Information Technology?

ver. 25 z drobnymi modyfikacjami!

Wojciech Myszka

2022-10-27 08:58:52 +0200

Wrocław University of Science and Technology

Instead of an introduction I

1. My office is in room 120 (building B1). You can contact me either in person or by e-mail (wojciech.myszka@pwr.edu.pl). My office hours:
   - Tuesday 09:15–11:45 (just after the class)
   - Thursday 09:15–11:45 (just after the class)

   If you **urgently**, i.e., outside office hours, need to met me in person — please send an e-mail to make an appointment.
2. I will use the USOS system for sending messages to you. You also can use it or use plain e-mail.
3. Lectures are **compulsory** for first semester students.

Instead of an introduction II

4. Final grade;
   There are two possibilities:
   - 4.1 Quiz at the end of the semester
   - 4.2 Student presentation (i.e., all students are divided into groups and each group presents a short — up to 30 minutes — presentation).
     Topics:
     - everything you are interested in...
     - ...but related to computers, information technology, the Internet, on-line life, social media, mobile technology,...

Instead of an introduction III

5. You can find auxiliary materials here:
   - 5.1 https://kmim.wm.pwr.edu.pl/myszka/didactics/information-technology/
   - 5.2 On the e-portal https://eportal.pwr.edu.pl/
     - course W10MBM-SI3075W # W # Information Technologies - Lecture
   - 5.3 Please let me know if something I am talking about is known to you. The same if something is not clear. It is better to interrupt me.
   - 5.4 And remember: I do prefer presentations as a way to grade you.

Information for your "sponsors"

(Yes, I know you are of legal age!)

1. There is an old Polish proverb:
   You can lead a horse to a water, but you can't make it to drink.
   Or better (classical version):
   You can show someone how to do something, but you can't make them do it.
2. You are **obligated to learn**. Teachers play a supporting role only.

Wrocław University of Science and Technology

Syllabus

1. Syllabus. Requirements.
2. Information
3. History and development of computer systems.
4. Elements of the computer system.
5. Binary logic, logic elements, binary arithmetic, half adder, and adder. Integer numbers. Floating-point numbers.
6. Computer internal design.
7. Software: Operating System & utilities.
8. Algorithms. Basic algorithmic structures.
9. The course ends with a series of short presentations prepared by the participants (groups of 3–4 students). I hope.
10. Another possibility is a quiz (test).

Wrocław University of Science and Technology

Information Technology

Nomenclature (mainly for Polish students)

1. Informatyka — Computer Science (Informatics???)
2. Technologie Informacyjne — Information Technology

Information Technology

Nomenclature (mainly for Polish students)

1. Informatyka — Computer Science (Informatics???)
2. Technologie Informacyjne — Information Technology

Information Technology (IT)

is a branch of engineering dealing with the use of computers and telecommunications equipment to store, retrieve, transmit and manipulate data... (Wiki)

Information Technology I

Assumptions

1. Virtually all students have already had contact with a computer and know the basics of the main applications (word processing, spreadsheet, web browser, instant messenger, social media app). And, probably, you are better than me!
2. The course aims to systematise and organise your knowledge and terminology.
3. We have only the lecture room and the projector...
4. The lecture will present a history of hardware and software development.

Information Technology II

Assumptions

5. We will focus on the issues of a computer's operation: mainly on the way it conducts calculations and on all consequences of binary arithmetic.
6. The important topics of the lecture will be :
   - the concept of an algorithm,
   - The task of algorithmization (and automation) of operations,
   - information about programmes and programming languages and examples.
7. The next topics will cover: computer networks, the Internet, and related problems (security, privacy, law, intellectual property, etc.).

Requirements

- Attendance: Lectures for freshers (first-year students) are ~~(still?)~~ obligatory!
- Quiz/Test/???
- Homework
- ...

Requirements

- Attendance: Lectures for freshers (first-year students) are ~~(still?)~~ obligatory!
- ~~Quiz/Test/???~~
- Homework
- ...

