How to test a philosophical theory empirically

Barry Smith

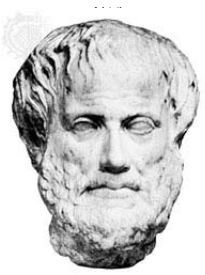
http://ifomis.de

Ontology as a branch of philosophy

the science of the kinds and structures of objects, properties, events, processes and relations in every domain of reality

Ontology a kind of generalized chemistry or zoology

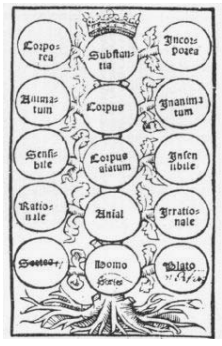
(Aristotle's ontology grew out of biological classification)



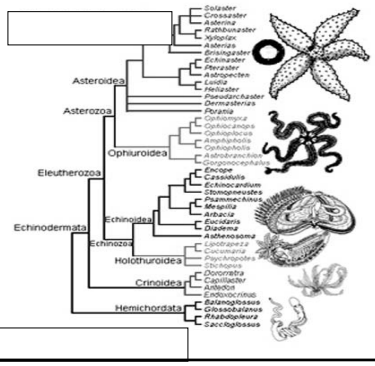
world's first ontologist

First ontology

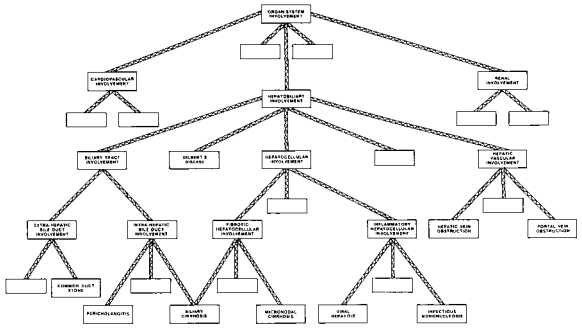
(from Porphyry's *Commentary* on Aristotle's *Categories*)



Linnaean Ontology



Medical Diagnostic Hierarchy



Ontology is distinguished from the special sciences

it seeks to study *all* of the various types of entities existing at all levels of granularity

and to establish how they hang together to form a single whole ('reality' or 'being')

Sources for ontological theorizing:

the study of ancient texts

thought experiments (we are philosophers, after all)

the development of formal theories

the results of natural science

now also:

working with computers

The existence of computers

and of large databases

allows us to express old philosophical problems in a new light

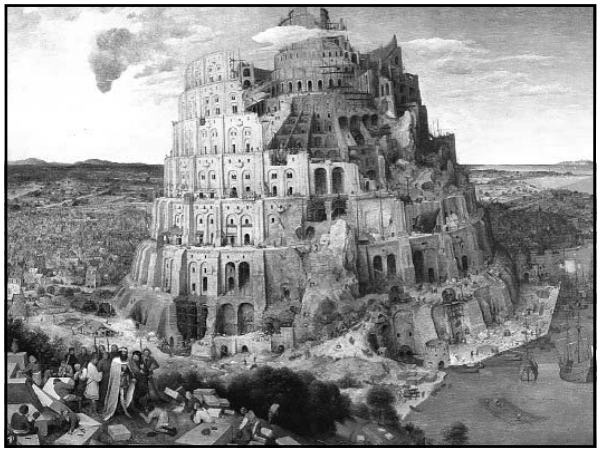
Example:
The problem of the unity of science

The logical positivist solution to this problem addressed a world in which sciences are identified with

printed texts

What if sciences are identified with

Large Databases ?



The Database Tower of Babel Problem

Each family of databases has its own idiosyncratic terms and concepts

by means of which it represents the information it receives

How to resolve the incompatibilities which result when databases need to be merged?

Compare: how to unify biology and

The term ‘ontology’

now used by information scientists to describe the building of standardized taxonomies

which are designed to make databases mutually compatible

An ‘ontology’

is a dictionary of terms formulated in a canonical syntax and with commonly accepted definitions and axioms

How has this idea been realized?

How have information systems engineers built ontologies?

