

Informatik-Grundbegriffe

Autor: E. Engelhardt

Stand: 03. Juni 2009

Was ist Informatik?

- ◆ Einfache Definition:
 - Informatik ist die Wissenschaft von der systematischen Verarbeitung von Informationen, insbesondere der automatischen Verarbeitung mit Hilfe von Digitalrechnern.
- ◆ Informatik ist weder Natur- noch Technikwissenschaft, sondern Strukturwissenschaft (wie Mathematik), d.h.:
 - Informatik behandelt und untersucht vom Menschen geschaffene formale Strukturen (Datenstrukturen, Sprachstrukturen, Systemstrukturen)
 - Informatik ist eine Disziplin der Informationswissenschaften
- ◆ Zentrales Thema der Informatik ist die Formulierung und Realisierung von Algorithmen

Weitere Informationswissenschaften

- ◆ Informationstheorie
 - Verarbeitung und Übertragung von Informationen; Codierung, Störungen, Korrekturverfahren
- ◆ Kybernetik
 - Wissenschaft von der Steuerung, Regelung und Informationsübertragung in technischen und biologischen Systemen (Rückkopplung, Lernen, adaptive Verfahren)
- ◆ Systemtheorie
 - untersucht die Rolle der Information unter Berücksichtigung gesellschaftlicher, soziologischer und politischer Aspekte
- ◆ Kognitionswissenschaften
 - Wissenschaft vom „Wissen“, Berücksichtigung psychologischer Aspekte
- ◆ Linguistik
 - Sprachwissenschaften

Teilgebiete der Informatik

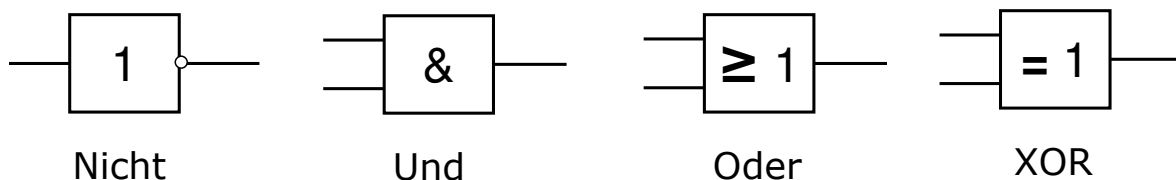
- ◆ Theoretische Informatik
 - Formale Sprachen, Automatentheorie, Theorie der Berechenbarkeit, Komplexitätstheorie, Algorithmenanalyse, Theorie der Programmierung
- ◆ Technische Informatik
 - Hardwarekomponenten, Schaltnetzwerke, Schnittstellentechnik, Computerkonstruktion, Steuerungen, Mikroprogrammierung
- ◆ Praktische Informatik
 - Algorithmen und Datenstrukturen, Softwaretechnik, Programmiersprachen und Compiler, Programmiermethoden, Betriebssysteme
- ◆ Angewandte Informatik
 - Informationssysteme, Prozess-Steuerung, Künstliche Intelligenz, Computergrafik, Simulation und Modellierung, spezifische Gebiete in Natur- und Ingenieurwissenschaften (Medieninformatik, Bioinformatik, ...)

Daten und Informationen

- ◆ Daten (von lateinisch *datum* = gegeben)
 - Angaben über reale Objekte
 - Eigenschaften (Attribute) von realen Objekten
 - Darstellung erfolgt mit Hilfe von Symbolen und Zeichenketten
- ◆ Informationen
 - sind abstrakt
 - zweckbezogener Inhalt von Daten
- ◆ Nachrichten
 - Träger von Informationen

Logische Grundschaltungen (1/3)

- ◆ Schöpfer der Aussagenlogik: englischer Mathematiker George Boole (1815 - 1864)
- ◆ Elementare logische Operationen
 - Negation (NICHT, NOT, \neg , Überstrich)
 - Konjunktion (UND, AND, \wedge , \cdot)
 - Disjunktion (ODER, OR, \vee , $+$)
 - Antivalenz, Exklusiv-ODER (XOR, \oplus)
- ◆ Symbole nach Norm DIN 40 900



Logische Grundschaltungen (2/3)

Wahrheitstabelle

a	b	$a \wedge b$ (a · b)	$a \vee b$ (a + b)	$a \oplus b$
1	1	1	1	0
1	0	0	1	1
0	1	0	1	1
0	0	0	0	0

