Workshop on Ontologies and their Applications,
September 28, 2004, São Luís do Maranhão (Brazil)

Ontological Foundations for Biomedical Sciences

Stefan Schulz, Kornél Markó, Udo Hahn

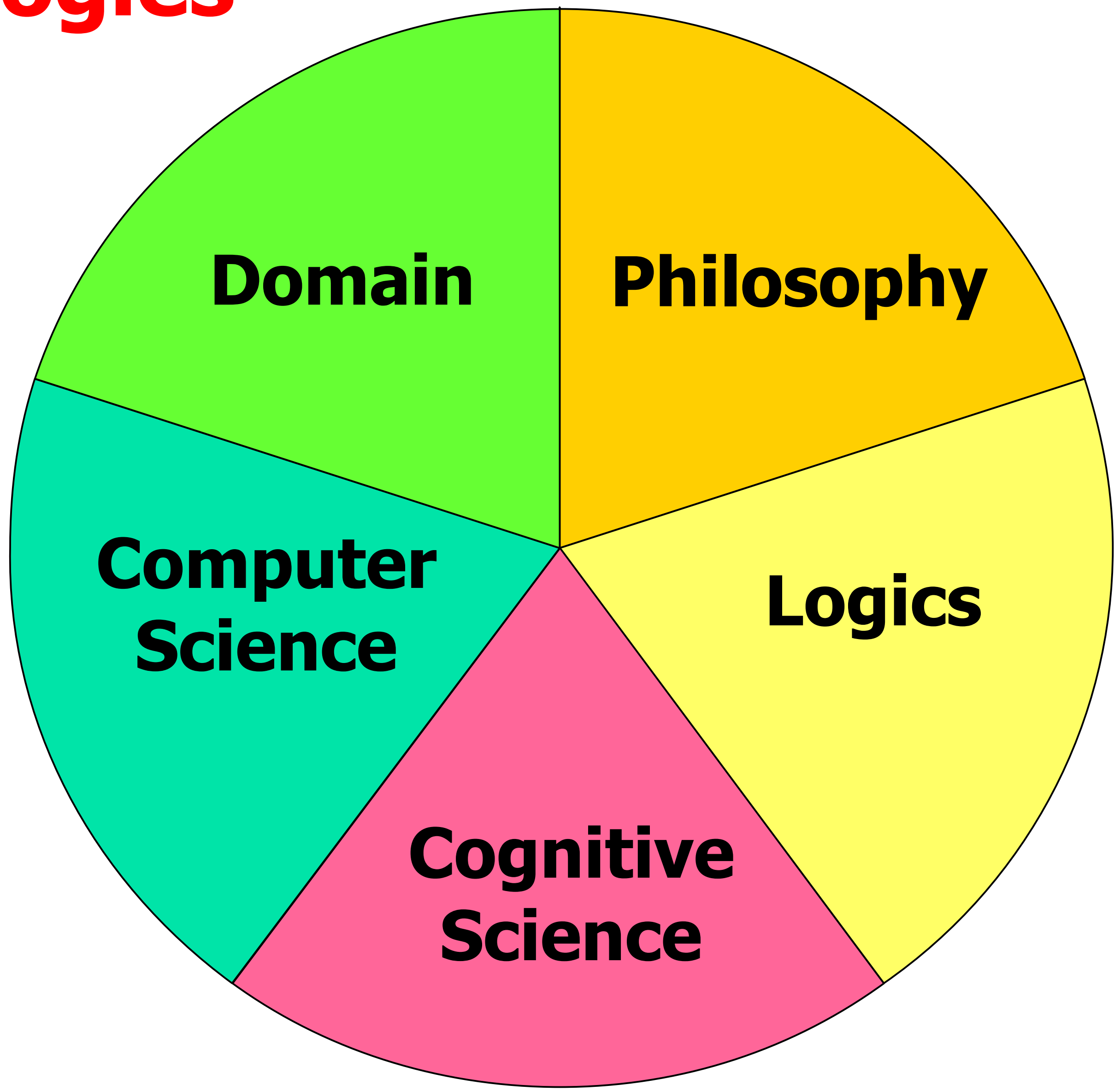
Department of Medical Informatics
University Hospital Freiburg (Germany)

Text Knowledge Engineering Lab
University of Jena (Germany)

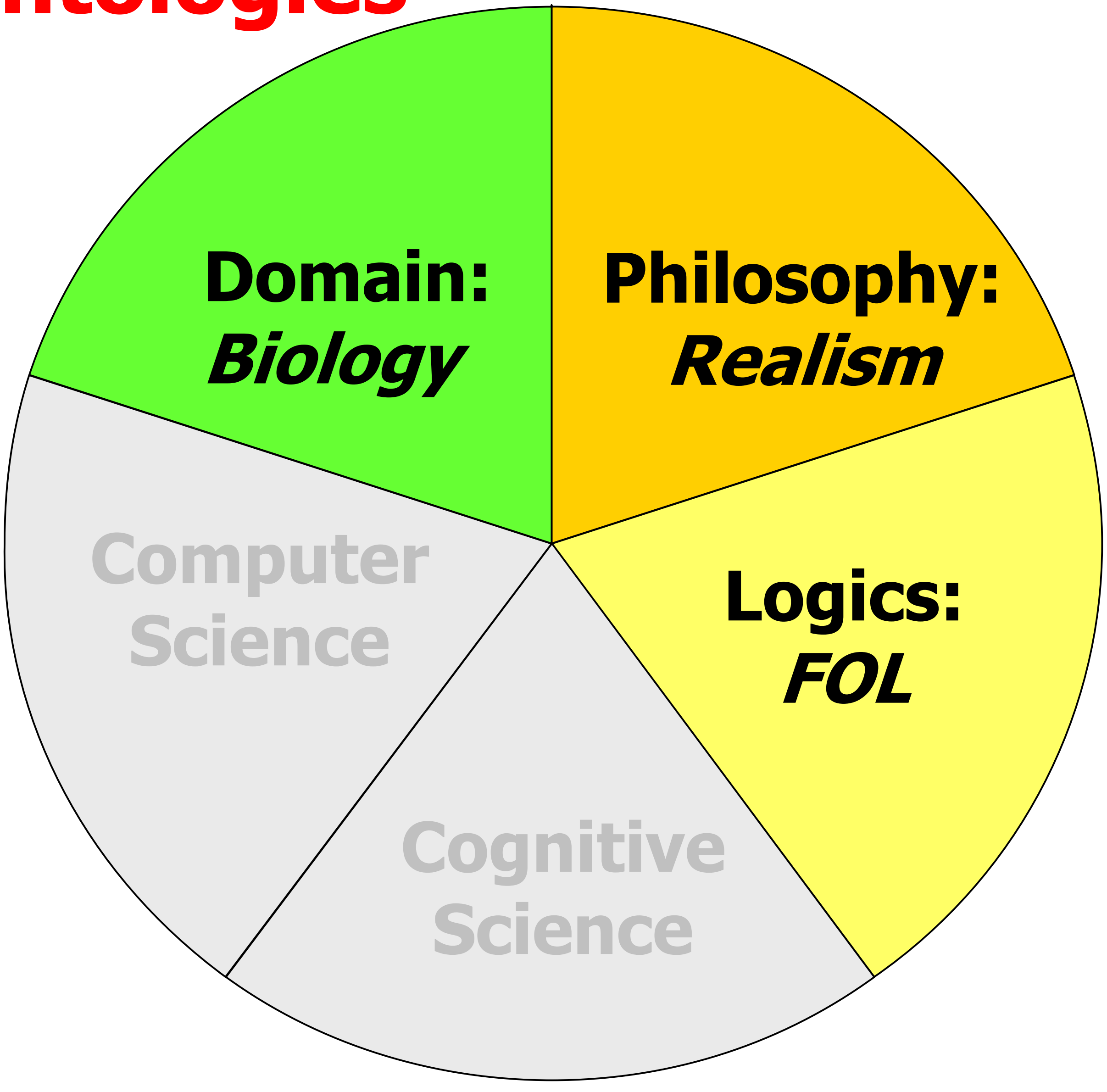
The World of Life Sciences…

…requires sophisticated organization

Ontologies


Domain
Philosophy
Computer Science
Logics
Cognitive Science

Bio-Ontologies



Top-Level Division

Biological Entitiy

Occurrents:

(Changes of) states of affairs of the physical world:

Examples: process, state, event, disease, procedure…

Continuants:

Entities of the physical world *(„Biomedical Structure")*:

Examples:
body, organ, tissue, molecule,..

disjoint

depend on

Representation of Continuants in Bio-ontologies. What exists ?

- Human Anatomy
  - Foundational Model of Anatomy (FMA)
  - Portions of SNOMED, OpenGalen, MeSH
- Other Organisms
  - Open Biological Ontologies (OBO)
    - Mouse (developmental stages), Zebrafish, Drosophila,…
  - UMLS Semantic Network
- Species-Independent
  - Gene Ontology: Cellular Component branch

Size: 14 (UMLS SN) – 10^3 (Adult Mouse) – 10^5 (FMA)

Deficiencies of existing Bio-Ontologies

- Redundancy
- Synonymy
- Ambiguity
- Underspecification

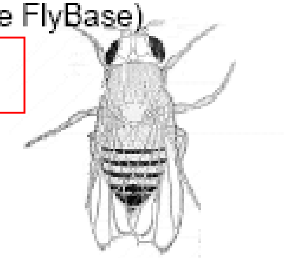
Redundancy

Mouse (embryonal stage TS11, source: MGI)
- cardiovascular system
- - heart
- - - cardiogenic plate

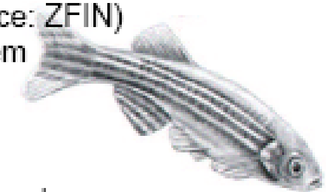
Mouse (embryonal stage TS18 , source: MGI)
- cardiovascular system
- - heart
- - - atrio-ventricular canal
- - - atrium
- - - bulboventricular groove
- - - bulbus cordis
- - - endocardial cushion tissue
- - - mesentery
- - - outflow tract
- - - pericardium
- - - primitive ventricle
- - - sinus venosus