Requirements

- Attendance: Lectures for freshers (first-year students) are ~~(still?)~~ obligatory!
- ~~Quiz/Test/???~~
- Homework
- ...
- Final presentation (one presentation for a group of 3–4 persons)

Literature I

John Bayko.
Great microprocessors of the past and present (V 13.4.0).
http://www.cpushack.com/CPU/cpu.html, 2005.
Mainly PAST.

Richard P. Feynman.
Feynman Lectures on Computation.
CRC Press, 2018.

James Gleick.
The Information: A History, A Theory, A Flood.
Vintage, New York, 2.5.2012 edition, March 2012.

David Harel.
Computers Ltd.: What They Really Can't Do.
Oxford University Press, USA, December 2003.

Literature II


David Harel and Yishai Feldman.
Algorithmics: The Spirit of Computing.
Addison-Wesley, 3 edition, June 2004.


Donald E. Knuth.
The Art of Computer Programming, Volumes 1-4A Boxed Set.
Addison-Wesley Professional, 1 edition, March 2011.


James F. Kurose and Keith W. Ross.
Computer Networking: A Top-Down Approach.
Addison-Wesley, 6 edition, February 2012.


Éric Lévénez.
Windows history.
http://www.levenez.com/windows/, 2021.

Literature III


Éric Lévénez.
Computer languages history.
http://www.levenez.com/lang/, 2022.


Éric Lévénez.
Unix history.
http://www.levenez.com/unix/, 2023.


Jason Robert Carey Patterson.
Modern microprocessors – A 90-minute guide!
http://www.lighterra.com/papers/modernmicroprocessors/, 2016.


Diarmuid Pigott.
The encyclopedia of computer languages.
https://web.archive.org/web/20081020140121/http://hopl.murdoch.edu.au/, 2006.

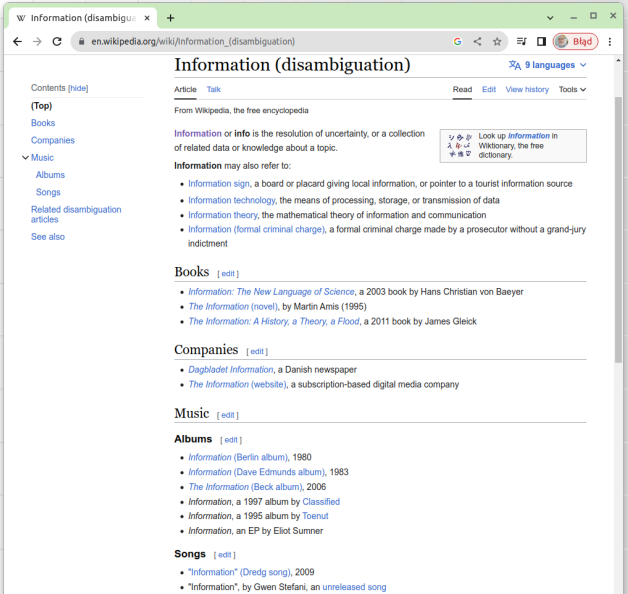
Literature IV


Abraham Silberschatz, Peter B. Galvin, and Greg Gagne.
Operating System Concepts.
Wiley, 8 edition, July 2008.

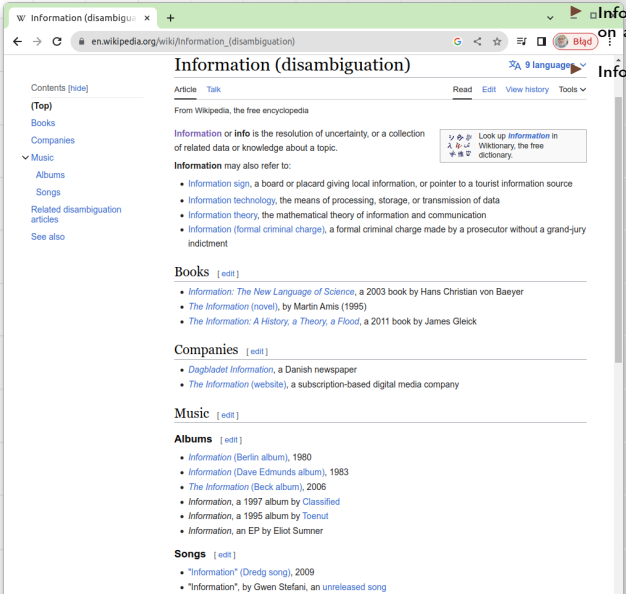

William Stallings.
Operating Systems: Internals and Design Principles.
Prentice Hall, 7 edition, February 2011.