- ◆ Weitere logische Schaltungen:
 - Sheffer-Funktion NAND (Not And)
 - Peirce-Funktion NOR (Not Or)

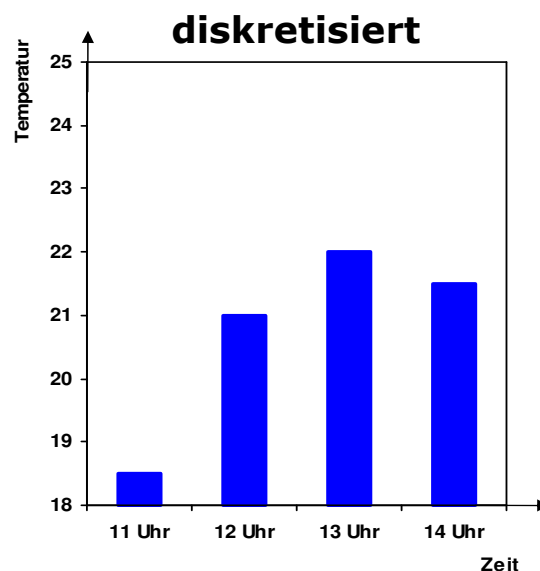
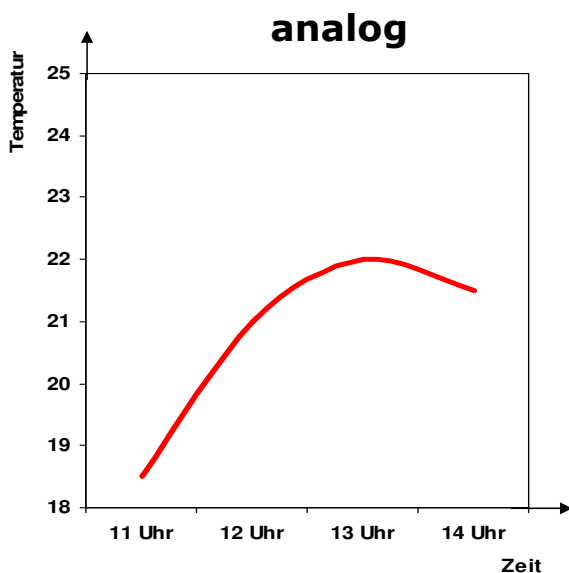
Logische Grundschaltungen (3/3)

- ◆ Realisierung von AND durch NOT und OR:
$$a \cdot b = \overline{\overline{a} + \overline{b}}$$
- ◆ Realisierung von OR durch NOT und AND:
$$a + b = \overline{\overline{a} \cdot \overline{b}}$$
- ◆ Realisierung der drei Grundschaltungen durch NAND und NOR:
 - $\overline{a} = a \text{ NAND } a$ $\overline{a} = a \text{ NOR } a$
 - $a \cdot b = (a \text{ NAND } b) \text{ NAND } (a \text{ NAND } b)$
 - $a \cdot b = (a \text{ NOR } a) \text{ NOR } (b \text{ NOR } b)$
 - $a + b = (a \text{ NAND } a) \text{ NAND } (b \text{ NAND } b)$
 - $a + b = (a \text{ NOR } b) \text{ NOR } (a \text{ NOR } b)$

Speicherung von Daten (1/4)

- ◆ Alles was zähl- und messbar ist, kann durch binärcodierte Daten ausgedrückt werden
 - Zahlen, Texte
 - physikalische Größen (optische, akustische, ...)
- ◆ Dadurch können alle diese Daten durch Computer verarbeitet werden (Universalität der Daten)
- ◆ Die physikalischen Größen werden diskretisiert und digitalisiert, d.h. unendlich viele Werte werden durch endlich viele Werte ersetzt (Diskretisierung oder Rasterung von analogen Daten);
- ◆ Die endlich vielen Werte werden digital (durch Zahlen) dargestellt (Digitalisierung)

Speicherung von Daten (2/4)



11.00 Uhr	18.5 °C
12.00 Uhr	21.0 °C
13.00 Uhr	22.0 °C
14.00 Uhr	21.5 °C

digitalisierte Darstellung

Speicherung von Daten (3/4)

- ◆ Vorteile binärcodierter Darstellung von Daten
 - Genauigkeit (beliebig genau angebar)
 - geringe Störanfälligkeit
 - die binären Symbole können relativ einfach durch physikalische Größen dargestellt werden
 - verlustlose Speicherbarkeit
- ◆ Möglichkeiten zur Darstellung von Binärzeichen