Mouse (embryonal stage TS26 , source: MGI)
- cardiovascular system
- - heart
- - - aortic sinus
- - - atrio-ventricular canal
- - - atrio-ventricular cushion tissue
- - - atrium
- - - bulbar cushion
- - - endocardial cushion tissue
- - - endocardial tissue
- - - mesentery
- - - pericardium
- - - trabeculae carneae
- - - valve
- - - ventricle

Drosophila (adult, source FlyBase)
- circulatory system
- - heart
- - - heart muscle
- - - adult aortic funnel
- - - adult ostia
- - - dorsal diaphragm
- - - heart chamber
- - - terminal opening



Zebrafish (adult, source: ZFIN)
- cardiovascular system
- - heart
- - - atrium
- - - bulbus arteriosus
- - - hypobranchial vessels
- - - sinus venosus
- - - ventricle



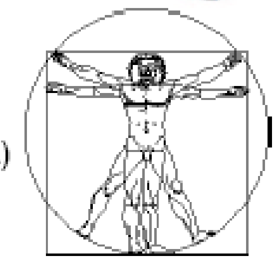
Human, Adult, (source: FMA)
- cardiovascular system
- - heart
- - - wall of heart
- - - right atrium
- - - left atrium
- - - right ventricle
- - - left ventricle

is-a organ chamber

- - - right side of heart
- - - left side of heart
- - - fibrous skeleton of heart
- - - papillary muscle
- - - cardiac valve
- - - tricuspid valve
- - - mitral valve
- - - aortic valve
- - - pulmonary valve

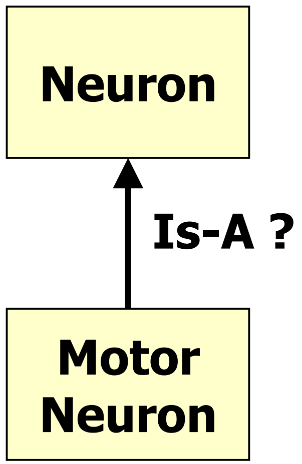
is-a cardiac valve

- - - interatrial septum
- - - (...)



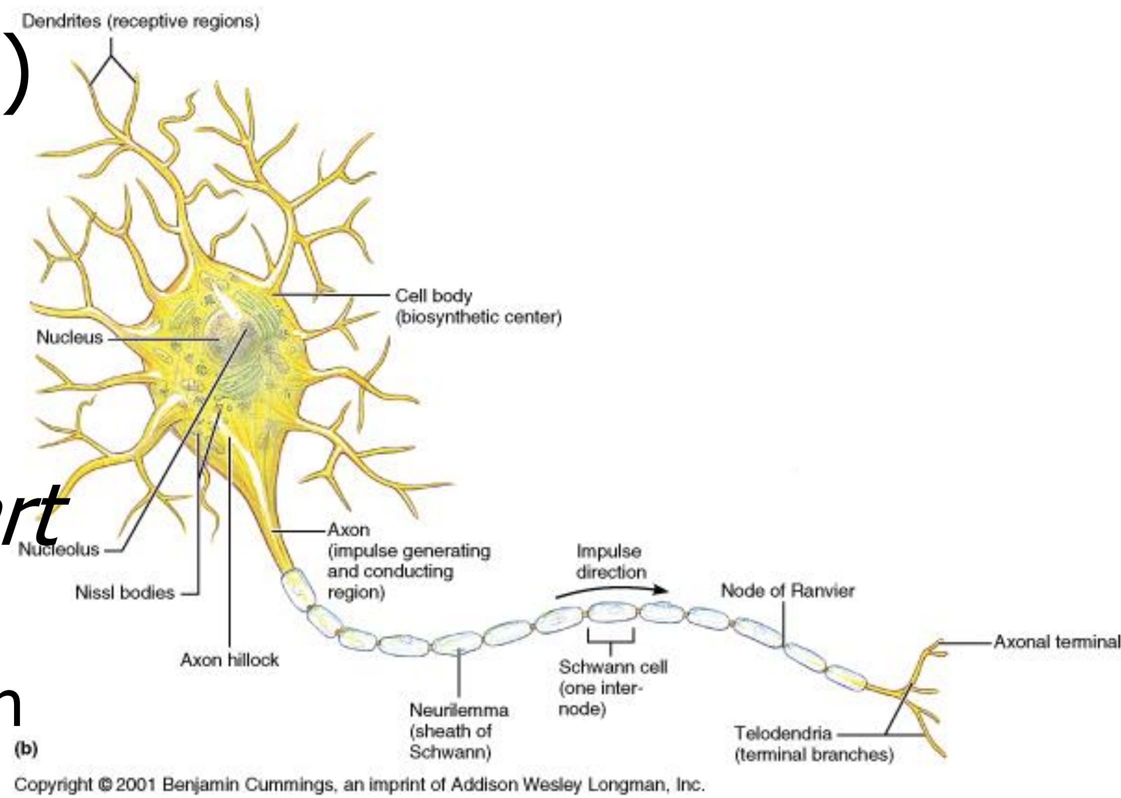
Synonymy

- "*Motor Neuron instance-of Neuron*" (FlyBase)
- "*Motor Neuron narrower Neuron*" (MeSH)
- "*Motor Neuron subclass-of Neuron*" (FMA, OpenGALEN)



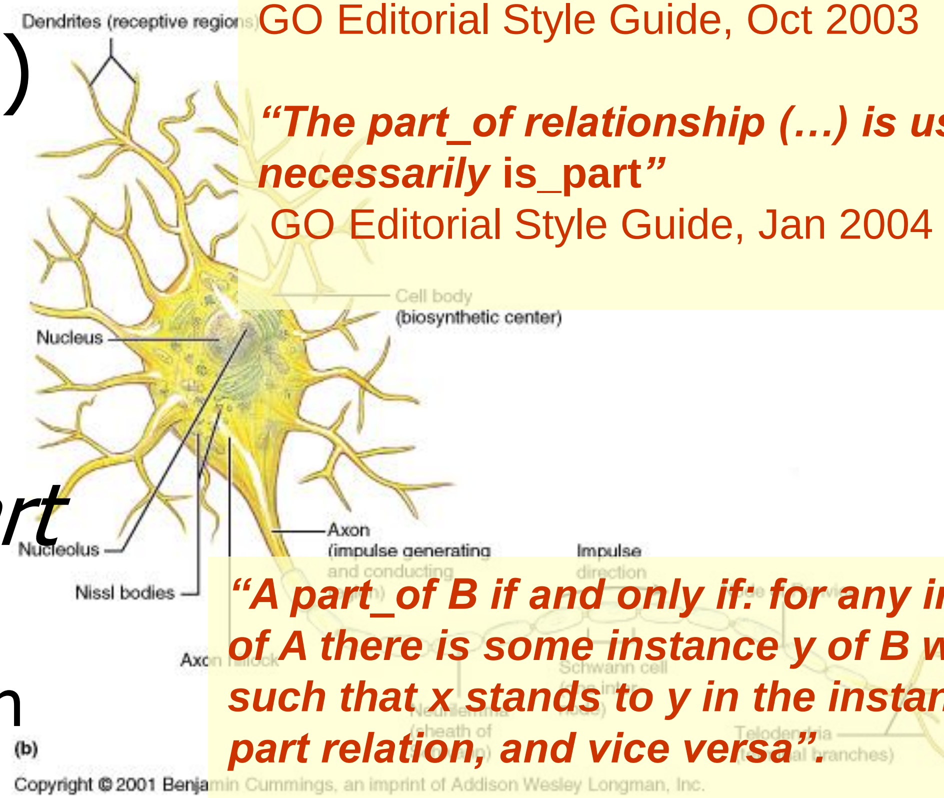
Ambiguity, Underspecification

- "*Cell has-part Axon*" (Gene Ontology)
  - Do cells without axons exist ?
  - Do axons without cells exist ?
- "*Neuron has-part Axon*" (FMA)
  - Does every neuron has an axon?



Ambiguity, Underspecification

- "*Cell has-part Axon*" (Gene Ontology)
  - Do cells without axons exist ?
  - Do axons without cells exist ?
- "*Neuron has-part Axon*" (FMA)
  - Does every neuron has an axon?



"Keep in mind that part_of means can be a part of, not is always a part of "
GO Editorial Style Guide, Oct 2003

***"The part_of relationship (…) is usually necessarily* is_part"**
GO Editorial Style Guide, Jan 2004

"A part_of B if and only if: for any instance x of A there is some instance y of B which is such that x stands to y in the instance-level part relation, and vice versa".

Rosse & Smith MEDINFO 2004

Semantic framework for biological structure…

- Foundational Relations
- General Attributes
- Theories

Semantic framework for biological structure…

- Foundational Relations
- General Attributes
- Theories

Bio-ontologies

	Occurrents (Changes of) states of affairs of the physical world:	**Continuants** Entities of the physical world
Universals **(Concepts, Classes of Individuals)**	Life, Appendectomy, Mitosis	Hand, Blood, Cell, Tree
Particulars **(Concrete Objects in the world)**	My Life, Appedectomy of Patient #123, this Mitosis	My Hand, Blood Sample #12345, this Cell, the Maple Tree in front of the house #xyz

Four disjoint partitions

Some Foundational Relations between Biological Continuants

Rel(x,y) y / x	Universals	Particulars
Universals	*Is-A*	
Particulars	*Instance-of*	*part-of, has-location has-branch, bounds, connects has-developmental-form*

Some Foundational Relations between Biological Continuants

Rel(x,y) y / x	**Universals**	**Particulars**
Universals	*Is-A*	
Particulars	*Instance-of*	*part-of, has-location has-branch, bounds, connects has-developmental-form*

Some Foundational Relations between Biological Continuants

Rel(x,y) y / x	**Universals**	**Particulars**
Universals	*Is-A, Part-Of, Has-Location Bounds, Has-Branch, Connects Has-Developmental-Form*	
Particulars	*Instance-of*	*part-of, has-location has-branch, bounds, connects has-developmental-form*

From Instance-to-Instance relations to Class-to-Class Relations

A, *B* are classes, *inst-of* = class membership

rel: relation between instances *Rel*: relation between classes

$$Rel\ (A, B) =_{def}$$

❶ $\exists x: inst\text{-}of\ (x, A) \wedge inst\text{-}of\ (y, B) \wedge rel\ (x, y)$ **OR**

cf.
Schulz & Hahn (KR 2004, ECAI 2004)
Rosse & Smith (MEDINFO 2004)