From where did they take the term ‘ontology’?

From Quine

… for Quineans, the ontologist studies, not reality,

but *scientific theories*

the study of ontology is confused with the study of Ontological Commitment

Quine:

each natural science has its own catalogue of types of objects to the existence of which it is committed

Quineanism:

ontology is the study of the ontological commitments or presuppositions embodied in the different natural sciences

In the hands of information scientists

this is transformed into the view that ontology should study

the concepts people use

Ontology becomes

the study of concept systems

Arguments for Ontology as Conceptual Modeling

Ontology is hard.

Life is short.

Let's do conceptual modeling instead

Ontological engineers

thus neglect the standard of *truth* in favor of other, putatively more practical standards:
above all programmability

For an information system ontology

there is no reality other than the one created through the system itself,

so that the system is, by definition, correct

Ontological engineering

concerns itself with concept systems

It does not care whether these are *true of* some independently existing reality.

‘Ontology’

is tremendously popular in information systems research today

… b u t

ATTEMPTS TO SOLVE THE TOWER OF BABEL PROBLEM VIA ONTOLOGIES AS CONCEPTUAL MODELS HAVE FAILED

To see why

let us consider some examples of concept systems in the medical domain

Attempts at such standardization include:

1. UMLS
2. SNOMED
3. GALEN

Example 1: UMLS

Universal Medical Language System

Very large taxonomy maintained by National Library of Medicine in Washington DC

Example 1: UMLS

134 semantic types

800,000 concepts

10 million interconcept relationships

UMLS is the product of fusion of several source vocabularies

(built out of concept trees)

Example 2: SNOMED-RT

Systematized Nomenclature of Medicine

A Reference Terminology with Legal Force

Example 2: SNOMED-RT

121,000 concepts,

340,000 relationships

"common reference point for comparison and aggregation of data throughout the entire healthcare process"

Problems with UMLS and SNOMED

Each is a 'fusion' of several source vocabularies

They were fused without an ontological system being established first

They contain circularities, taxonomic gaps, and unnatural ad hoc determinations

Example 3: GALEN

Generalised **A**rchitecture for

Languages, **E**ncyclopaedias and

Nomenclatures in Medicine

Applied especially to surgical procedures

Problems with GALEN

Ontology is ramshackle and has been subject to repeated fixes

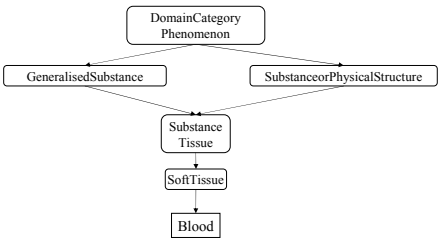
Unnaturalness makes coding slow and expensive, hence narrow scope

Not gained wide acceptance

Blood

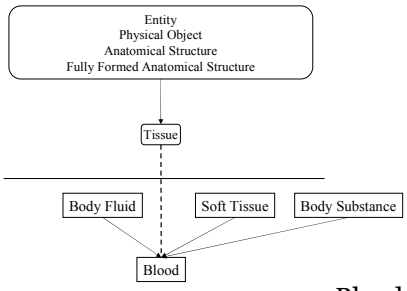
with thanks to Anita Burgun and Olivier Bodenreider

Representation of Blood in GALEN



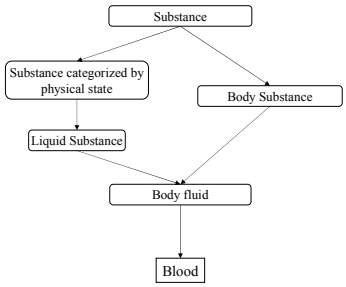
Blood has two states, LiquidBlood and CoagulatedBlood

Representation of Blood in UMLS



Blood as tissue

Representation of Blood in SNOMED



Blood as Fluid

How make ONE SYSTEM out of this?

To reap the benefits of standardization we need to resolve such incompatibilities?

But how?

