

# Information science and health education

Szeged 2011

# How informatics can contribute to health education?

- Technical layer (distance education, digital anatomy, hands on practice etc)
- Conceptual layer (how to represent medical knowledge)

# Knowledge representation why it matters?

- Information explosion
- Narrative vs. formal description
- Representation is purpose dependent.
- Fast
  - learning,
  - understanding,
  - problem solvingrequires suitable representation

# An example – narrative representation

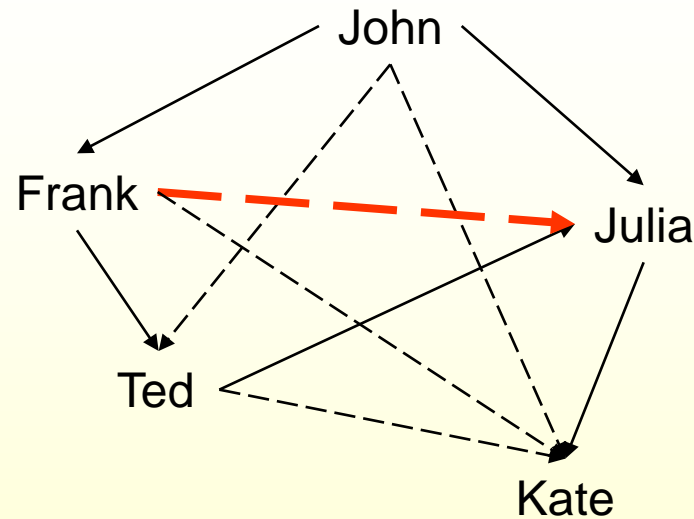
This is a narrative representation of a number of facts. This illustrates the usual way, how medical knowledge used to be described in medical textbooks:

*"John is taller than Julia, Ted is smaller than Frank, Julia is higher than Kate, but John is taller than Frank and Julia is smaller than Ted"*

Is Julia higher than Frank?

# Graph representation of the narrative description

- John is taller than Julia
- Ted is smaller than Frank
- Julia is taller than Kate
- John is taller than Frank
- Julia is smaller than Ted



# Simplified graph representation

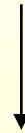
John



Frank



Ted



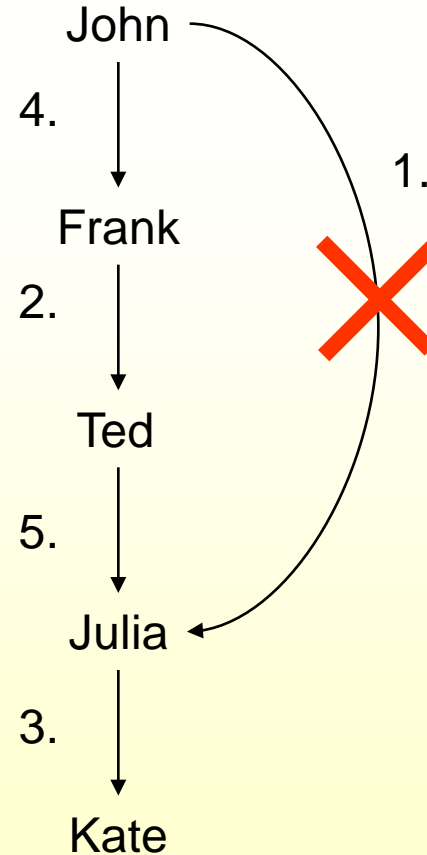
Julia



Kate

# Simplified graph representation

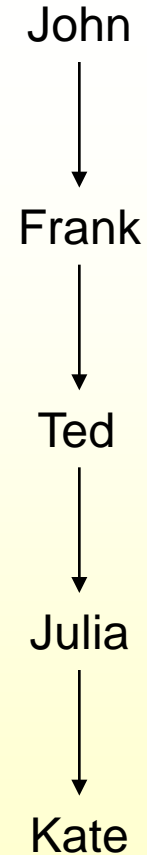
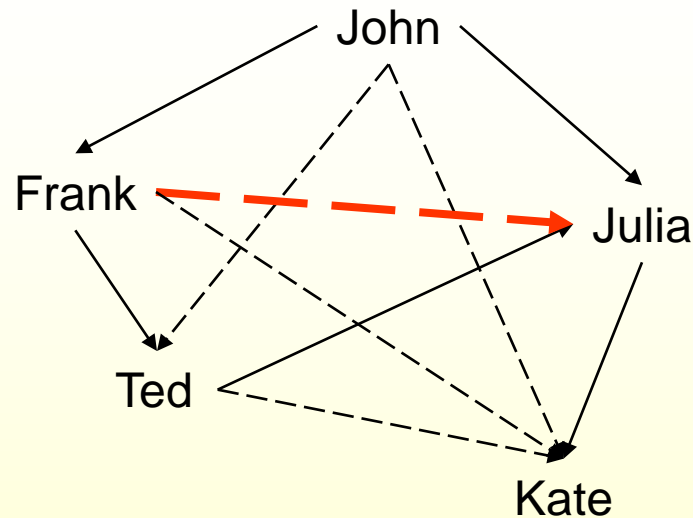
1. *John is taller than Julia*
2. Ted is smaller than Frank
3. Julia is taller than Kate
4. John is taller than Frank
5. Julia is smaller than Ted