Niklaus Wirth.
Algorithms + Data Structures = Programs.
Prentice Hall, 1st edition, February 1976.

Information



Information



- Information is a collection of related data or knowledge on a topic.
- Information may also refer to:
  - Data, or data used in computing
  - Physical information contained in a system
  - Information theory, the mathematical theory of information and communication
  - Directory assistance, a phone service used to find a specific telephone number and/or address of a residence, business, or government entity.
  - Information (formal criminal charge), a formal criminal charge made by a prosecutor without a grand jury indictment
  - Fisher information, in statistics

Wrocław University of Science and Technology

Information

Information

(shortened as info or info.) is that which informs, i.e., that from which data can be derived. Information is conveyed either as the content of a message or through direct or indirect observation of some thing. That which is perceived can be construed as a message in its own right, and in that sense, information is always conveyed as the content of a message. Information can be encoded into various forms for transmission and interpretation. For example, information may be encoded into signs and transmitted via signals.

Wrocław University of Science and Technology

Technology

Technology

Technology (from Greek, techne, “art, skill, cunning of hand”; and logia) is the collection of tools, including machinery, modifications, arrangements, and procedures used by humans. Engineering is the discipline that seeks to study and design new technologies.

(Wikipedia)

Wrocław University of Science and Technology

Technology

Technology

Technology (from Greek, techne, “art, skill, cunning of hand”; and logia) is the collection of tools, including machinery, modifications, arrangements, and procedures used by humans. Engineering is the discipline that seeks to study and design new technologies.

(Wikipedia)

technology n [C,U]

New machines, equipment, and ways of doing things that are based on modern knowledge about science and computers.

(from dictionary in my—very old—mobile)

Wrocław University of Science and Technology

How to measure Information?

- Is it possible at all?

How to measure Information?

- Is it possible at all?
- Units?

How to measure Information?

- Is it possible at all?
- Units?
- Information vs. uncertainty (lack of information).

How to measure Information?

- Is it possible at all?
- Units?
- Information vs. uncertainty (lack of information).
- Does information resolve the uncertainty?

Information

An example

- Let us think about a young couple waiting for their first baby.
- In general, before the examination, there is the same probability that they wait for a girl and a boy. (It is uncertainty).

Can we measure the amount of uncertainty?

Information

An example

- Let us think about a young couple waiting for their first baby.
- In general, before the examination, there is the same probability that they wait for a girl and a boy. (It is uncertainty).

Can we measure the amount of uncertainty?

- Let us assume, for a moment, that in His (or in Her) family, for ages, there were always more boys than girls (for example, 9 to 1).

Information

An example

- Let us think about a young couple waiting for their first baby.
- In general, before the examination, there is the same probability that they wait for a girl and a boy. (It is uncertainty).

Can we measure the amount of uncertainty?

- Let us assume, for a moment, that in His (or in Her) family, for ages, there were always more boys than girls (for example, 9 to 1).
- So, they can be **almost** sure that they will have a boy.

Information

An example

- Let us think about a young couple waiting for their first baby.
- In general, before the examination, there is the same probability that they wait for a girl and a boy. (It is uncertainty).

Can we measure the amount of uncertainty?

- Let us assume, for a moment, that in His (or in Her) family, for ages, there were always more boys than girls (for example, 9 to 1).
- So, they can be **almost** sure that they will have a boy.
- The amount of uncertainty is very small...

Information

An example

- Let us think about a young couple waiting for their first baby.
- In general, before the examination, there is the same probability that they wait for a girl and a boy. (It is uncertainty).

Can we measure the amount of uncertainty?

- Let us assume, for a moment, that in His (or in Her) family, for ages, there were always more boys than girls (for example, 9 to 1).
- So, they can be **almost** sure that they will have a boy.
- The amount of uncertainty is very small...

Information

An example

- Let us think about a young couple waiting for their first baby.
- In general, before the examination, there is the same probability that they wait for a girl and a boy. (It is uncertainty).

