



Informatik I

Information & Daten

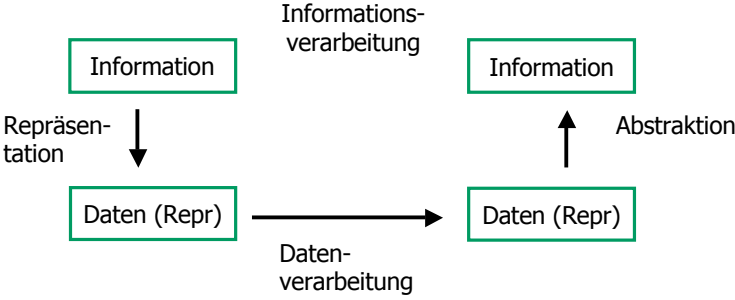
Repräsentation von Daten

G. Zachmann
 Clausthal University, Germany
zach@in.tu-clausthal.de

Daten & Informationen

- Menschen sind an **Informationen** interessiert
- Computer verarbeiten **Daten**
 - lesen, verknüpfen, schreiben





```

graph TD
    I1[Information] -- Repräsentation --> D1[Daten (Repr)]
    D1 -- Datenverarbeitung --> D2[Daten (Repr)]
    D2 -- Abstraktion --> I2[Information]
    I1 -- Informationsverarbeitung --> I2
  
```



The diagram illustrates the relationship between Information and Daten (Repr) through various processes. It shows a cycle where Information is represented as Daten (Repr), which is then processed (Datenverarbeitung) to produce more Daten (Repr), which is then abstracted back into Information. Additionally, Information is processed (Informationsverarbeitung) to produce more Information.

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 2

- Was ist Information?
 - Definitionen der Informatik beinhalteten: Die Wissenschaft von der systematischen Verarbeitung von *Informationen*.
 - Begriff Information ist von zentraler Bedeutung für die Informatik
 - Leider bis heute wenig präzisiert
- "Definition" Information und Repräsentation:
 - Information nennen wir den abstrakten Gehalt („Bedeutungsinhalt“, „Semantik“) eines Dokumentes, einer Aussage, Beschreibung, Anweisung, Nachricht oder Mitteilung.
 - Die äußere Form der Darstellung nennen wir Repräsentation (konkrete Form der Nachricht).

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 3

Semantische Äquivalenz

- Achtung: keine 1:1-Beziehung zwischen Information und Repräsentation
- Dieselbe Information kann auf verschiedene Weisen repräsentiert (dargestellt) werden
 - Begriffe: deutsch, englisch, Schreibschrift, Druckschrift, ...
 - Zahlen: Dezimalzahlen, Römische Zahlen, ...
- Dieselben Daten können verschiedene Informationen darstellen
 - IC = "Intercity" oder "Integrierte Schaltung" (*integrated circuit*)
- Zwei Darstellungen heißen „**semantisch äquivalent**“ genau dann, wenn ihre Interpretation dieselbe abstrakte Information liefert.

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 4

Bits

- Informationen werden durch zwei sich ausschließende Zustände repräsentiert

Ein Bit (= Binary Digit) ist entweder „0“ oder „1“

Früher auch: „O“ und „L“

- technisch leicht und eindeutig zu unterscheiden

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 5

Interpretation		Techn. Realisierung	
0	1	0	1
nein	ja	ungeladen	geladen
falsch	wahr	0 Volt	5 Volt
weiß	schwarz	unmagnetisiert	magnetisiert
gerade	ungerade	kein Licht	Licht
nicht best.	bestanden	kein Loch	Loch
...