# Three representation compared

- John is taller than Julia
- Ted is smaller than Frank
- Julia is taller than Kate
- John is taller than Frank
- Julia is smaller than Ted



# A medical example

- Features of human pathogen bacteria is described in traditional text books.
- It is possible to create a database that stores all described examinable feature for all species
- Such a database revealed that all features of some pairs of species described in different part are totally identical
- Using conventional representation it was impossible to discover that these are the same species described under different names.

# What formal representation is about?

- Formal language:
  - Defined and finite set of symbols
  - Syntactic (typographic) rules
- Formal system
  - Based on some formal language
  - Basic set of "Canonic" statements or axioms
  - Transformation (inference) rules that leads to further canonic statements (theorems)

# Medicine as empirical science

- Empirical sciences thought to be narrative and non-axiomatic
- Yes, but there experiences must be described comparably.
- Comparable descriptions depend on:
  - Unambiguous conceptual system (ontology): What things are?
  - Comparable data structures: What features of a certain thing should be described and how?

# What makes descriptions comparable?

Imagine you want to buy a car.

You study various catalogues describing:

Type of fuel, engine power, capacity (number of persons), colour etc.

If different catalogues describe

- different features
- same features with different terms

You cannot make any comparison.

# A model for comparable descriptions

<car> has\_colour<red>

<car> has\_colour<white>

In general:

<entity>has\_feature<value>

Feature set

Value set for each feature

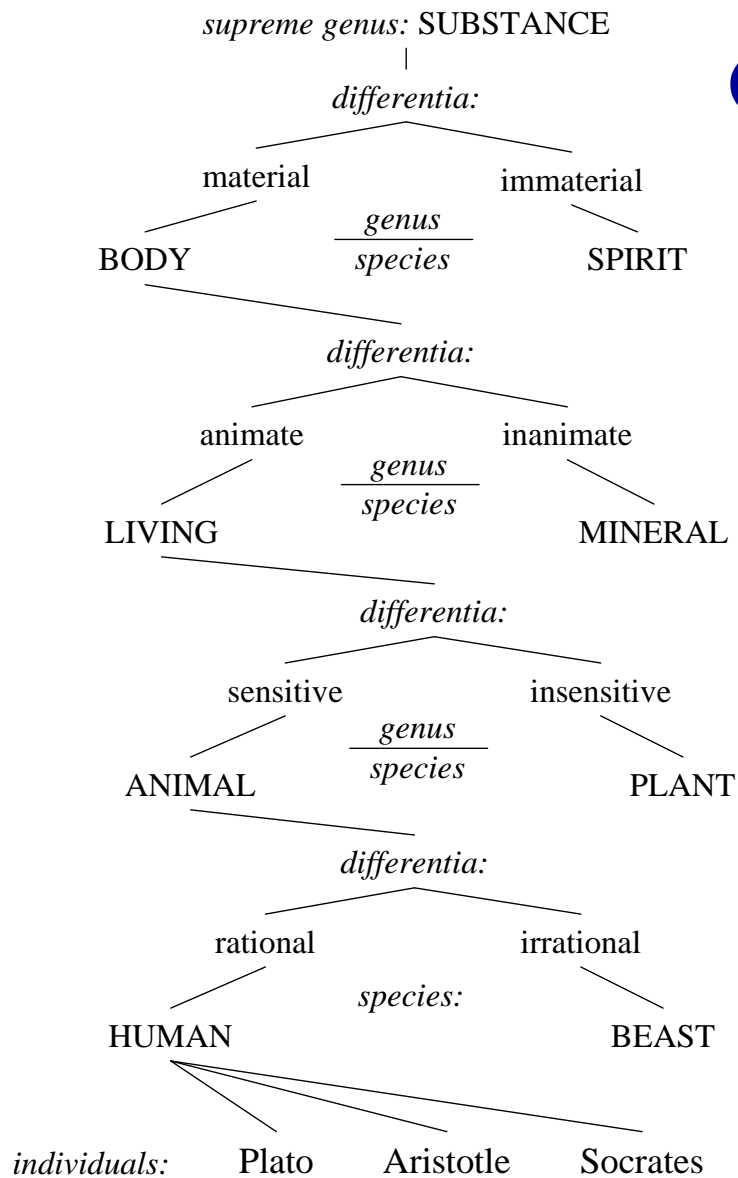
# From comparable descriptions to formal systems