***Not* just by looking at the concepts underlying the respective systems**

For how, just by looking at separate concepts, could we establish how these concept systems relate to each other?

different conceptual systems

need not interconnect at all

Database standardization

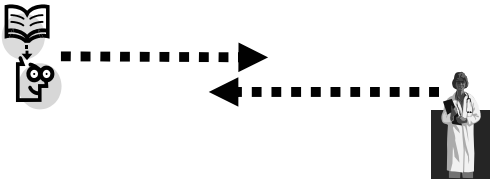
is desparately needed in medicine

… to enable the huge amounts of data

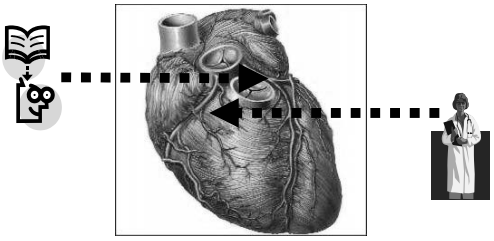
resulting from clinical trials by different groups working on the same drugs/therapies/diagnostic methods

…to be *fused together*

the only way to make them interconnect



is by looking *not just at concepts*



but also at the *reality* beyond

How to solve the Tower of Babel Problem

Look not at concept systems alone

but at how concept systems relate to the world beyond

Concept systems which are transparent to reality

have a reasonable chance of being integrated together into a single ontological system

This means we need to return to the traditional view of ontology

… as a maximally opportunistic theory of reality

Maximally opportunistic

means:

don't just look at concepts

look at the objects themselves

towards which such concepts are directed

… look at the objects

from every available direction

- **both formal and informal**
- **scientific and non-scientific**
- **empirical and theoretical**

attempting always to establish how these objects hang together ontologically

Maximally opportunistic

means:

- **look at concepts *critically***
- **and always in such a way as to**
- **include independent ways to access the objects themselves**

How to test a philosophical theory empirically?

IFOMIS

Institute for Formal Ontology and Medical Information Science

Faculty of Medicine

University of Leipzig

IFOMIS

in collaboration with those groups of ontological engineers who have recognized that they can improve their methods

by drawing on the results of the philosophical work in ontology carried out over the last 2000 years

… above all:

LADSEB, Padua/Trento

ITBM-CNR, Rome

ONTEK Corporation, Georgia

Language and Computing EV, Belgium

It will develop medical ontologies

at different levels of granularity:

- **cell ontology**
- **drug ontology ***
- **protein ontology**
- **gene ontology ***

*** = already exists (but in a variety of mutually incompatible forms)**

and also

- **anatomical ontology ***
- **epidemiological ontology**
- **disease ontology**
- **therapy ontology**
- **pathology ontology ***

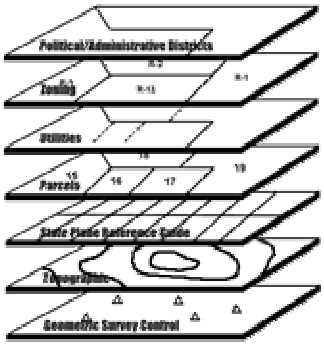
together with

physician's ontology

patient's ontology

and even

hospital management ontology *



Ontology

like cartography

must work with maps at different scales

How fit these maps (conceptual grids) together into a single system?

