  $\Rightarrow$

```
x != 0 && z / x > y
```

$\Rightarrow$  Keine Division durch 0