- Unambiguous conceptual systems
    - Defined categories
    - Unique symbols
- } Defined, finite set of symbols
- Comparable data structures
    - Fixed syntactic rules

# Axiomatic vs. empirical sciences

- E.g. Math: Limited number of axioms  
(Euclidean geometry: 5 axioms, 5 inference rules, 23 basic concepts – point, surface etc.)
- Life sciences: high number of axioms, basic concepts, few theorems, few derived definitions, no domain specific inference rules.





# Categories and instances

The Aristotelian definition of categories:

'Genus proximum' – the immediate superior category (*A is a B, ...*)

+

'Differentia specifica' – the differentiating properties (*..., which is ...*)

E.g. *Animal is a living being, which is sensitive*

# Two valued (bivalent) vs. fuzzy logic

- Aristotelian logic:
  - All statements are either true or false
  - No statement can be true and false at the same time
  - Nothing is in between true and false
- Fuzzy logic: all statements has a truth value. For "absolutely" true statements this value is 1, "absolutely" false statements is 0.  
(similar to likelihood function)

# Aristotelian definitions

Based on the hierarchy of the categories things should be defined by determining the '*genus proximum*' i.e. the immediate superior category and the '*differentia specifica*' i.e. the feature that discriminates the given entity from other things within the same superior category

It is impossible to define all categories in a domain. All system of definitions must start from a number of undefined 'primitive' categories.

# Aristotelian categories

For any entity  $x$  and category  $\mathbf{A}$ , the truth value of the statement:

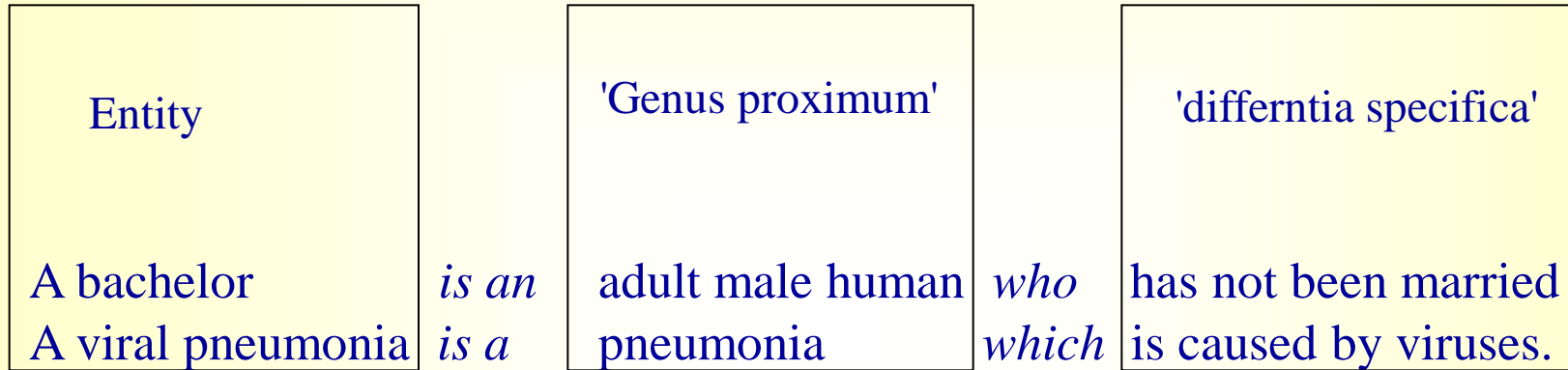
$x$  is an instance of  $\mathbf{A}$

- can either be true or false,
- can not be true and false at the same time and
- can not be something in between true and false

A bachelor is an adult male human      who has not been married  
A viral pneumonia is a pneumonia      which is caused by viruses.