From Instance-to-Instance relations to Class-to-Class Relations

A, B are classes, $\textit{inst-of}$ = class membership

rel: relation between instances Rel: relation between classes

$Rel\ (A, B) =_{def}$

❶ $\exists x: \textit{inst-of}\ (x, A) \wedge \textit{inst-of}\ (y, B) \wedge rel\ (x, y)$ **OR**

❷ $\forall x: \textit{inst-of}(x, A) \rightarrow \exists y: \textit{inst-of}\ (y, B) \wedge rel\ (x, y)$ **OR**

cf.
Schulz & Hahn (KR 2004, ECAI 2004)
Rosse & Smith (MEDINFO 2004)

From Instance-to-Instance relations to Class-to-Class Relations

A, B are classes, *inst-of* = class membership

rel: relation between instances *Rel*: relation between classes

$Rel\ (A, B) =_{def}$

➊ $\exists x: inst\text{-}of\ (x, A) \wedge inst\text{-}of\ (y, B) \wedge rel\ (x, y)$ **OR**

➋ $\forall x: inst\text{-}of(x, A) \rightarrow \exists y: inst\text{-}of\ (y, B) \wedge rel\ (x, y)$ **AND**

➌ $\forall y: inst\text{-}of(y, B) \rightarrow \exists x: inst\text{-}of\ (x, A) \wedge rel\ (x, y)$

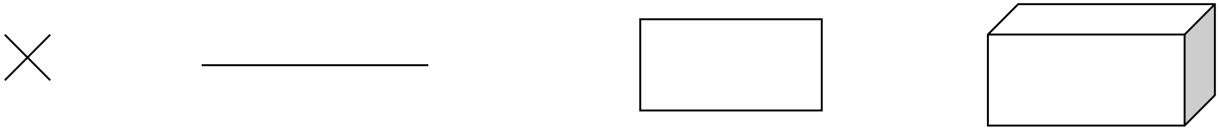
cf.
Schulz & Hahn (KR 2004, ECAI 2004)
Rosse & Smith (MEDINFO 2004)

Semantic framework for biological structure…

- Foundational Relations
- General Attributes
- Theories

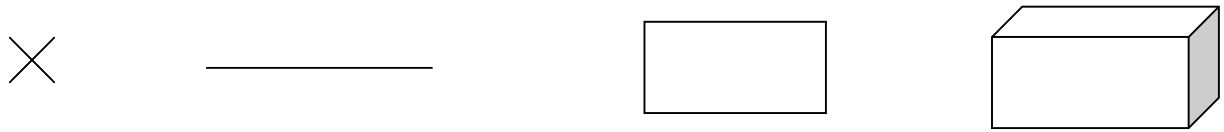
General Attributes (mutually disjoint classes)

- Dimensionality: Point, 1-D, 2-D, 3-D

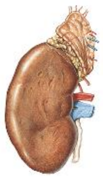


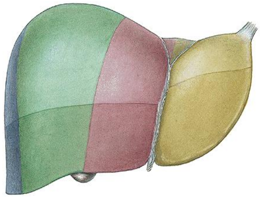
General Attributes (mutually disjoint classes)

- Dimensionality: Point, 1-D, 2-D, 3-D



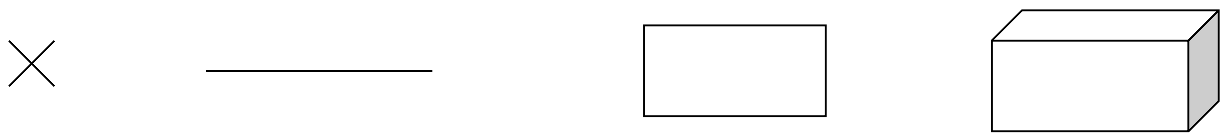
- Solids vs. hollow spaces, vs. Boundaries



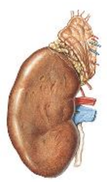
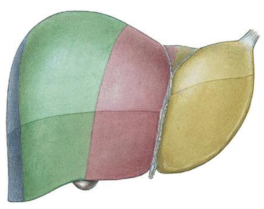


General Attributes (mutually disjoint classes)

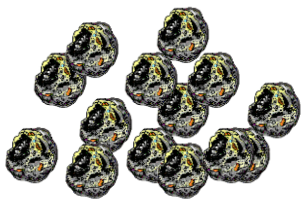
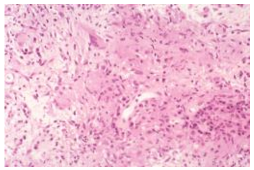
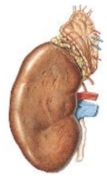
- Dimensionality: Point, 1-D, 2-D, 3-D



- Solids vs. hollow spaces, vs. Boundaries

- Collections vs. Masses vs. Count Objects

cf. Schulz & Hahn, FOIS 01

Semantic framework for biological structure…

- Foundational Relations
- General Attributes
- Theories

Theories

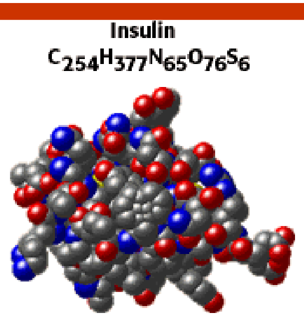
- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
  - **Granularity**
  - **Species**
  - **Development**
  - **Canonicity**

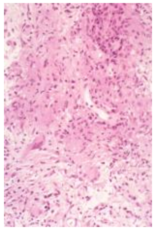
Theories

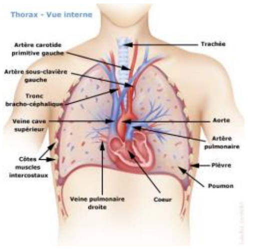
- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
    - **Granularity**
    - **Species**
    - **Development**
    - **Canonicity**

Granularity

- Level of detail (molecular, cellular, tissue, organ)





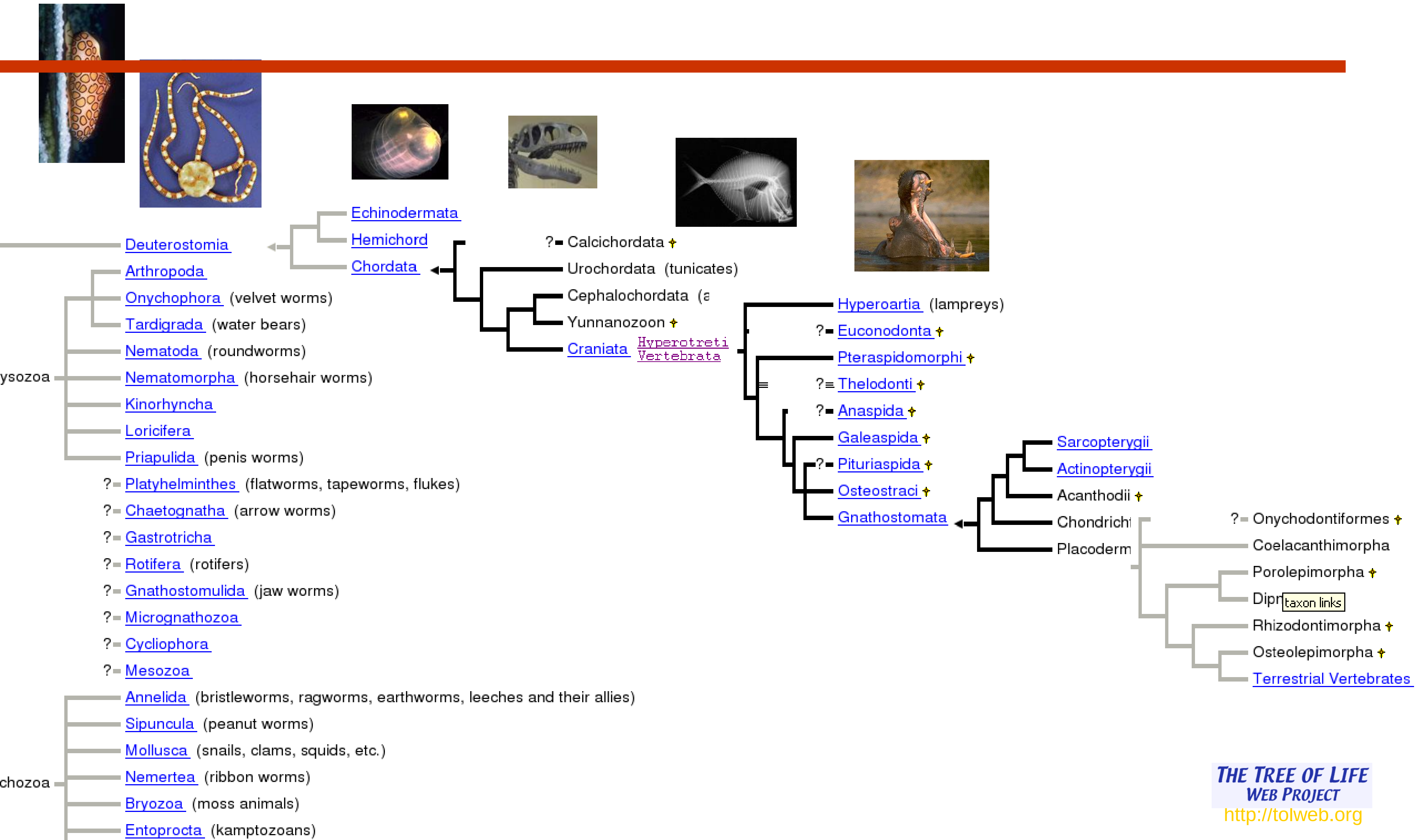


- Change in Granularity level may be non-monotonous
  - Change of sortal restrictions:
    - 3-D → 2-D boundary
    - Count concept → Mass concept
  - Change of relational attributions:
    - disconnected → connected