Binärwert 1	Binärwert 0	Anwendung
Spannung	keine Spannung	Schaltkreise
Elektrische Ladung	keine elektrische Ladung	Schaltkreise
Pos. Magnetisierung	Negat. Magnetisierung	Magnetspeicher
Loch	kein Loch	Lochkarten und Lochbänder
hell	dunkel	Bildpunkt
Frequenz 1	Frequenz 2	Frequenzmodulation
Amplitude 1	Amplitude 2	Amplitudenmodulation

Speicherung von Daten (4/4)

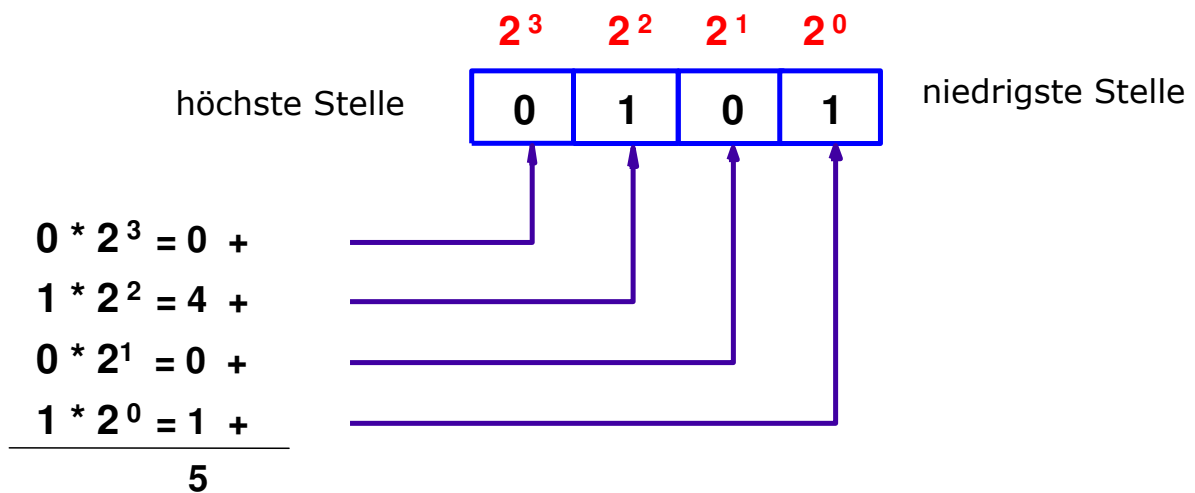
- ◆ *Bit* (binary digit) ist der kleinstmögliche Speicherplatz
- ◆ Eine Folge von acht Bit wird *Byte* genannt (ein halbes Byte wird *Nibble* genannt)
- ◆ Byte ist die übliche Maßeinheit für die Speicherkapazität
- ◆ Einteilung orientiert sich an der physikalischen Nomenklatur (**SI-Präfixe** mit 10er-Potenzen; eine Einheit= $10^3=1.000$), wobei eine Einheit= $2^{10}=1.024$ ist (**Binärpräfixe**)

Unterschied

1 kB (Kilobyte)	= 10^3	~ 2^{10} Byte (Kibibyte; KiB/KB)	2,4%
1 MB (Megabyte)	= 10^6	~ 2^{20} Byte (Mebibyte; MiB)	4,9%
1 GB (Gigabyte)	= 10^9	~ 2^{30} Byte (Gibibyte; GiB)	7,4%
1 TB (Terabyte)	= 10^{12}	~ 2^{40} Byte (Tebibyte; TiB)	10,0%
1 PB (Petabyte)	= 10^{15}	~ 2^{50} Byte (Pebibyte; PiB)	12,6%
1 EB (Exabyte)	= 10^{18}	~ 2^{60} Byte (Exbibyte; EiB)	15,3%
1 ZB (Zettabyte)	= 10^{21}	~ 2^{70} Byte (Zebibyte; ZiB)	18,1%
1 YB (Yottabyte)	= 10^{24}	~ 2^{80} Byte (Yobibyte; YiB)	20,9%

Binärsystem (1/4)

- ◆ Das Binärsystem oder Dualsystem ist ein Stellenwertsystem, wie das Dezimalsystem (Position legt Wert fest)
 - bei additiven Systemen werden die Ziffern addiert (Mesopotamien, Ägypten)
 - Römische Zahlen sind Mischsystem

