

Grundzüge der Wirtschaftsinformatik *Introduction to Business Information Systems*

Unit 3

Prof. Dr. Martin Hepp
<http://www.heppnetz.de>
mhepp@computer.org

<http://www.heppnetz.de/teaching/gwi/>

Link to the Previous Unit

- **Last Unit:** Central Processing Units – how does a computer execute a program?
Main memory (RAM) – temporary storage; loses content when power supply is interrupted.
- **Today:** How can data be stored persistently? How do we make sure that a fact can be found when needed?

Logistics

- **Lecture**
 - Tuesdays, 13:15 - 14:45, Auditorium Maximum (Building 33)
- **Tutorial and Exercises (Begin: October 17)**
 - Wednesdays, 11:30 – 13:00, Building 33 Room 2401 (in German)
 - Thursdays, 09:45 - 11:15, Building 43 Room 4/126 (in German) Waiting for feedback on available classrooms...
 - Thursdays, 15:00 - 16:30, Building 33 Room 2216 (in German)

Assignment from Last Week

- WI2: pp. 93-211, 387-515
- IBIS, pp. 55-78
- Review the slides

WI1 = Hansen/Neumann: Wirtschaftsinformatik 1; WI2 = Hansen/Neumann: Wirtschaftsinformatik 2; IBIS = Wigand et al: Introduction to Business Information Systems.

Structure of the Lecture

- Unit 1:** Introduction
- Unit 2:** Central Processing Units
- Unit 3:** [Storage and Data Structures](#)
- Unit 4:** Input and Output Devices
- Unit 5:** Software
- Unit 6:** Networks, Data Interchange, and the Internet
- Unit 7:** Design, Development, Deployment, and Operations of Information Systems
- Unit 8:** Office Applications
- Unit 9:** Enterprise Applications
- Unit 10:** Supply Chain Applications and E-Business
- Unit 11:** Management Support Systems
- Unit 12:** Exam Review

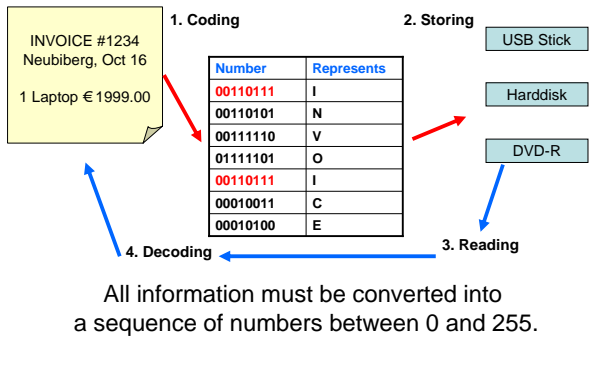
Storage and Data Structures

If we want to store information in a computer system, we need to

- Develop a **representation structure** and
- Have a **media** to write to and read from



Storage and Data Structures



Data vs. Information

Data		
8-28-2003	Miller, John	37.40
8-28-2003	Smith, Bill	23.20
8-27-2003	Burger, Mary	11.11
8-26-2003	Miller, John	40.00

Information

Turnover
111.71

Total sales by customer:

Miller, John	77.40
Smith, Bill	23.20
Burger, Mary	11.11

Customers on August 28:
Miller, John
Smith, Bill

Storage

- Exploiting physical phenomena for storing and retrieving data
- Examples:
 - Iron and iron oxide (ferrous oxide) can keep magnetism
 - A laser can be used to modify the surface of a media so that it reflects (or no longer reflects) light

Atomic Data

- Characters and Symbols
 - A-Z, 0-9
- Numbers
 - Integer
 - 1, 2, 3
 - Floating-point
 - 3.1415
 - 6.7

Data vs. Information

- Information:** Facts that they have value beyond their representation
 - Peter was born on July 4, 1980
 - This PC costs 1000 euros.
 - The temperature in Munich is 11 degrees Celsius.
- Data:** Representation of facts

Encoding Characters and Strings

- Define the set of relevant characters (an alphabet)
 - Examples
 - Just A-Z and 0-9
 - A-Z, a-z, ÄÖÜ, äöü, ß
 - Mathematical symbols
- Create a mapping between characters and binary values

Binary	Char	Hex	Alpha
100-0000-1000	64	40	D
100-0000-1001	65	41	E
100-0000-1010	66	42	F
100-0000-1011	67	43	G
100-0000-1100	68	44	H
100-0000-1101	69	45	I
100-0000-1110	70	46	J
100-0000-1111	71	47	K
100-0001-0000	72	48	L
100-0001-0001	73	49	M
100-0001-0010	74	4A	N
100-0001-0011	75	4B	O
100-0001-0100	76	4C	P
100-0001-0101	77	4D	Q
100-0001-0110	78	4E	R
100-0001-0111	79	4F	S
100-0001-1000	80	50	T
100-0001-1001	81	51	U
100-0001-1010	82	52	V
100-0001-1011	83	53	W
100-0001-1100	84	54	X
100-0001-1101	85	55	Y
100-0001-1110	86	56	Z
100-0001-1111	87	57	[
100-0010-0000	88	58	\
100-0010-0001	89	59]
100-0010-0010	90	5A	^
100-0010-0011	91	5B	_
100-0010-0100	92	5C	`
100-0010-0101	93	5D	{
100-0010-0110	94	5E	}
100-0010-0111	95	5F	~
100-0010-1000	96	60	
100-0010-1001	97	61	
100-0010-1010	98	62	
100-0010-1011	99	63	
100-0010-1100	100	64	
100-0010-1101	101	65	
100-0010-1110	102	66	
100-0010-1111	103	67	