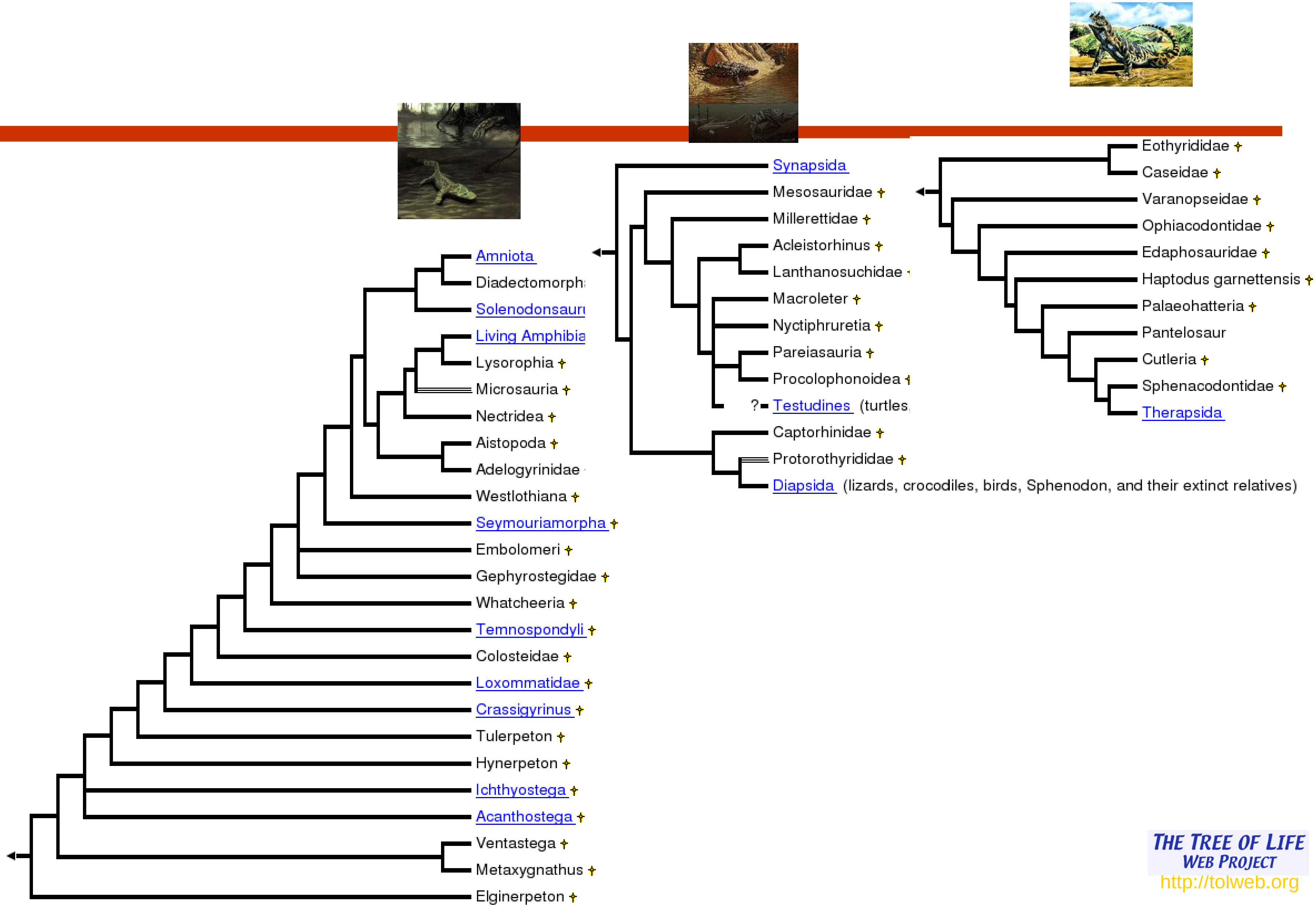
Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
  - **Granularity**
  - **Species**
  - **Development**
  - **Canonicity**