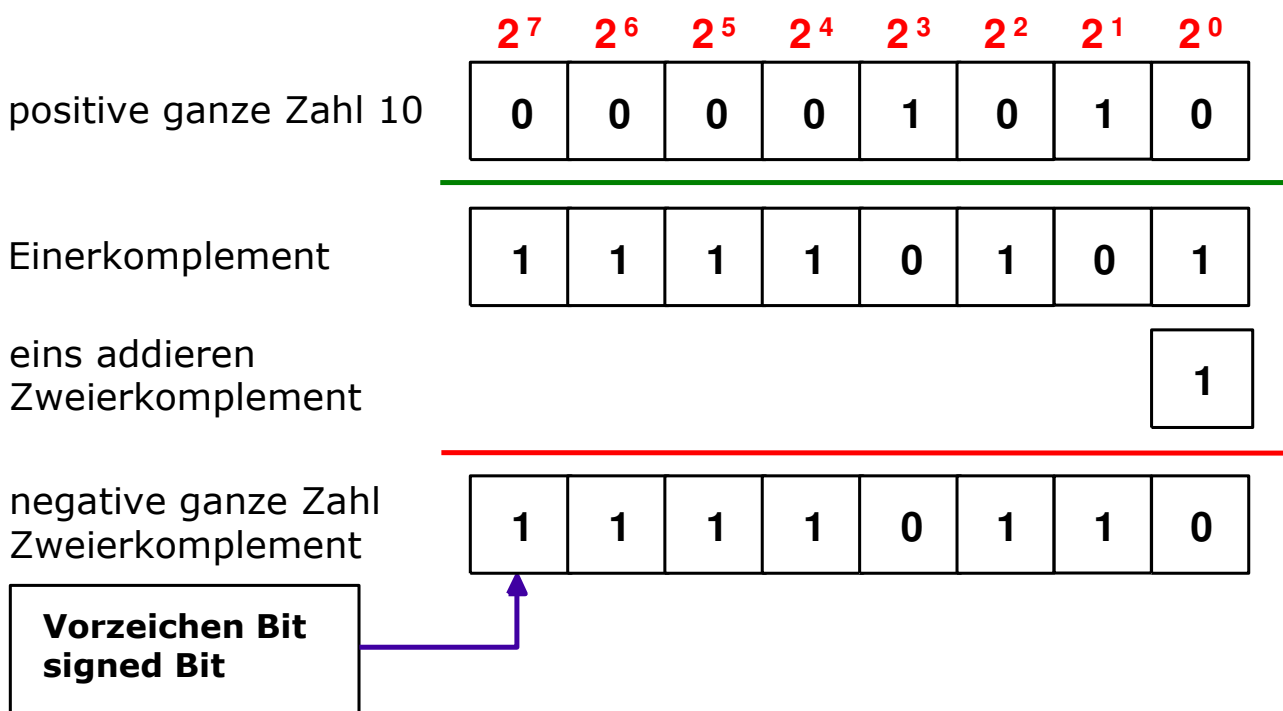


Folie 13 von 30

Informatik-Grundbegriffe

Binärsystem (2/4)

- ◆ Negative Zahlen werden durch Zweierkomplement gebildet



Folie 14 von 30

Informatik-Grundbegriffe

Binärsystem (3/4)

Allgemeine Summenformel zum Binärsystem

Größe mit n Stellen darstellbare vorzeichenlose Zahl

$$\sum_{i=0}^{n-1} 2^i = 2^n - 1$$

Anzahl verschiedener Zahlen (inklusive der Null)

$$\sum_{i=0}^{n-1} 2^i + 1 = 2^n$$

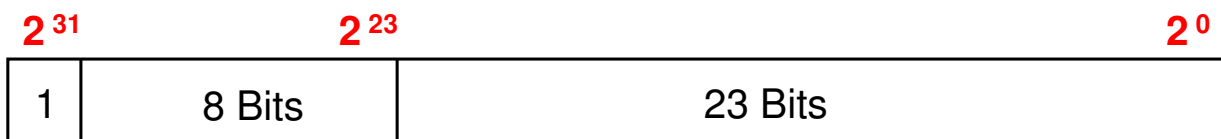
Betragsgröße mit n Stellen darstellbare vorzeichenbehaftete Zahlen

Positive Zahl: $2^{n-1} - 1$

Negative Zahl: $2^{n-1} - 2^n = -2^{n-1}$

Binärsystem (4/4)