Can we measure the amount of uncertainty?

- Let us assume, for a moment, that in His (or in Her) family, for ages, there were always more boys than girls (for example, 9 to 1).
- So, they can be **almost** sure that they will have a boy.
- The amount of uncertainty is very small...
- On the other hand, in normal circumstances, the ratio is 1 to 1. It is difficult to predict the gender of the newborn.

Wrocław University of Science and Technology

Entropy

Let us assume that we are waiting for one of n messages. The probability of the i-th message is $p(i)$. According to the Shannon summary uncertainty (in information theory called **entropy**) H is equal to:

$$H = -\sum_{i=1}^{n} p(i) \log p(i).$$

~~(Minus sign means that it is lack of information?)~~

Entropy

Special cases

If all messages have equal probability $\left(p(i) = \frac{1}{n}\right)$ formula is much simpler

$$H = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = -\frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{n} =$$

$$-\frac{1}{n}(n \log \frac{1}{n}) = -(\log(1) - \log(n)) = \log(n)$$

Entropy

Special cases

If all messages have equal probability $\left(p(i) = \frac{1}{n}\right)$ formula is much simpler

$$H = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = -\frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{n} =$$
$$-\frac{1}{n}(n \log \frac{1}{n}) = -(\log(1) - \log(n)) = \log(n)$$

The unit of entropy/information is:

- **bit** (2-based logarithm),
- **nat** (e-based logarithm),
- **hartley** (10-based logarithm).

Entropy

Special case: $n = 2$

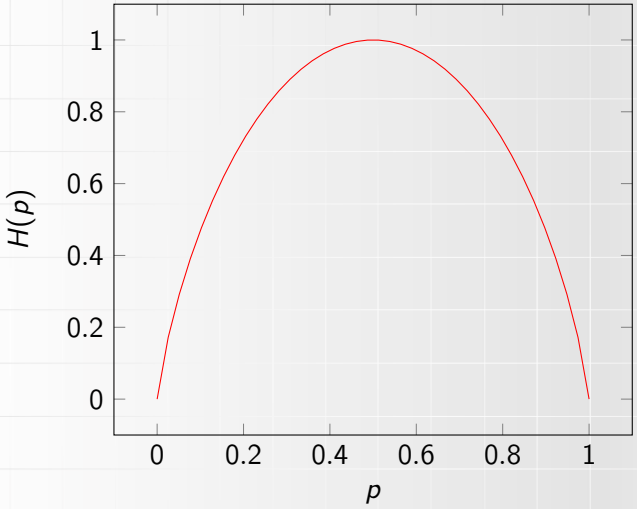
$$H_{\mathrm{b}}(p) = -p \log_2 p - (1-p) \log_2 (1-p).$$

When we use binary logarithm, the unit of H is called **bit**. And now, when probabilities are equal, we get exactly 1 bit of information.

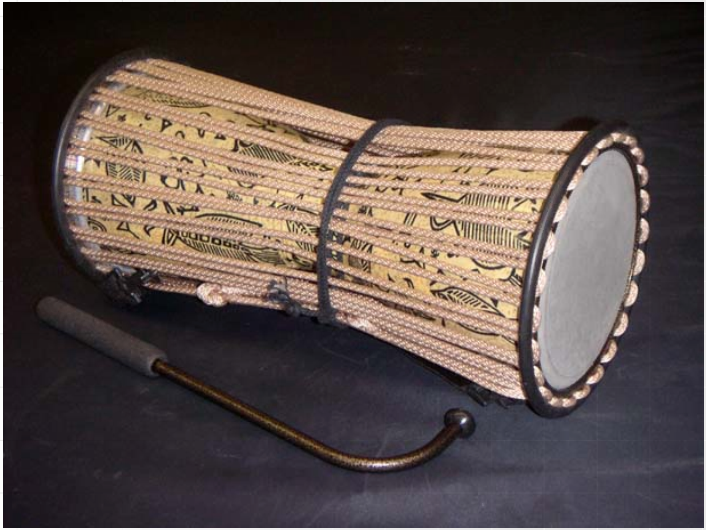
$$H_b = -0.5 \log_2(0.5) - (1-0.5) \log_2(1-0.5) = 0.5 + 0.5 = 1$$

