Introdution to medical informatics. Data and information. 1/23

Shiray Andrey

O.O.Bogomolets National Medical University

December 2, 2011

Outline

- Course overview
- Medical informatics. Why? What for?
- Fundamental role of informatics in modern science.
- Information. Data. Knowledge.

Course overview

- Information. Data. Knowledge
- Biostatistics
- Formal Logic. Expert and decision support systems in medicine
- Evidence-based medicine

Medical informatics – the application of informatics to the field of biology and medicine/healthcare.

Informatics, Computers science

- **Informatics** – the study of the processing, management, and retrieval of information
- **Computers science** – is the study of the theoretical foundations of information and computation and of practical techniques for their implementation and application in computer systems
- **Cybernetics** is the interdisciplinary study of the structure of regulatory systems.

Medical informatics. Why? What for?

Medical informatics and borrows concepts/ideas from:

- Medicine
- Informatics
- Logic
- Statistics

Fundamental role of informatics in modern science

- Any object in physical reality can be described with some information

Fundamental role of informatics in modern science

- Any object in physical reality can be described with some information
- Any physical process can be described as informational process
- Physical laws $\sim$ program/algorithm

Fundamental role of informatics in modern science

- Any object in physical reality can be described with some information
- Any physical process can be described as informational process
- Physical laws $\sim$ program/algorithm
- Universe is a gigantic computer!
  - Seth Lloyd, Programming the Universe
  - Roger Penrose, The Emperor's New Mind

Fundamental role of informatics in modern society

That's why informatics is an essential part of science disciplines in any University.
You will need basic informatics skills in modern society:

- We live in so called information society. (or at least tend to develop it)
- You may need to analyze systems/products
- You may need apply informatics skills in your research projects
- You may need to implement prototypes
- You may need to oversee programmers in your future job

Fundamental concepts of informatics

- Information
- Computation
- Data
- Knowledge

Information – primitive notion[1] of modern science.

So it is not defined formally in terms of previously defined concepts, but is only motivated informally, usually by an appeal to intuition and everyday experience

[1] fundamental undefined concept; primitive, undefined, term

Properties

Properties of information:

- **Objectivity** – information doesn't depend upon storage method or your opinion
- **Authenticity** – how trustworsy it is
- **Accuracy**
- **Actual**

Measures of Information. Shannon's measure

In *information theory*, **entropy** is a measure of the uncertainty associated with a random variable(in our case received message)

Shannon's measure[2]: $I(x) = \log(\frac{1}{P(x)}) = -\log(P(x))$

Example:
On tossing a coin, the chance of 'tail' is 0.5. When it is proclaimed that indeed 'tail' occurred, this amounts to $I('tail') = log_2(\frac{1}{0.5}) = \log_2 2 = 1$ bits of information.

[2] also caled self-information

Data

Data refers to qualitative or quantitative attributes of a variable or set of variables.

- For example – results of measurements: $(35.7, 36.7, 42, 38, 39.4, 37.5, 40.1)$
- Raw data is the lowest level of abstraction. Raw data on its own carries no meaning.

Data

Data refers to qualitative or quantitative attributes of a variable or set of variables.

- For example – results of measurements:
  $(35.7, 36.7, 42, 38, 39.4, 37.5, 40.1)$
- Raw data is the lowest level of abstraction. Raw data on its own carries no meaning.
- Raw data needs some way of interpretation.

Data

Data refers to qualitative or quantitative attributes of a variable or set of variables, stored in some way.

- For example – results of measurements: $(35.7, 36.7, 42, 38, 39.4, 37.5, 40.1)$
- Raw data is the lowest level of abstraction. Raw data on its own carries no meaning.
- Raw data must be interpreted and take on a meaning
- Example: $(35.7, 36.7, 42, 38, 39.4, 37.5, 40.1)$ – results of patients thermometrics

Knowledge

Knowledge is a familiarity with someone or something, that can include facts/data, descriptions, information, and/or skills acquired through experience or education.

Main difference between knowledge and data: knowledge has dynamic nature by itself, it can change via learning. This the highest level of information abstraction.

Computation

- **Computation** – any type of information processing.

Computation

- **Computation** – any type of information processing.
- A computation can be seen as a purely physical phenomenon occurring inside a closed physical system called a computer.

Computation

- **Computation** – any type of information processing.
- A computation can be seen as a purely physical phenomenon occurring inside a closed physical system called a computer.
- To describe a computation we need some formal way to state rules for information processing - **algorithm**.

Algorithm

Algorithm is an effective method expressed as a finite list of well-defined instructions for computation.

Starting from an initial state and initial input (perhaps empty), the instructions describe a computation that, when executed, will proceed through a finite number of well-defined successive states,using finite amount of resources, eventually producing "output"

Properties of Algorithms

- **Finiteness** – An algorithm must always terminate after a finite number of steps
- **Definiteness** – Each step of an algorithm must be precisely defined; the actions to be carried out must be rigorously and unambiguously specified for each case
- **Discrete** – moving in separate(discrete) steps: 1, 2, 3, ...
- **Input**
- **Output**