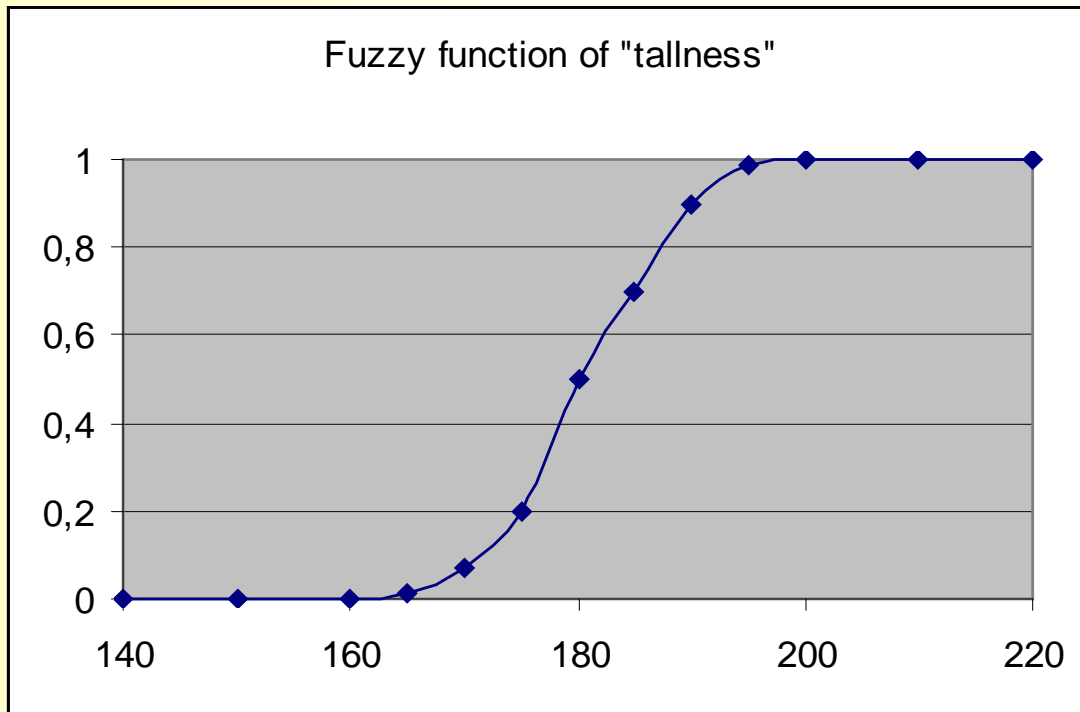
*John has viral pneumonia (at a certain time point) .T. OR .F.*

## Typical Aristotelian categories:



# Fuzzy categories

For any entity  $x$  and category  $\mathbf{A}$ , the truth value of the statement:  
 $x$  is an instance of  $\mathbf{A}$   
has a fuzzy function value between 0 and 1.



If 0 is interpreted as 'false',  
1 as 'true' than fuzzy categories violate the Aristotelian rules.

"John (185 cm) is a tall man". This statement has a truth value 0.7 something in between true and false

But this is a matter of interpretation:

*"The 'tallness' of John is 0.7 "* .T. OR .F.

is an other expression of the same statement. This statement can be either true or false, and satisfies the Aristotelian rules.

The conflict between Aristotelian and fuzzy logic can be more or less resolved by such rephrasing of the statements.

# Everyday categories that are hard to define

Furniture  
Salad  
Cloth  
Food  
Toy  
Fruit  
etc.

Humans can rate entities as better or worse examples. The best examples are the *PROTOYPES* of the given category (E. Rosch)

PROTOTYPICAL CATEGORIES



# Everyday categories that are hard to define

Wittgenstein's example: game

There is no common exclusive property that is characteristic to all and only games.

For any game there is an other game such that they share some common property.

Everything is a game that resembles to (has sufficient number of common properties) with another game.

FAMILY RESEMBALNCE

# An example

"Autism is a disorder which *usually* appears within the first three years of life and *may* result in learning difficulties, speech problems and difficulty relating to people."

This statistically provable statement. Assuming that there is no other disorder with these criteria, the statement fulfils the criteria of an Aristotelian definition:

'genus proximum' – autism is a disorder

'differentia specifica'

- usually appears within the first three years of life AND
- may result in learning difficulties AND
- may result in speech problem AND
- may result in difficulty relating to people

# A medical definition definition

"Autism is a disorder which usually appears within the first three years of life and may result in learning difficulties, speech problems and difficulty relating to people."

Patient XY has a disorder that appeared at her fifth year. She has no learning difficulties, no speech problems and has no difficulty in relating to people.

Relying on the definition above, is it possible to exclude that she is an autistic?

# The answer is NO !

"Autism is a disorder which *usually* appears within the first three years of life and *may* result in learning difficulties, speech problems and difficulty relating to people."

XY has a disorder

which USUALLY but not always appears within the first three years..

and MAY but not necessarily result in...

# Summing up

- Conventional medical knowledge can be described in formal languages
- The logic behind comes from philosophy and mathematics, but IT makes formal medical descriptions manageable (high number of axioms and 'primitive' categories)

# Conclusions

- Traditional medical descriptions can be represented in a number of different ways
- Different representation have different advantages
- Modern IT offers new ways of representation of medical knowledge
- It is the time now for re-engineering of medical knowledge.
- This is a task for new generation of physicians