Source: Wikipedia

Common Codesets

- ASCII: „American Standard Code for Information Interchange (ASCII)“
- Extended ASCII (8 Bit): ISO 8859
 - 256 characters → 0-255
- Unicode: 1,114,112 (= $2^{20} + 2^{16}$) possible code points
 - Currently about 100,000 of those used
 - Including Chinese, Korean, and Japanese characters
- UTF-8 and UTF-16
 - Varying number of bytes per character, depending on its frequency of usage

Encoding Real Values

- Fixed-point representation
 - E.g., one byte or half a byte per digit
- Floating-point representation
 - **Mantissa** or significant: string of digits
 - Base 2 or base 10
 - **Exponent**: The power of the base by which the significant is multiplied
- Example
 - $2.753E3 \Rightarrow 2.753 * 10^3$

On the Importance of Codeset Standardization

“ I have also approved recommendations of the Secretary of Commerce regarding standards for recording the Standard Code for Information Interchange on magnetic tapes and paper tapes when they are used in computer operations.

All computers and related equipment configurations brought into the Federal Government inventory on and after July 1, 1969, must have the capability to use the Standard Code for Information Interchange and the formats prescribed by the magnetic tape and paper tape standards when these media are used. ”

—Lyndon B. Johnson, Memorandum Approving the Adoption by the Federal Government of a Standard Code for Information Interchange, March 11th, 1968 at Gerhard Peters (ed.), The American Presidency Project, [9]

Source: Wikipedia

Precision Problems

- What is $1/3 * 3$?
 - 1?
 - 0.99999?
- Significant rounding errors can occur in computer systems, because such precision problems sum up

Encoding Integer Values

- Non-negative Integers
 - Simple binary numbers
 - Combine multiple bytes to store larger numbers
 - Example: 16 Bit Number for values 0 – 65535
 - Represented value: lower byte + $256 * \text{higher byte}$
- Positive and Negative Values
 - Simple: Use highest bit for storing the sign (+/-)
 - Two-complement

Complex Data

- Strings
 - „Peter Miller“, „Universität der Bundeswehr“
- Structured Data
 - Qty / Description / Price

Complex Data: Fixed Length vs. Explicit Delimiters

- **Fixed Length:** All fields have the same length
 - 255 characters for each street address
- **Explicit Delimiters:** A special character that is not included in the regular alphabet indicates the end of one data field
 - Example: Comma-separated Values (CSV)
„Peter Miller, Hauptstrasse 8, Neubiberg“

Machine-readable Content vs. Unstructured Data



ORDER
QTTY=3
ITEMNO=1234

Dataset and Files

- Dataset: Set of data elements (atomic or complex) that belong together
 - Address: Name, Street, ZIP, City
 - Invoice: Customer, Items, Total
- File: Collection of datasets of the same type
 - All addresses
 - All invoices

Further Examples

Good	Bad
Structured text message	Fax Image
Vector drawing of a floor plan	Photo

Locating a Dataset in a File

- Fixed length: Dataset Number * length of Dataset
 - Example:
 - Each address be 250 characters long
 - 1st address starts at first byte (0), 2nd at byte 250, 3rd at byte 500, ...
- Variable length
 - Explicit dataset delimiter (different from field delimiter)
 - Table with pointers to beginning of datasets

Storage

- Exploiting physical phenomena for storing and retrieving data
- Examples:
 - Iron and iron oxide (ferrous oxide) can keep magnetism
 - A laser can be used to change the surface of a media so that it reflects (or no longer reflects) light

Access Methods

- **Sequential:** records must be retrieved in order
 - Devices used are called sequential access storage devices (SASD)
- **Direct:** records can be retrieved in any order
 - Devices used are called direct access storage devices (DASDs)

Hard Disks



<http://computer.howstuffworks.com/hard-disk.htm/printable>

Hard Disks



<http://computer.howstuffworks.com/hard-disk.htm/printable>

Speed

$$r = 3.5" / 2$$
$$\text{diameter} = 3.5"$$
$$\text{circumference} = 2\pi r$$

$$\Rightarrow 3.5" * 3.1415 = \text{ca. } 11"$$

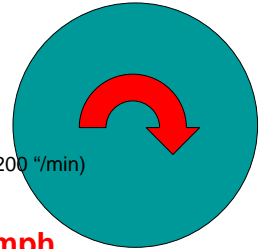
$$\Rightarrow 7200 \text{ rpm}$$

$$\Rightarrow 11" * 7,200 \text{ inch/minute (79,200 "/min)}$$

$$\Rightarrow 11" * 7,200 * 60 \text{ inch/hour}$$

$$\Rightarrow 4,752,000 \text{ inch/hour}$$

$$\Rightarrow 4,752,000 / 63,360 \Rightarrow \mathbf{75 \text{ mph}}$$



Hard Disks

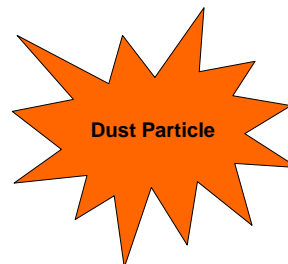
- Platters
- Read/Write Head
- Actuator Arm