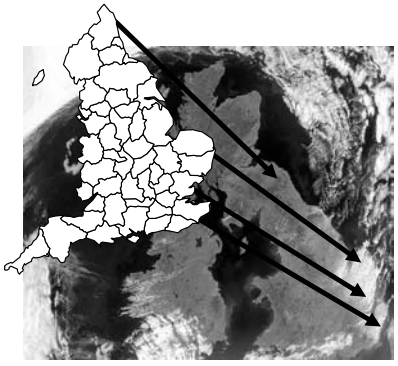
Consider them

as grids

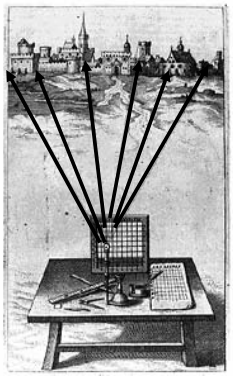
transparent to reality

allowing our directedness towards objects beyond

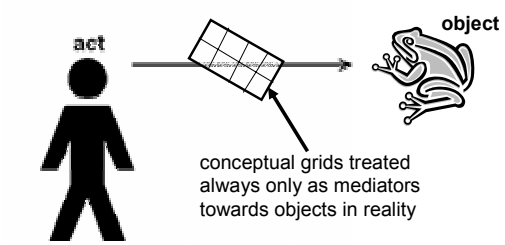
Cartographic Projection



Optical Projection



***intentionality* = the directedness towards objects via conceptual grids**



Intentional directedness

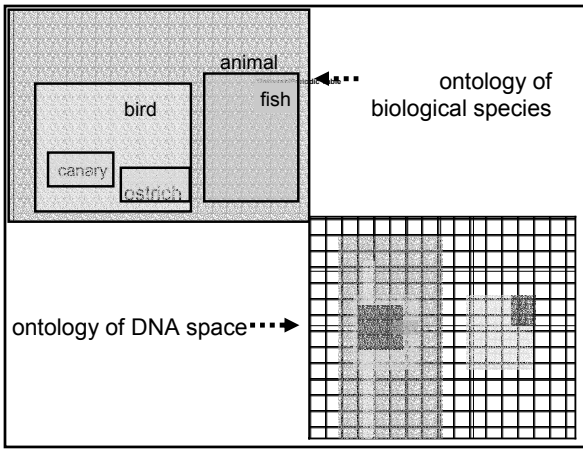
… is effected via conceptual grids

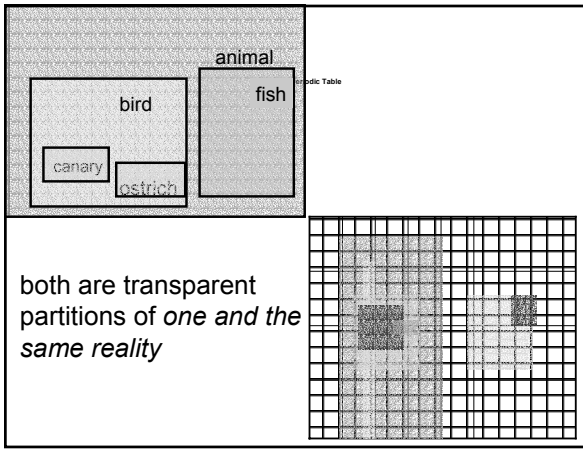
we are able to reach out to the objects themselves *because our conceptual grids are transparent*

there are many compatible map-like partitions

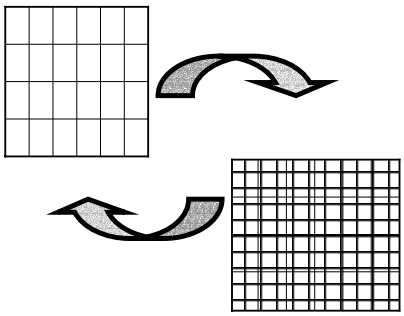
at different scales,

which are all *transparent* **to the reality beyond**





Ontological Zooming



The job of the ontologist

is to understand how different partitions *of the same reality* interrelate

The Tests

Uniform top-level ontology for medicine applicable at distinct granularities

Test-case development of partial medical domain ontologies applied to:

- **Standardization of clinical trial protocols**
- **Clinical trial *Merkmal*-dictionary**
- **Processing of unstructured patient records (www.landc.be)**

The Goals

Uniform top-level ontology for medicine

ONE YEAR

Applicable at distinct granularities (e.g. gene ontology)

FOUR YEARS

Standardization of clinical trial protocol

TWO YEARS

Clinical trial *Merkmal*-dictionary

TWO YEARS

Processing of unstructured patient records (www.landc.be)

THREE YEARS

Measures of Success

Uniform top-level ontology for medicine

NO COMPETITOR

applicable at distinct granularities

NO COMPETITOR

Standardization of clinical trial protocol

NO SERIOUS COMPETITOR

Clinical trial *Merkmal*-dictionary

NO COMPETITOR

Processing of unstructured patient records

MANY COMPETITORS, BUT GOOD MEASURES OF EFFECTIVENESS

The End