Linnean Taxonomy of Species

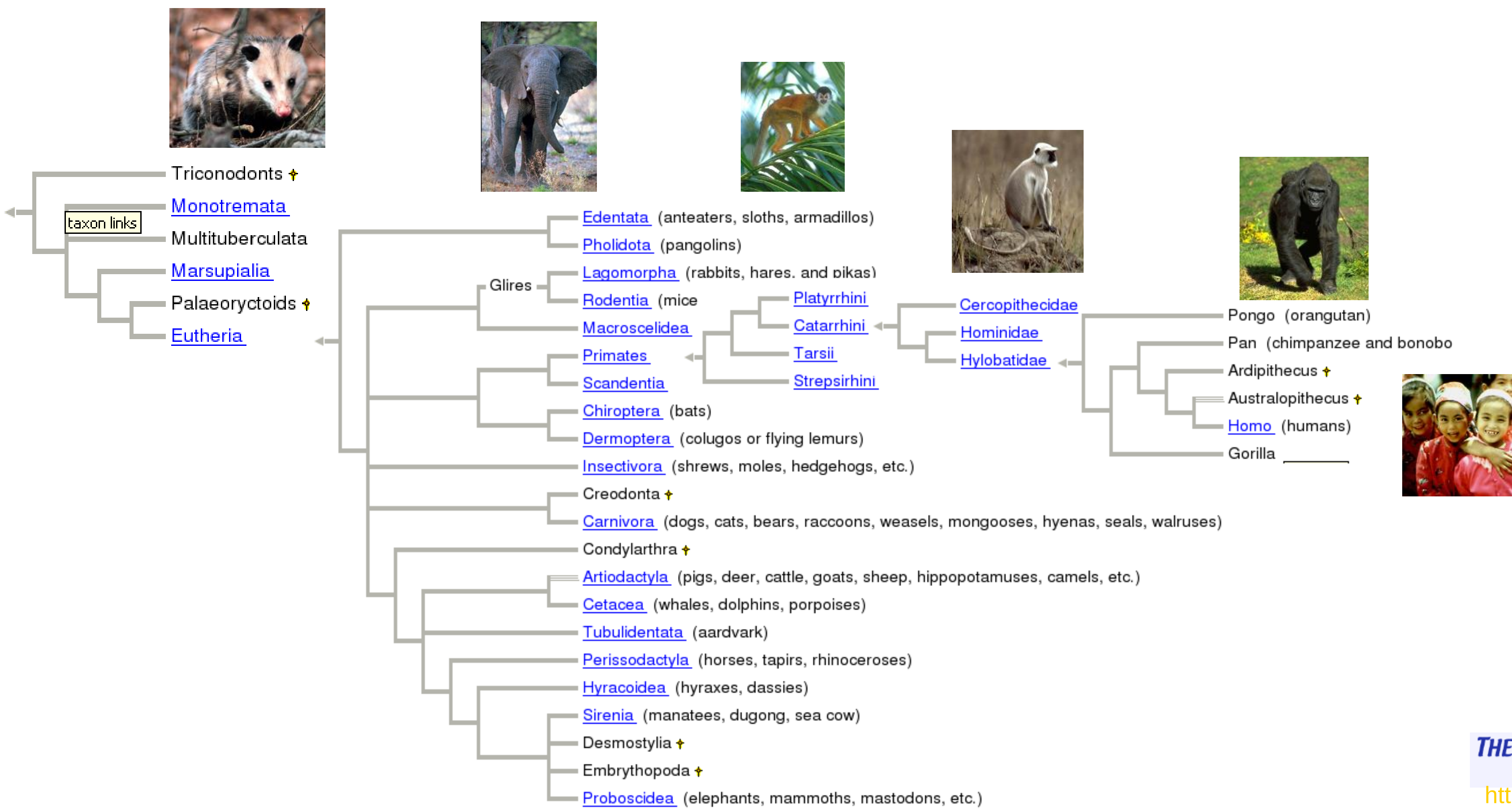


Linnean Taxonomy of Species



THE TREE OF LIFE WEB PROJECT
http://tolweb.org

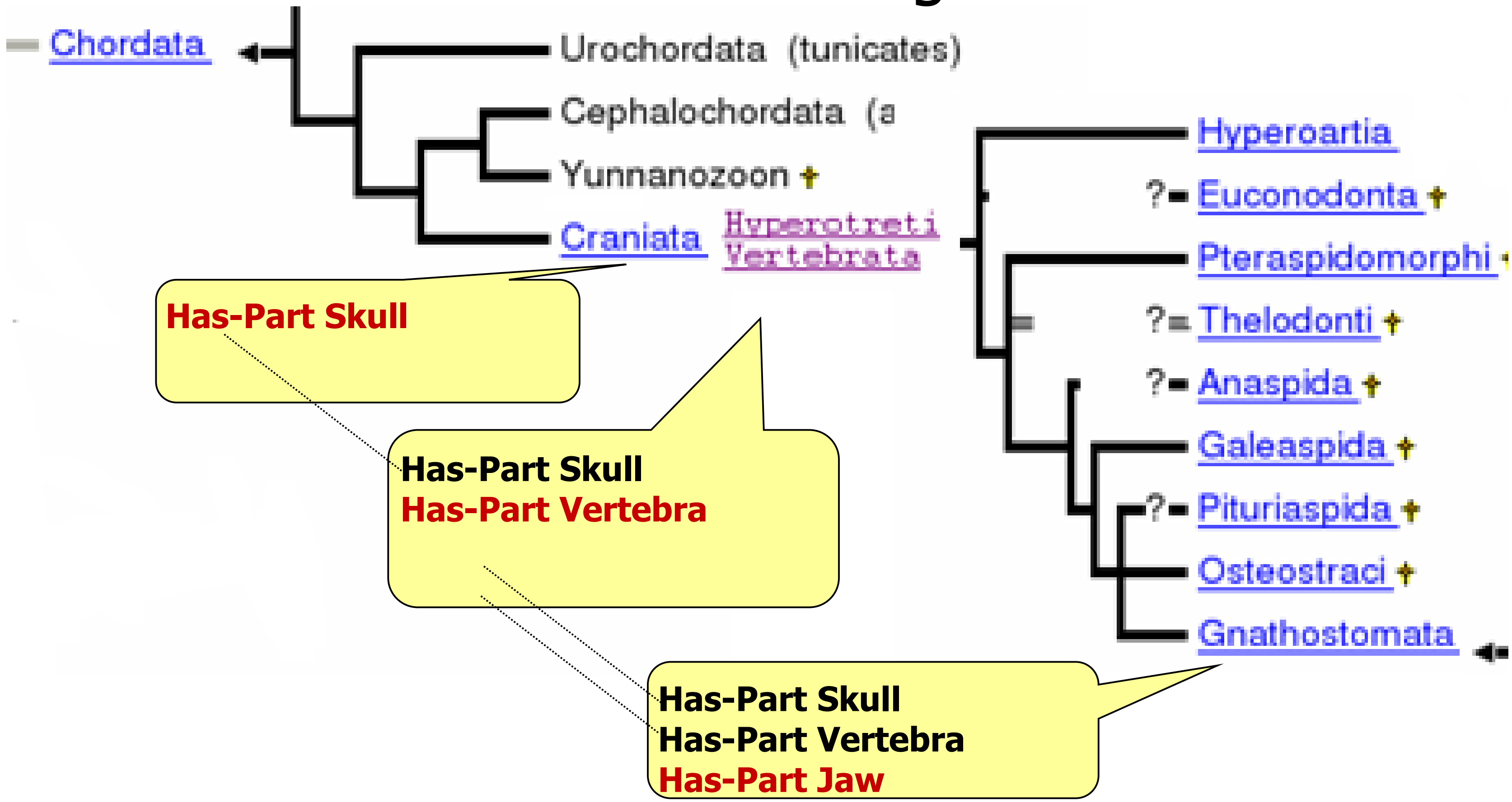
Linnean Taxonomy of Species



THE TREE OF LIFE WEB PROJECT
http://tolweb.org

Species

Introduction of axioms at the highest common level

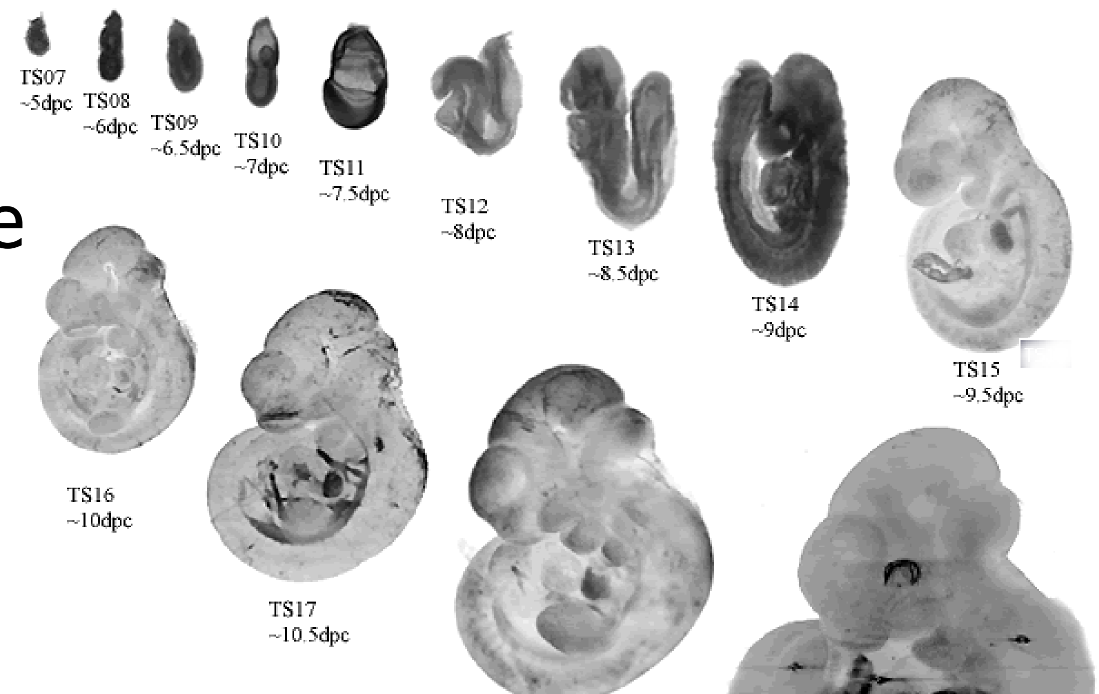


Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
    - **Granularity**
    - **Species**
    - **Development**
    - **Canonicity**

Development

- Represents time-dependent "snapshots" from the life cycle of an organism, e.g., zygote, embryo, fetus, child, adult
- Development stages are species-dependent
  e.g. metamorphosis

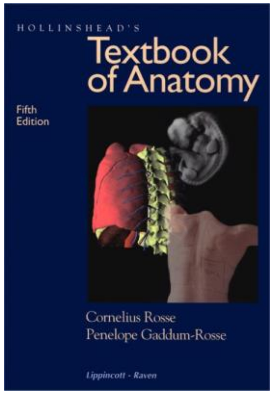


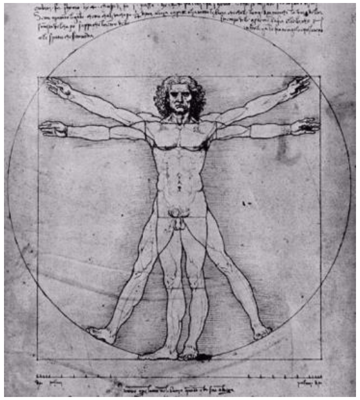
Theories

- A set of formal axioms which describe a restricted (local) domain.
- Four orthogonal theories for Biological Structure
    - **Granularity**
    - **Species**
    - **Development**
    - **Canonicity**

Canonicity

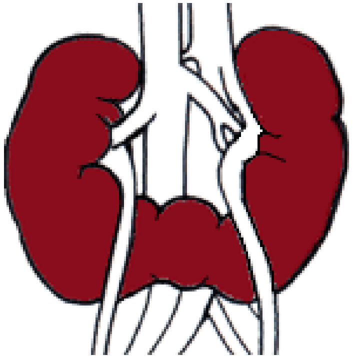
- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure





Canonicity

- Degrees of "Wellformedness" of Biological Structure:
  - Canonic structure
  - Structural Variations



Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure

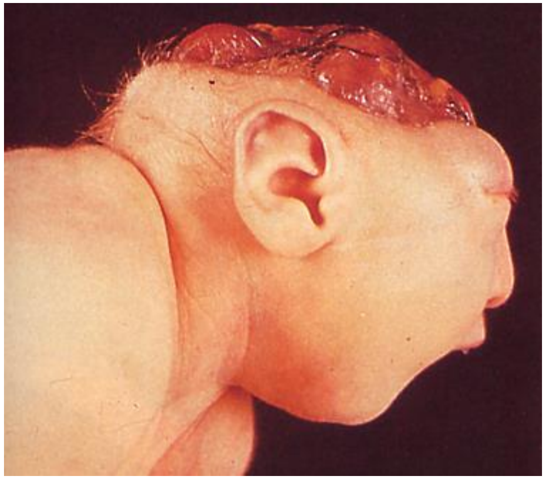
© Layla Aerts - HIB

acquired

congenital

Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure



Canonicity

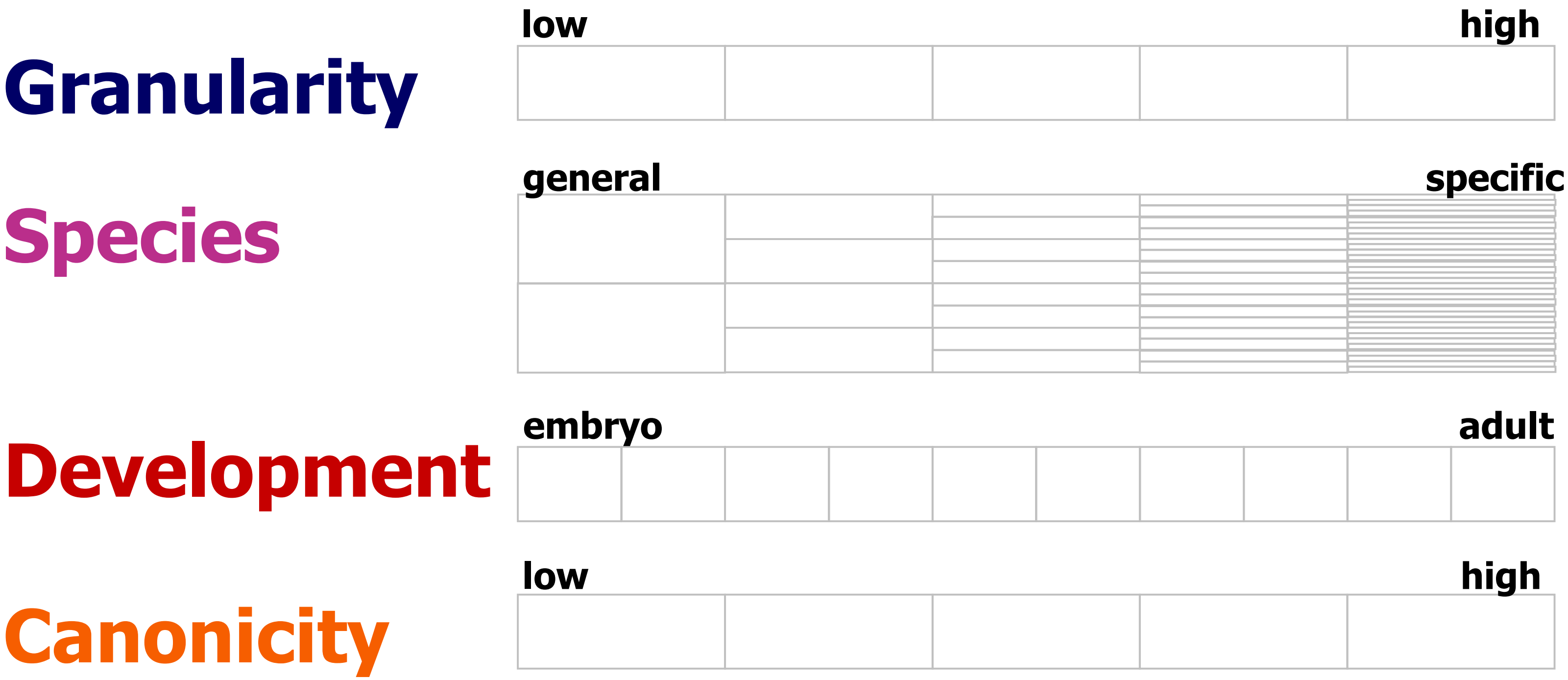
- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure
  - Derivates of biological structure