- ◆ Darstellung von 32-Bit Gleitkommazahlen nach IEEE-754-Standard



s e: Exponent
└─ Vorzeichenbit

f: signifikante Stellen Mantisse

$$x = (-1)^s * (1.f) * 2^{e-127}$$

Größter relativer Fehler:

$$2^{-24}$$

Genauigkeit:

≈ 7 Dezimalstellen

Kleinste positive Zahl:

$$2^{-126} \approx 1.2 * 10^{-38}$$

Größe positive Zahl:

$$(2 - 2^{-23}) * 2^{127} \approx 3.4 * 10^{38}$$

Darstellung der Null:

$$s = 0, e = 0, f = 0$$

Unendlich (-/+):

$$s = 0/1, e \text{ alle Bits gesetzt } f = 0$$

NaN (not a number):

$$s = 0/1, e \text{ alle Bits gesetzt } f > 0$$

Darstellung von Texten (1/3)

- ◆ Textzeichen werden kodiert
- ◆ Häufig verwendete Codes sind:
 - ASCII **A**merican **S**tandard **C**ode of **I**nformation **I**nterchange
 - ANSI **A**merican **N**ational **S**tandard **I**nstitute
 - Unicode 16-Bit-Code (wird u.a. in Java verwendet)
 - EBCDIC **E**xtended **B**inary **C**oded **D**ecimal **I**nterchange **C**ode
- ◆ ASCII-Code und ANSI-Code sind in der Kodierung der Buchstaben und Ziffern gleich, unterscheiden sich aber bei den Sonderzeichen (auch Umlaute, sz, usw.)
- ◆ Im MS-DOS; UNIX, ... wird ASCII verwendet und im Win3.x, Win95/98, WinNT, ... wird ANSI verwendet
- ◆ Die druckbaren Zeichen beginnen ab 32 (dezimal)

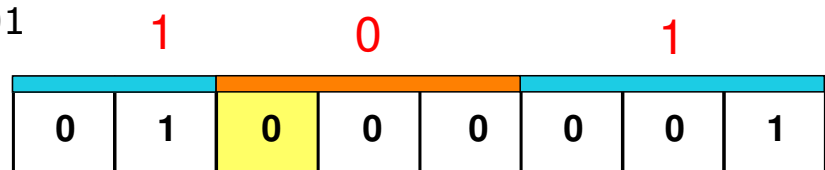
Darstellung von Texten (2/3)

Das Zeichen 'A'

Hexadezimal: 0x41

Dezimal: 65 Oktal: '0101

Oktalziffern



Das Zeichen 'a'

Dezimal: 97 Oktal: '141

Hexadezimal: 0x61

Bit für Groß- Kleinbuchstaben



Hexadezimalziffern (Sedezimalziffern)

Darstellung von Texten (3)

ASCII-Tabelle

Hex	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	NL	VT	NP	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL
8	Ç	ü	é	â	ä	à	å	ç	ê	ë	è	ï	î	ì	Ä	Å
9	É	æ	Æ	ô	ö	ò	û	ù	ÿ	Ö	Ü	ø	£			f
A	á	í	ó	ú	ñ	Ñ	ª	º	¿			½	¼	¡	«	»
B	°	±	²	³	´				¹	º	»	¼				¿
C	À	Á	Â	Ã	Ä	Å			È	É	Ê	Ë	Ì	Í	Î	
D						Õ			Ù	Ú	Û	Ü				ß
E		ß					µ									
F		±									.			²	þ	

Hardware und Software

- ◆ Hardware bezeichnet die physischen Bestandteile von Rechnern (Rechnergehäuse, Platinen, Stromversorgung, Kabel, Tastatur, Bildschirm, ...)
- ◆ Rechnerbestandteile (Tastatur, Drucker, Plotter, ...) enthalten oft auf Chips gebrannte Programme (Firmware), die oft mit zur Hardware gezählt werden
- ◆ Software sind Programme