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 6

Bitfolgen

- Mit einem Bit können nur $2^1 = 2$ Werte dargestellt werden
- Mit n Bits können 2^n verschiedene Werte dargestellt werden
- Müssen k ($= 2^n$) Werte dargestellt werden, benötigt man $n = \log_2(k)$ Bits

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 7

Hexziffern


- Lange Folge von Nullen und Einsen für den Menschen sehr unübersichtlich
0011110101000101001010100110
- Deswegen: aufteilen in 4er-Gruppen
0011 1101 0100 0101 0010 1010 0110
- 16 verschiedene Bitfolgen mit 4 Bits
- jede bekommt ein Symbol, Ziffern '0'...'9', Zeichen 'A'...'F'
- Zuordnen des Namens zu jeder 4er-Gruppe
- kompakte Hexdarstellung

0000 = 0	0100 = 4	1000 = 8	1100 = C
0001 = 1	0101 = 5	1001 = 9	1101 = D
0010 = 2	0110 = 6	1010 = A	1110 = E
0011 = 3	0111 = 7	1011 = B	1111 = F

0011 1101 0100 0101 0010 1010 0110
3 D 4 5 2 A 6

3D452A6


G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 8



Bytes

- CPU liest oder schreibt Daten immer in Gruppen von Bits
 - wäre zu langsam, Bits einzeln zu behandeln
 - derzeit üblich: 8, 32, 64 Bits
 - Rechner werden entsprechend genannt: 8-Bit-Rechner, 32-Bit-Rechner, etc.
- **Byte** ist definiert als Gruppe von 8 Bit
 - darstellbar mit zwei Hex-Ziffern (00, ..., FF)
 - kann $256 = 16^2 = 2^8$ verschiedene Werte annehmen

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 9



Files

- Repräsentation auf externem Speicher
- Deutsch: Datei
- beliebig lange Folge von Bytes
- meist auf Festplatten, Disketten, CD-ROM's, ...
- jede Information muß sich auf irgendeine Weise als Folge von Bytes darstellen lassen
 - Art der gespeicherten Information sollte durch Dateinamen charakterisiert werden
 - Üblich sind zwei Teile: Name.Suffix
 - Beispiel: Info1.ppt
 - Suffix bezeichnet File-Typ

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 10

Darstellung logischer Werte

- es gibt genau zwei Wahrheitswerte: wahr (true) und falsch (false) (jedenfalls in der klassischen Aussagenlogik)
- zur Repräsentation genügt ein Bit
- da meist ganze Bytes als kleinste Einheit verwendet werden, kodiert man häufig:

falsch: 0000 0000
 wahr: 1111 1111
- in fast allen Sprachen (C/C++, Java, Python, ...) gilt die Konvention

0000 0000: falsch
 alles andere: wahr

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 11

Logische Verknüpfungen

- Eine Aussage ist wahr oder falsch
 - Ich studiere in Bonn.
 - Ich studiere Philosophie.
- Aussagen können verknüpft werden
 - Ich studiere nicht in Bonn.
 - Ich studiere in Bonn Philosophie.
 - Ich studiere in Bonn oder ich studiere irgendwo Philosophie.
- "Boolesche" Operationen
 - George Boole, engl. Mathematiker, 1815-1864

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 12

		Negation NICHT <i>not</i>	Multipli- kation UND <i>and</i>	Addition ODER <i>or</i>	Exklusives Oder <i>xor</i>
x	y	x'	$x*y$	$x+y$	$x\oplus y$
0	0	1	0	0	0
0	1	1	0	1	1
1	0	0	0	1	1
1	1	0	1	1	0

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 13

- Logische Operationen auf Bitfolgen
 - Bitfolgen können auch als Folgen logischer Werte verstanden werden
 - Boolesche Operationen sind dann bitweise auszuführen

NOT 0100 1100 = 1011 0011
 0100 1100 **AND** 1000 1111 = 0000 1100

G. Zachmann Informatik 1 - WS 05/06 Repräsentation von Daten 14

- Verschiedene Schreibweisen:

	Schreibweise	andere Schreibweisen	C, C++, Java	
			logische Verkn.	bitweise Verkn.
NICHT	x'	\bar{x} $\neg x$!x	~x
UND	$x*y$	$x^{\wedge}y$	x&& y	x&y
ODER	$x+y$	$x^{\vee}y$	x y	x y
XOR	$x\oplus y$			x^y

Und fast allen anderen heutigen Sprachen

"The battle over syntax was won by C"