Entropy

$H(p)$



Drums that talk



Wrocław University of Science and Technology

Drums that talk

- Europeans in sub-Saharan Africa
- One of the explorers received the information that every village had "facility of musical correspondence."
- By the time explorers discovered the talking drums in 1841, Samuel F. B. Morse was struggling with his own percussive code. His first idea was to send numbers, a digit at a time, with dots and pauses. The sequence ●●● ●● ●●●●● would mean 325. Every English word would be assigned a number, and the telegraphists at each end of the line would look them up in a special dictionary. Morse set about creating this dictionary himself, wasting many hours inscribing it on large folios.

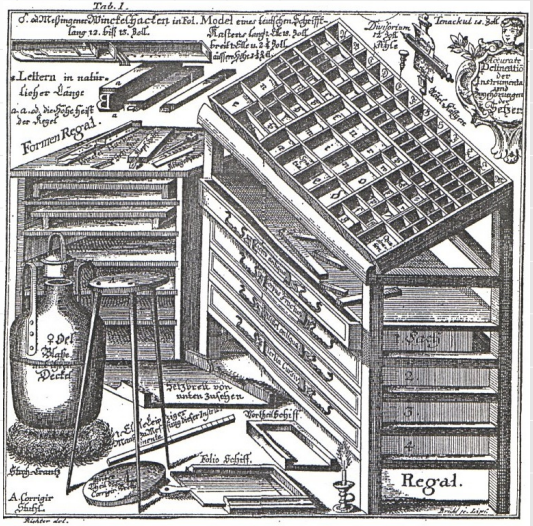
Drums that talk

Morse code

- *The dictionary or vocabulary consists of words alphabetically arranged and regularly numbered, beginning with the letters of the alphabet, so that each word in the language has its telegraphic number and is designated at pleasure, through the signs of numerals.*
- In the name of speed, Morse and Vail realised that they could save strokes by reserving the shorter sequences of dots and dashes for the most common letters. But which letters would be used most often?

Drums that talk

The type cases



Drums that talk

The type cases

Drums that talk

- In search of data on the relative frequencies of letters, Vail was inspired to visit the local newspaper office in Morristown, New Jersey, and look over the type cases. He found a stock of **twelve thousand** E's, **nine thousand** T's, and only **two hundred** Z's. He and Morse rearranged the alphabet accordingly.

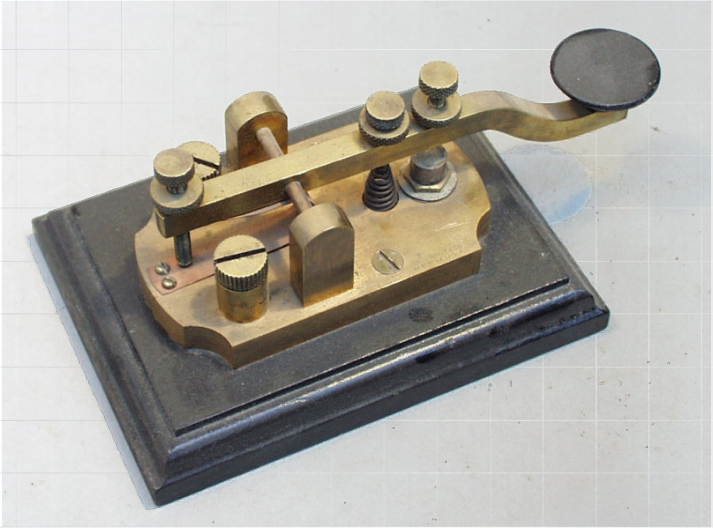
Drums that talk

Morse code

A	•–	J	•–––	S	•••	1	•———–
B	–•••	K	–•–	T	–	2	••–––
C	–•–•	L	•–••	U	••–	3	•••––
D	–••	M	––	V	•••–	4	••••–
E	•	N	–•	W	•––	5	•••••
F	••–•	O	–––	X	–••–	6	–••••
G	––•	P	•––•	Y	–•––	7	––•••
H	••••	Q	––•–	Z	––••	8	–––••
I	••	R	•–•	0	–––––	9	––––•