Software



System- und Basissoftware

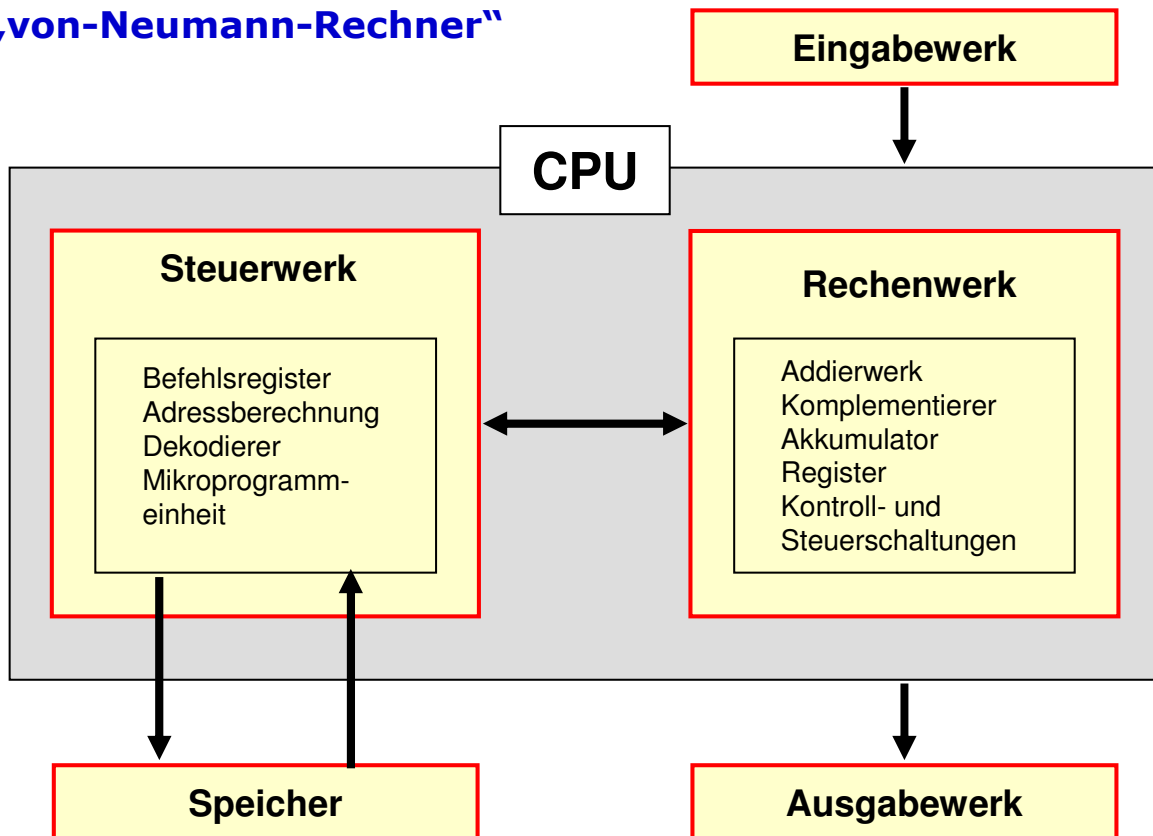
Betriebs-, Programmier- und Datenbanksysteme, Kommunikations- und Grafiksoftware

Anwender-Software

Bearbeitung eines konkreten Anwenderproblems (Textverarbeitung, Tabellenkalkulation, Spiele, ...)

Architektur (1/3)

„von-Neumann-Rechner“



Folie 21 von 30

Informatik-Grundbegriffe

Systematisierung nach Leistung (2/3)

- ◆ Personalcomputer (PC)
 - Konzentration der Rechenleistung auf einen Arbeitsplatz
 - in der Regel für einen Nutzer
- ◆ Workstation
 - hohe Rechen- und Grafikleistung an einem Arbeitsplatz
 - für Forschungs- und Entwicklungslabore, CAD-Einsatz
 - für Vernetzung konzipiert, Multi-User und Multitasking
 - meist UNIX-Derivat als Betriebssystem
- ◆ Midrange-Systeme (z.B. AS/4000)
 - Robustes Computersystem mittlerer Leistung
 - Komplettlösung für mittlere und große Unternehmen
 - 5 - 200 dezentral angeordnete Arbeitsplätze