Canonicity

- Five canonicity levels: each level introduces axioms valid for higher levels

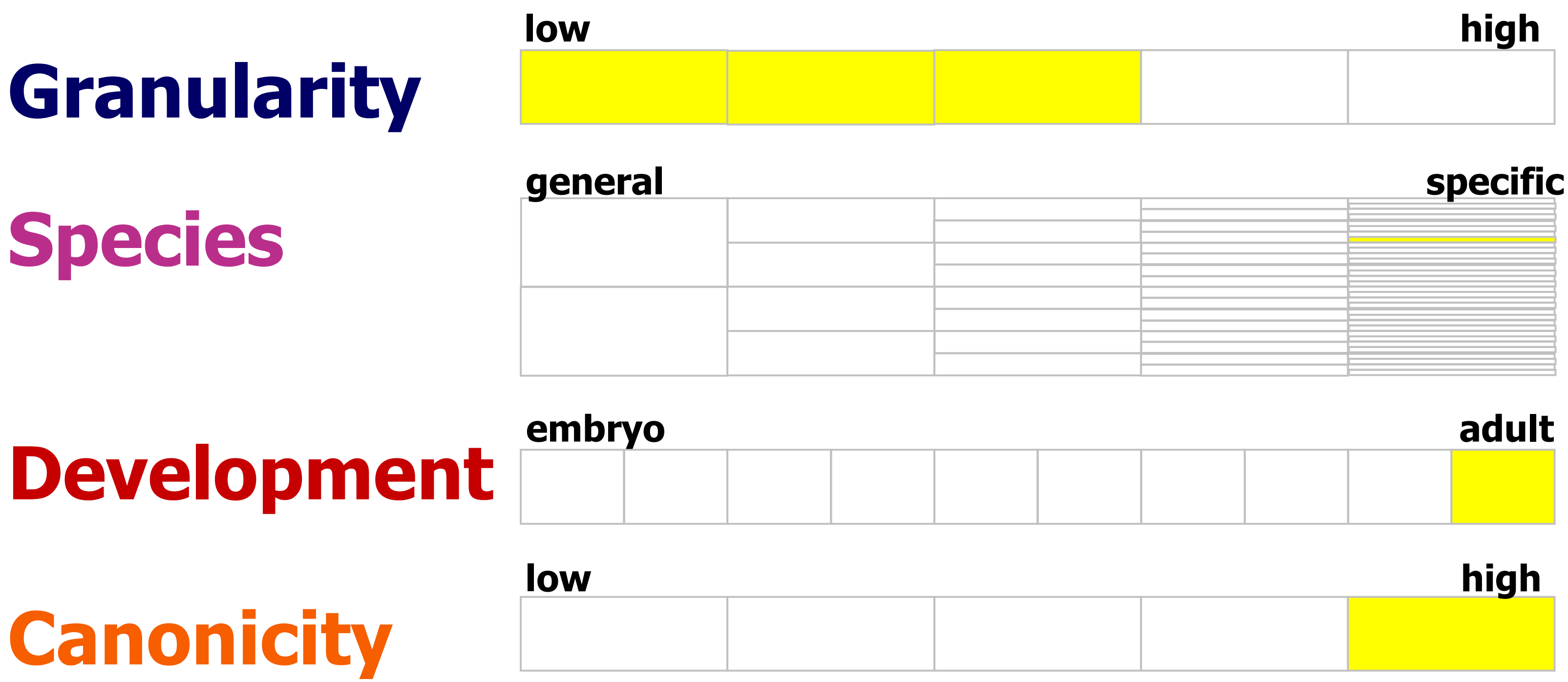
Level	1	2	3	4	5
Theory	any amount of matter, if of biological origin	any living or dead organism	any living organism	living organism without pathologic modifications	ideal organism
Set of Axioms	n_1	n_2 $n_1 \subset n_2$	n_3 $n_2 \subset n_3$	n_4 $n_3 \subset n_4$	n_5 $n_4 \subset n_5$

Examples

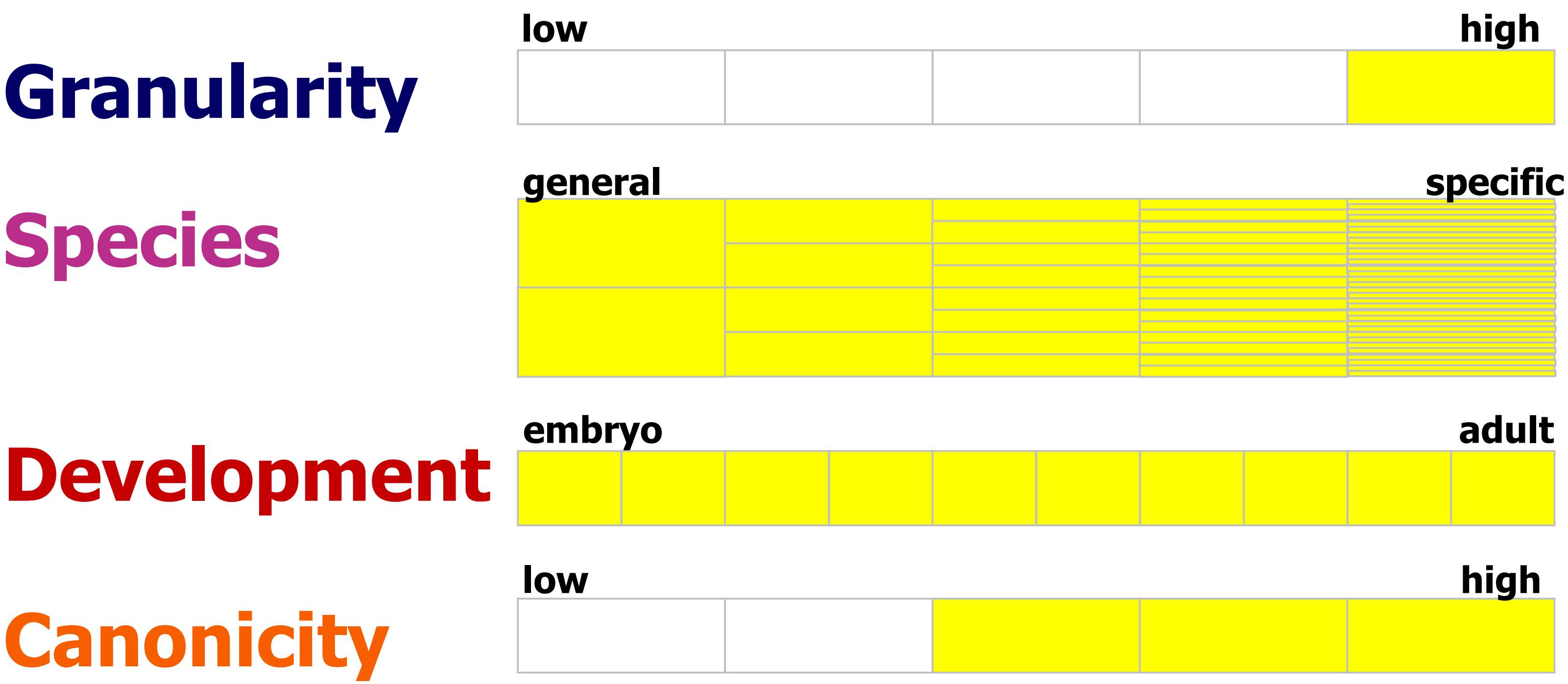


Granularity
low
high
Species
general
specific
Development
embryo
adult
Canonicity
low
high