Drums that talk

Morse key



Drums that talk

- But talking drums do not use such codes!
- In solving the enigma of the drums, the explorers found the key in a central fact about the relevant African languages: They are tonal languages, in which meaning is determined as much by rising or falling pitch contours as by distinctions between consonants or vowels.
- This feature is missing in most Indo-European languages.
- One can imagine how comical the confusion could become:
  - alambaka boli [– _ – – _ _ _] = *he watched the riverbank*
  - alambaka boli [– – – – _ – _] = *he boiled his mother-in-law*

Drums that talk

- A double stroke on the high-tone lip of the drum [– –] matched the tonal pattern of the Kele word:
  - for father, **sango**,
    but naturally, it could just as well be
  - **songe**, the moon;
  - **koko**, fowl;
  - **fele**, a species of fish;

  or any other word of two high tones. Even the limited dictionary of the missionaries at Yakusu contained 130 such words.

Drums that talk

A drummer would invariably add 'a little phrase' to each short word.

- Songe, the moon, is rendered *as songe li tange la manga*—"the moon looks down at the earth."
- Koko, the fowl, is rendered *koko olongo la bokiokio*—"the fowl, the little one that says kiokio."

Extra drumbeats, far from being extraneous, provide context.

Wrocław University of Science and Technology

Drums that talk

- Every ambiguous word begins in a cloud of possible alternative interpretations; then the unwanted possibilities evaporate. This takes place below the level of consciousness.
- A birth announcement in Bolenge, a village in the Belgian Congo, went as follows: *The mats are rolled up, we feel strong, a woman came from the forest, she is in the open village, that is enough for this time.*
- This is an example call to a funeral of a fisherman: *In the morning at dawn, we do not want gatherings for work, we want a meeting of play on the river. Men who live in Bolenge do not go to the forest, do not go fishing. We want a meeting of play on the river, in the morning at dawn.*

The message

- Let us come back to our couple. After the examination, they got the answer: It will be... (And remember: they got exactly **one bit** of information!)

The message

- Let us come back to our couple. After the examination, they got the answer: It will be... (And remember: they got exactly **one bit** of information!)
- So, they resend this message to their parents or/and friends...

The message

- Let us come back to our couple. After the examination, they got the answer: It will be... (And remember: they got exactly **one bit** of information!)
- So, they resend this message to their parents or/and friends...
- It is enough to send 0 (girl) or 1 (boy), but instead, they write: **boy** or **girl** (and sometimes even more: *We will have the son/the daughter*).

The message

- Let us come back to our couple. After the examination, they got the answer: It will be... (And remember: they got exactly **one bit** of information!)
- So, they resend this message to their parents or/and friends...
- It is enough to send 0 (girl) or 1 (boy), but instead, they write: **boy** or **girl** (and sometimes even more: *We will have the son/the daughter*).

Why?

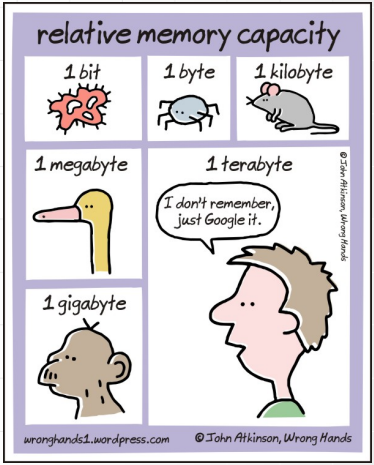
The message

- Let us come back to our couple. After the examination, they got the answer: It will be... (And remember: they got exactly **one bit** of information!)
- So, they resend this message to their parents or/and friends...
- It is enough to send 0 (girl) or 1 (boy), but instead, they write: **boy** or **girl** (and sometimes even more: *We will have the son/the daughter*).

Why?

Homework

relative memory capacity



from http://wronghands1.wordpress.com/2013/11/01/relative-memory-capacity/

Information redundancy

Hello world

How do you do?

Can you read this?