Folie 22 von 30

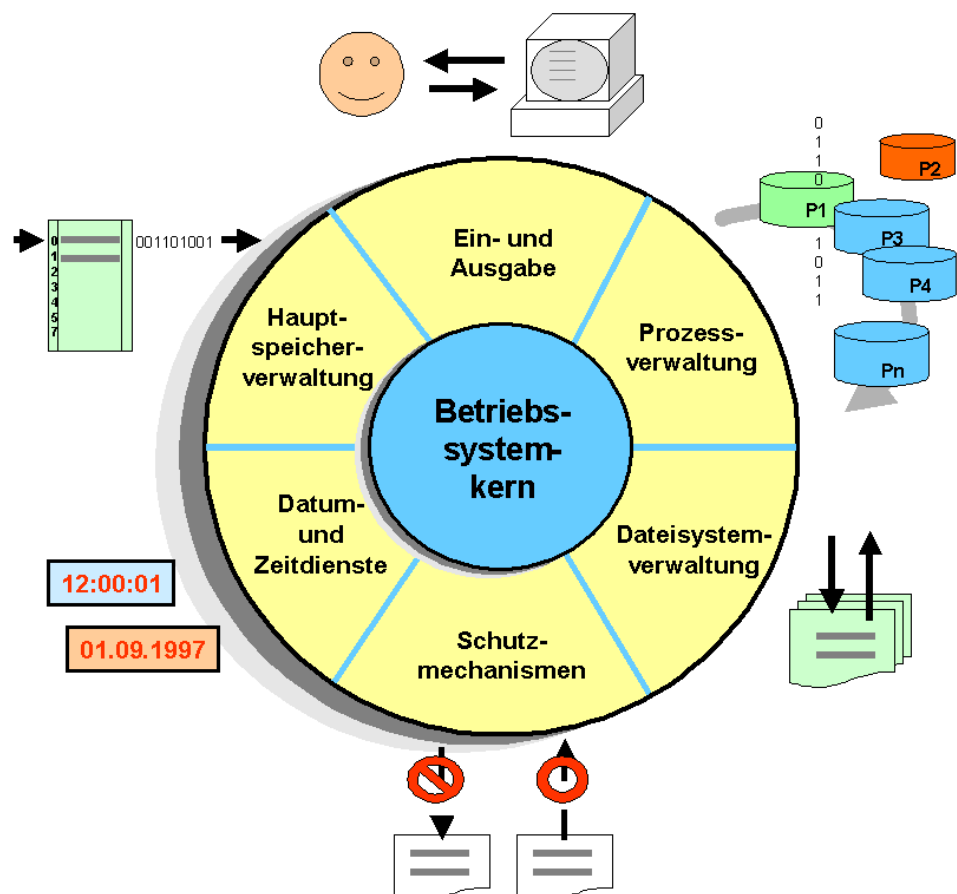
Informatik-Grundbegriffe

Systematisierung nach Leistung (3/3)

- ◆ Mainframe (z.B. S390)
 - Großrechner der oberen Leistungsklasse für Entwicklungsabteilungen und Forschungszentren
 - Zentrale Verwaltung von Banken und Versicherungen
 - Nutzung über Terminals auch über große Flächen
 - Hohe Rechenleistung und umfangreiche Massendatenspeicher
- ◆ Supercomputer (Cray, Deep Blue, ...)
 - Hochleistungsrechner immer um etwa Faktor 100 über den besten Workstations
 - Oft viele Prozessoren, die aufwändig gekühlt werden
 - Zunehmend auch Parallelschaltung tausender „normaler“ Prozessoren
 - Einsatz für Forschung, Simulationen, Wetterprognose, ...

Was ist ein Betriebssystem?

- ◆ Schnittstelle zwischen Mensch und Computer
- ◆ Ermöglicht Kommunikation mit Computer
- ◆ Eingabe von Kommandos und Start von Programmen

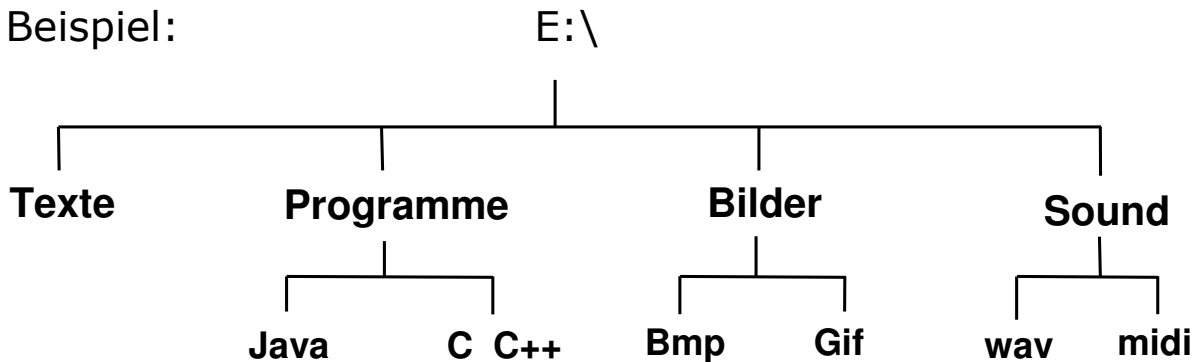


Dateisysteme (1/2)