http://www.flickr.com/photos/gek_at2000/485151116/

<http://computer.howstuffworks.com/hard-disk.htm/printable>

Dust and Abrasion



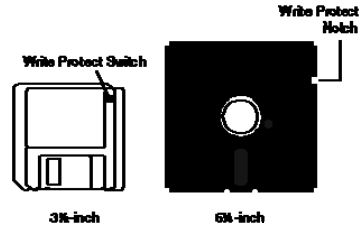
a few micrometers
(ca. 1/1,000,000 yard)

Headcrash



<http://www.flickr.com/photos/124330160/89745500/>

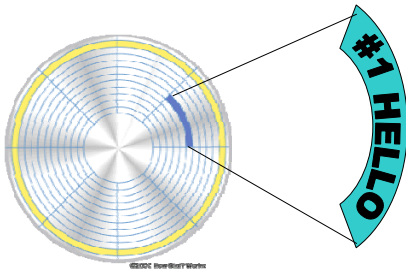
Floppy Disk



http://www.webopedia.com/TERM/f/floppy_disk.html

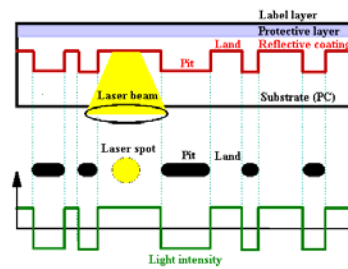
Hard Disks

- Tracks
- Sectors



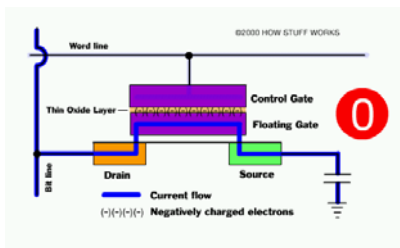
<http://computer.howstuffworks.com/hard-disk.htm/printable>

CD-ROM



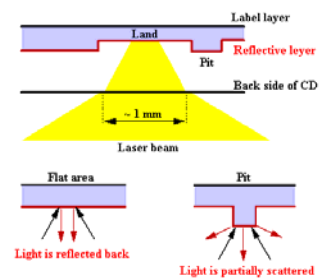
http://www.usbyte.com/common/compact_disk_3.htm

Flash Memory



<http://computer.howstuffworks.com/flash-memory.htm/printable>

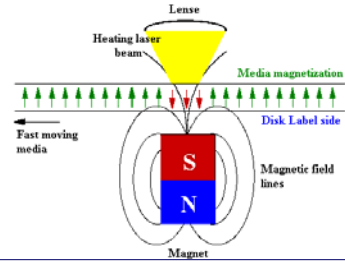
CD-ROM



http://www.usbyte.com/common/compact_disk_3.htm

Only ONE Track!

Magneto-Optical Drives



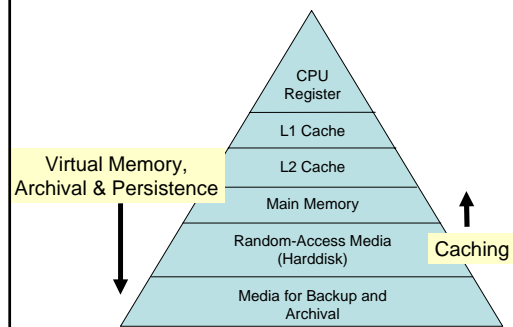
<http://www.usbyte.com/common/MOsystems.htm>

CLV vs. CAV

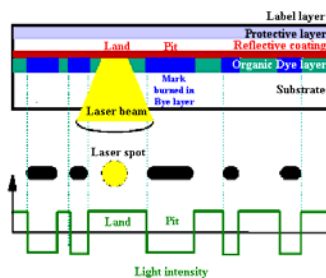
CLV: Constant Linear Velocity: inches per second remains constant

CAV: Constant Angular Velocity: rotations per second remains constant

Memory Hierarchy



CD-R



http://www.usbyte.com/common/recordable_CD.htm

Assignment for Next Week

- WI2: pp. 213-300
- Review the slides

WI1 = Hansen/Neumann: Wirtschaftsinformatik 1; WI2 = Hansen/Neumann: Wirtschaftsinformatik 2; IBIS = Wigand et al: Introduction to Business Information Systems.

Thank you!

The slides and additional materials will be available at
<http://www.heppnetz.de/teaching/qwi/>
shortly.

Don't forget: Tutorials will start this week!