'Cdnuolt blveiee taht I cluod aulaclty uesdnatnrd waht I was rdanieg. The phaonmneal pweor of the hmuan mnid, aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it dseno't mtaetr in waht oerdr the ltteres in a wrod are, the olny iproamtnt tihng is taht the frsit and lsat ltteer be in the rghit pclae. The rset can be a taotl mses and you can sitll raed it whotuit a pboerlm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe. Azanmig huh? yaeh and I awlyas tghuhot slpeling was ipmorantt ... '

What is data compression?

I have created the file 18000 bytes long that contains 2000 lines of the form
boy girl
Next, I have tried to compress this file using some compression tools.

- gzip — 93 bytes
- xz — 132 bytes
- rar — 115 bytes
- 7zip — 181 bytes

Can you explain this?

What is a compression (cont)?

Next, the file of the same length (18000) is filled with random letters. And now:

- gzip — 14541 bytes
- xz — 14580 bytes
- rar — 14884 bytes
- 7zip — 14631 byres

What is a compression (cont)?

Next, the file of the same length (18000) is filled with random letters. And now:

- gzip — 14541 bytes
- xz — 14580 bytes
- rar — 14884 bytes
- 7zip — 14631 byres

In the random text, there is almost no redundancy, so there is nothing to compress.

Wrocław University of Science and Technology

Compression

The compression method used to compress these example files is lossless. The process is reversible, and you can always recover the original file.

- The idea of lossless compression is as follows: every repeated text is replaced by a shorter one.
- "boy" becomes α.
- "girl" — β.
- Space and new line characters are left untouched.
- Instead of 9 characters, we have only four: $\alpha_{\sqcup}\beta$<nl>
- The sequence of the same words is once again replaced by a new short word plus information about the number of repetitions.

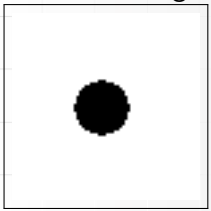
Wrocław University of Science and Technology

Lossy compression

Here is a very simple picture: black dot on white background.

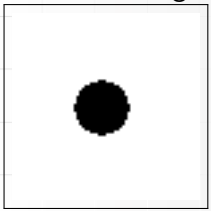
Lossy compression

Here is a very simple picture: black dot on white background.



Lossy compression

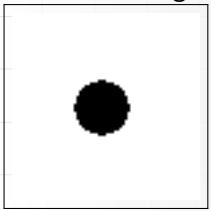
Here is a very simple picture: black dot on white background.



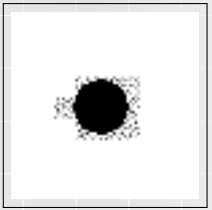
Next, the picture was converted to JPG, and all non-white dots have been boosted to black for better visibility.

Lossy compression

Here is a very simple picture: black dot on white background.

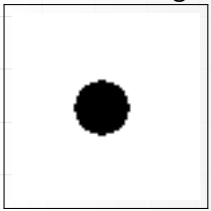


Next, the picture was converted to JPG, and all non-white dots have been boosted to black for better visibility.

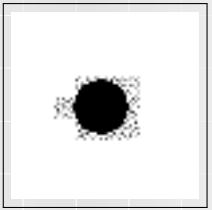


Lossy compression

Here is a very simple picture: black dot on white background.



Next, the picture was converted to JPG, and all non-white dots have been boosted to black for better visibility.



The number of errors depends on compression level: „better compression" means smaller file and more errors.

Real example

On the next two slides, there are two pictures sized 2600×1733 pixels that contain about 4.5 million pixels.
The first one is an original with a size of 238410 (2.3M) bytes.
The second one is processed with a high level of compression. Its size is 94980 bytes (96K).
You can use these images to compare distortions introduced during compression.
Hint: Use zoom.

Wrocław University of Science and Technology

Real example

Real example

Lossy vs lossless compression

- graphics
  - JPG, JPEG
  - DjVu
  - Wavelet
- video
  - MPEG [124]
  - H.264
  - Ogg Theora
- music
  - MP3
  - AAC
  - WMA

- General
  - RLE (BMP)
  - DEFLATE (ZIP, PNG)
- graphics
  - PNG
  - GIF
- video
  - QuickTime
  - H.264 loosless
- music
  - FLAC
  - ALAC (Apple Loosles)