- ◆ FAT File Allocation Table (DOS, Win*)
- ◆ NTFS New Technology File System (WinNT)
- ◆ HPFS High Performance File System (OS/2)
- ◆ Ext2 Extended File System (Linux)
- ◆ NFS Network File System

- ◆ Sind meist hierarchisch aufgebaut

- ◆ Beispiel:



Dateisysteme (2/2)

- ◆ Daten werden zu organisatorischen Einheiten in Dateien zusammen gefasst (**Daten-Kartei**, engl.: file)
- ◆ Dateien können auf zwei Arten im Dateisystem angegeben werden
 - absolute Pfadangabe (Angabe aller Stufen)
 - relative Pfadangabe (relativ zum „Standort“ des Nutzers)
- ◆ Zum Dateinamen gehört immer auch der Pfad
- ◆ Viele Anwendungsprogramme haben Standardverzeichnisse eingestellt
- ◆ Dateien können dann nur mit dem Dateinamen (ohne Pfad) angegeben werden, wenn:
 - Datei im selben Verzeichnis ist (wie Programm oder User)
 - der Pfad in der Variablen „PATH“ hinterlegt ist (alle in der Variablen „PATH“ angegebenen Pfade werden durchsucht)

Systemumgebung (Environment)

- ◆ Wichtige Parameter werden in Systemvariablen (Environment-Variablen) hinterlegt
 - PATH Pfade mit ausführbaren Dateien
 - TEMP / TMP temporärer Speicherplatz
 - HOME User-HOME-Verzeichnis
 - MAIL Verzeichnis mit Mails
 - ...
- ◆ Dadurch wird Unabhängigkeit von der konkreten Umgebung erreicht
- ◆ In Programmen sollten nie absolute Pfade angegeben werden
- ◆ Die Environment-Variablen können ausgewertet werden (natürlich auch in selbst programmierten Programmen)

Dateien (1/3)

- ◆ Datei (Daten Kartei) enthält Daten
 - Maschinencodeprogramme
 - reine ASCII-Texte (ANSI-Texte)
 - Shell-Scripte (UNIX-Shell-Programme, BAT-Datei, ...)
 - Quellcode-Programme (C/C++, Java, Basic, ...)
 - Informations-Dateien (config.sys, ini-Dateien, ...)
 - Dateien mit bestimmten Formaten spezieller Anwendungen
 - Formate von Textbearbeitungsprogrammen
 - Bildformate
 - Soundformate
 - Filmformate
 - komprimierte Daten
 - verschlüsselte Daten
- ◆ Oft werden die Formate durch Dateitypen (Extension) kenntlich gemacht (.txt, .bmp, .c, .obj, .exe, ...)

Dateien (2/3)

- ◆ Ausführbare Dateien
 - Maschinencode-Programme (enthalten direkt auf dem betreffenden Prozessor ausführbaren Code und die zugehörigen Daten)
 - Shell-Scripte: sie enthalten lesbaren und editierbaren Code, der vom Betriebssystem interpretativ in Maschinencode umgesetzt wird (UNIX-Shellscripte, BAT-Dateien)
 - Interpreter, die entsprechende dazugehörige Dateien interpretativ in Maschinencode umsetzen (Basic, Perl, ...), die in editierbarer Form vorliegen
- ◆ Ausführbare Dateien können von den Window-Oberflächen oder als Kommandozeilen-Programme (wenn es keine Window-Programme sind) gestartet werden
 - Programme können auch oft aus anderen Anwendungen heraus gestartet werden

Dateien (3/3)

Ausführung von Programmen in einer Kommandozeile

Eingabeanforderungen

- ◆ Kommandos können unter DOS groß oder klein oder gemischt groß/klein geschrieben werden
- ◆ Unter UNIX ist Groß/Klein-Schreibung zu beachten

>	DOS
\$	UNIX Bourne-Shell
%	UNIX C-Shell
#	UNIX Korn-Shell
> Kommando [Parameter]	

- ◆ Die Schreibweise der Parameter hängt vom Kommando ab (evtl. case-sensitiv)!
- ◆ Das ist unabhängig vom Betriebssystem!