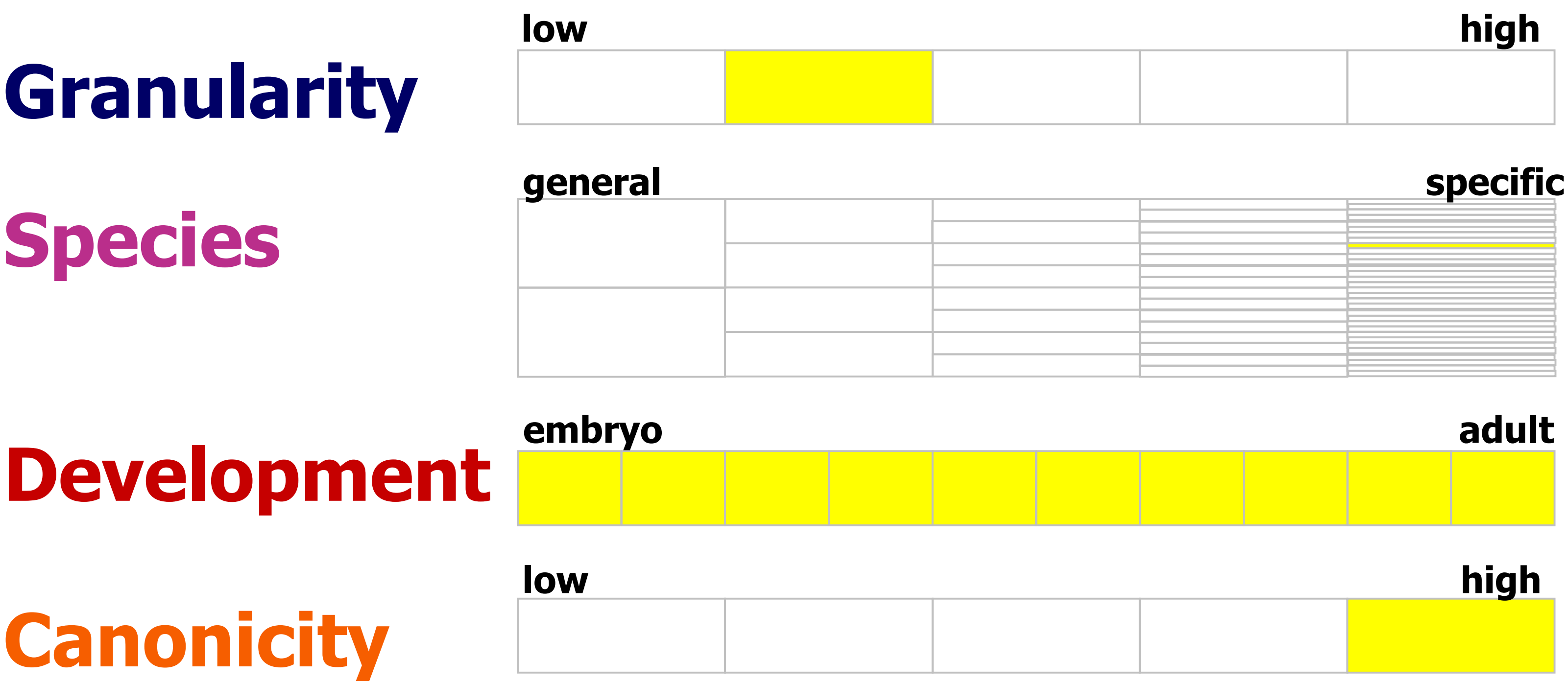
Coverage:
Foundational Model of Anatomy



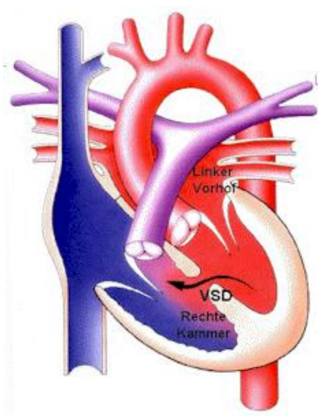
Coverage: Gene Ontology



Coverage: Mouse Anatomy



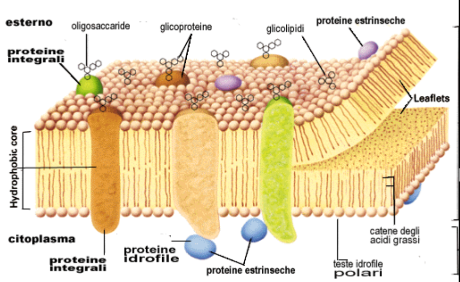
Examples



Connects (RightVentricle, Left Ventricle)

Granularity	**= normal**	**false**
Species	**= mammal**	
Development	**= adult**	
Canonicity	**= 4-5**	

Granularity	**= any**	**true**
Species	**= vertebrate**	
Development	**= early embryo**	
Canonicity	**= any**	



Is-A (Membrane, 3-D object)

Granularity	**= normal**	**true**
Species	**= any**	
Development	**= any**	
Canonicity	**= any**	

Granularity	**= lowest**	**false**
Species	**= any**	
Development	**= any**	
Canonicity	**= any**	

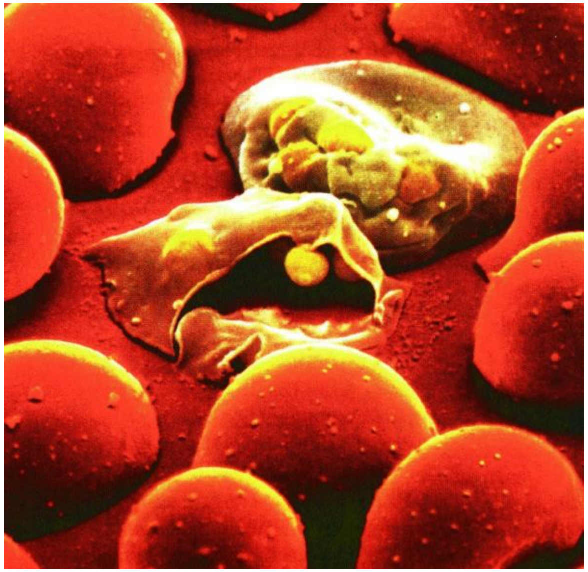
Conclusion

- Integration of bio-ontologies requires
  - Uncontroversial semantics of relations and attributes
  - Clear commitment to theories, such as granularity, species, development and canonicity
- Redundancy can be avoided
  - Encoding axioms at the highest common level in the species taxonomy (e.g. vertebrates, arthropods, primates) and benefit from inheritance in subsumption hierarchies

…requires sophisticated organization

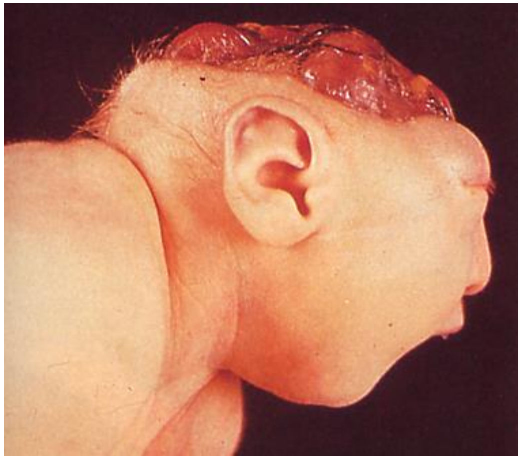
- Formalization and Standardization of Clinical Terminologies
- Basis for the Annotation of Genes and Gene Products
- Semantic reference for scientific communication
- Machine-supported reasoning and decision-support

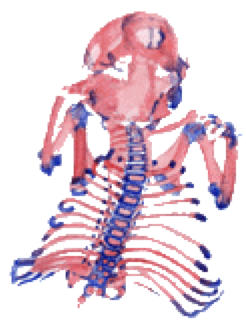
Bio-ontologies !

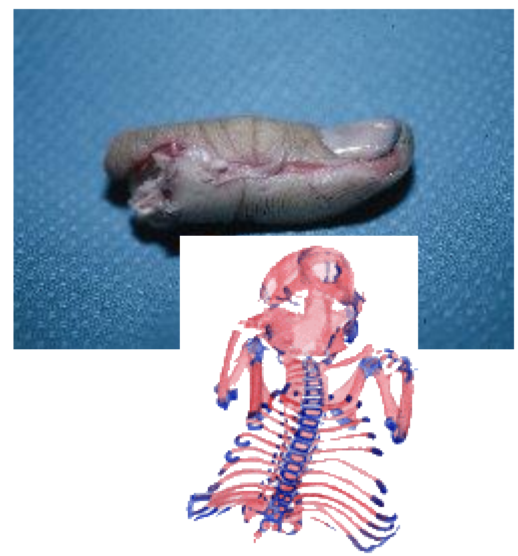


Upper level classification of entities

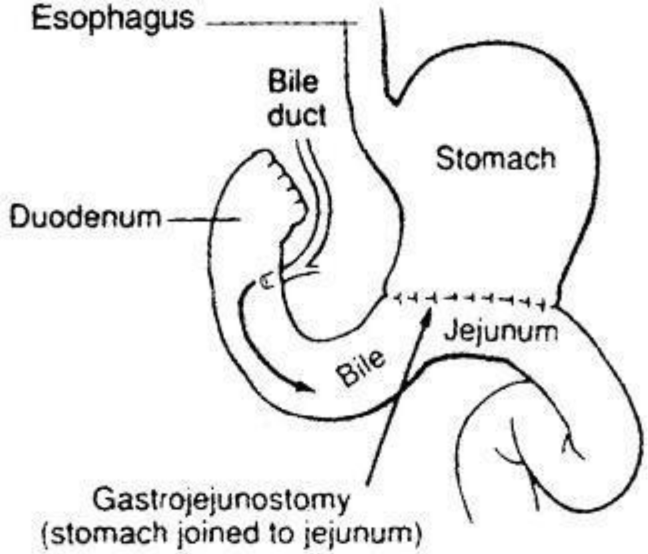
	Individuals (concrete objects)	Universals (Concepts, Classes of Individuals)
Continuants (physical objects,…)	• my left hand • a blood sample • a concrete cell	• Hand, • Blood • Cell
Occurrents (events, processes, actions…)	• Peter's diabetes • appendectomy of Patient #12345	• Diabetes mellitus • Appendectomy

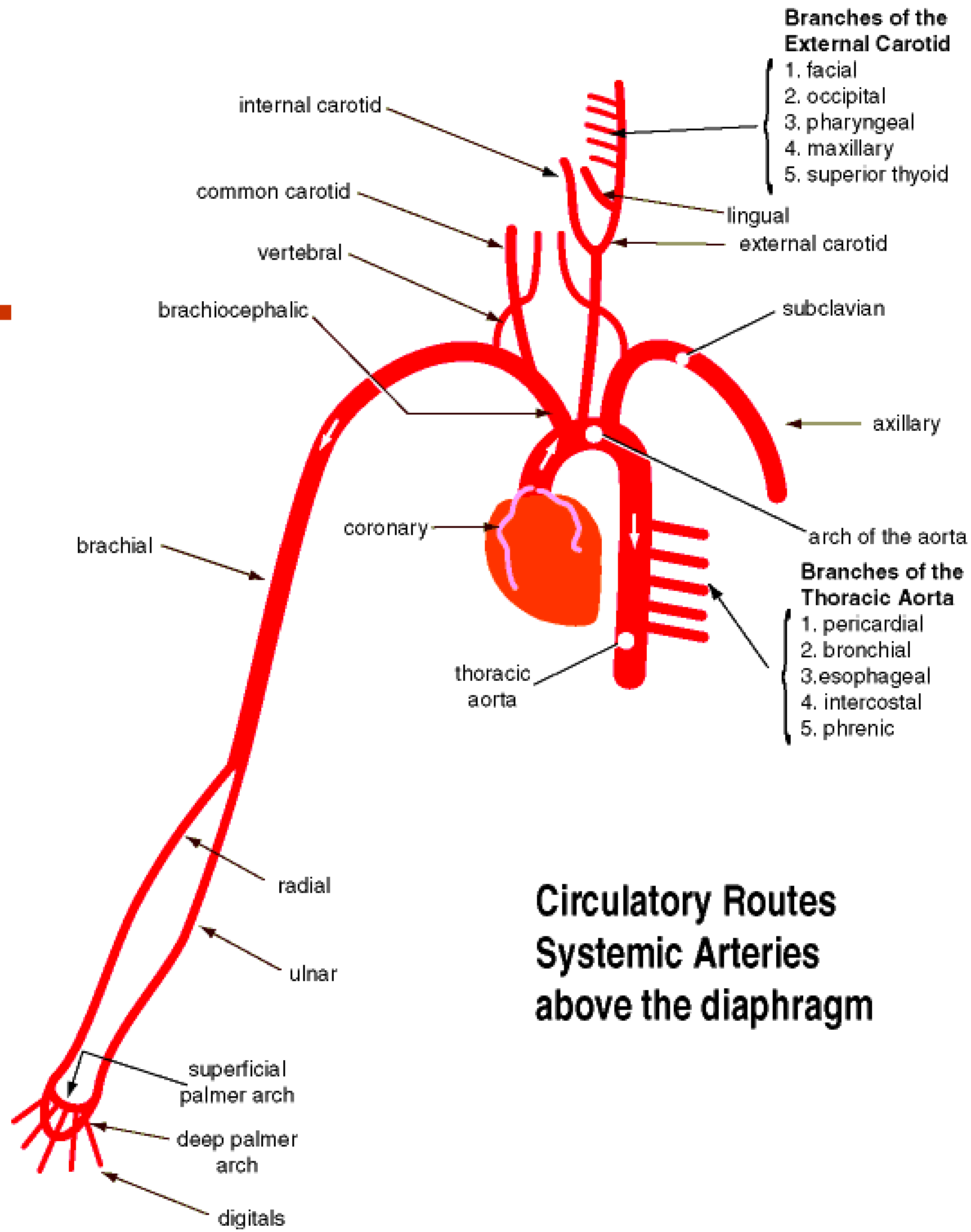







w^{1118}
© biol.org

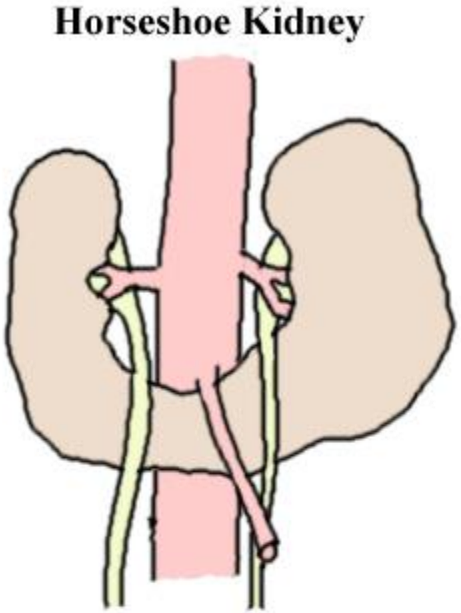

Esophagus
Bile duct
Stomach
Duodenum
Jejunum
Bile
Gastrojejunostomy
(stomach joined to jejunum)

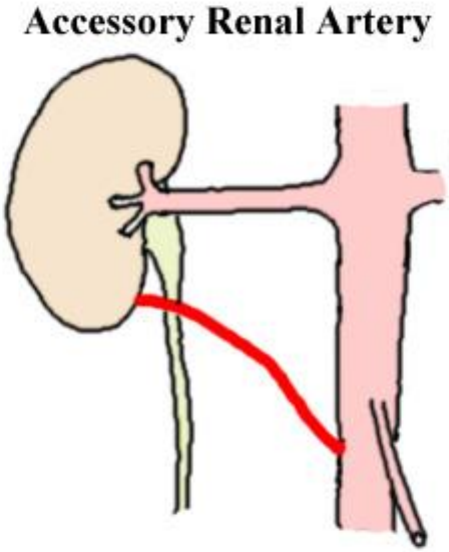


Circulatory Routes
Systemic Arteries
above the diaphragm


b
c

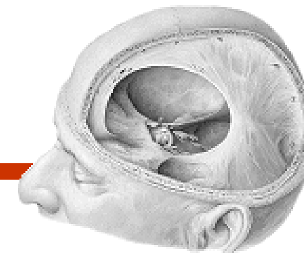

b

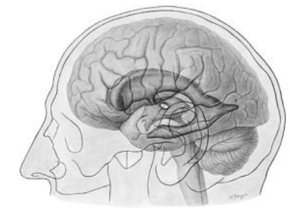

Horseshoe Kidney

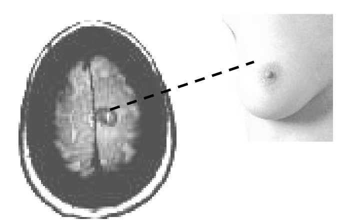

Accessory Renal Artery

Mereotopological Quiz

- Cranial Cavity *has-location* Head
  Is Cranial Cavity part-of Head ? ☺☹
- Brain *has-location* Cranial Cavity
  Is Brain part of Cranial Cavity ? ☺☹
- Glioblastoma *has-location* Brain
  Glioblastoma part-of Brain? ☺☹
- Brain metastasis *has-location* Brain
  Brain metastasis part-of Brain? ☺☹
- Embryo *has-location* Uterus
  Embryo part-of Uterus ? ☺☹







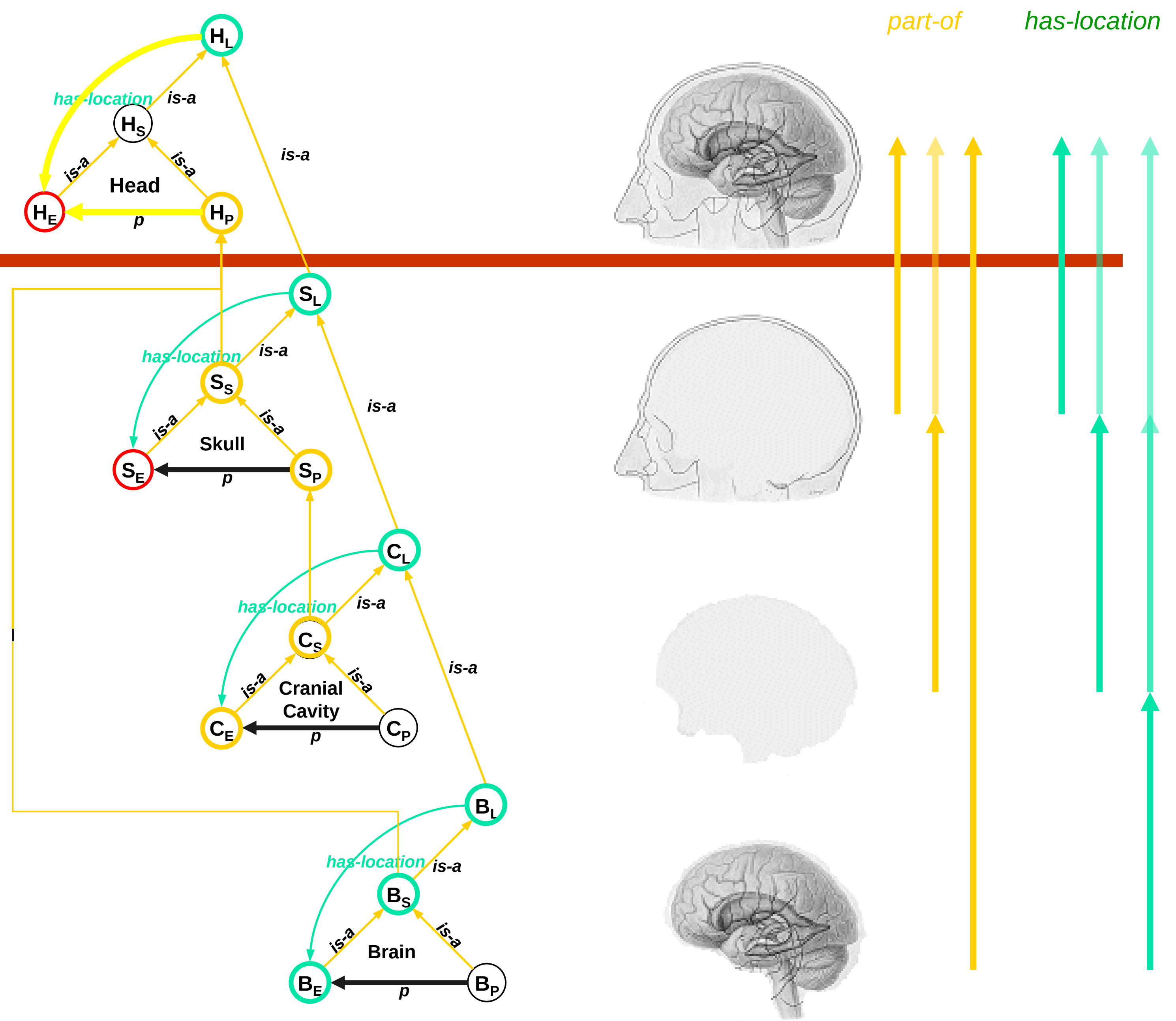
Images from: Sobotta CD-ROM

Head
(Solid)

Skull
(Solid)

Cranial
Cavity
(Hole)

Brain
(Solid)



transitive closure by taxonomic subsumption

Subtheories of an Ontology of Biological Structure

1. Taxonomy 2. 3. Topology

„is-a“ „part-of“ „connection“

Mereology

part-of

Hand

part-of

Thumb

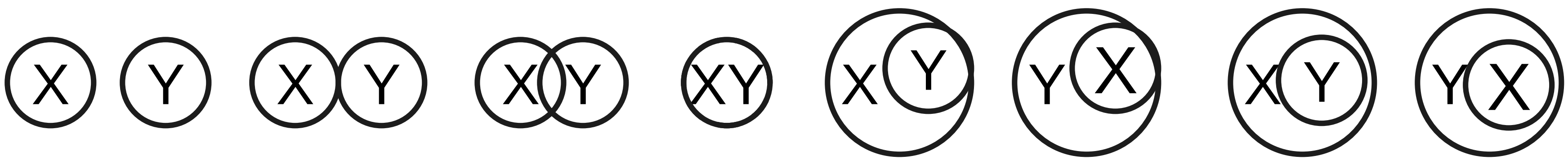
part-of

Thumbnail

Subtheories of an Ontology of Biological Structure

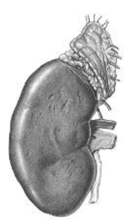
„is-a“ „part-of“ „connection“

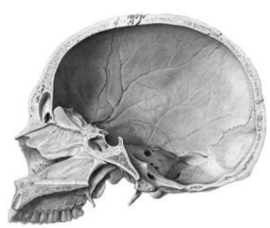
- Canonical relationships



(Schulz et al. AMIA 2000)

- Topological Primitives:



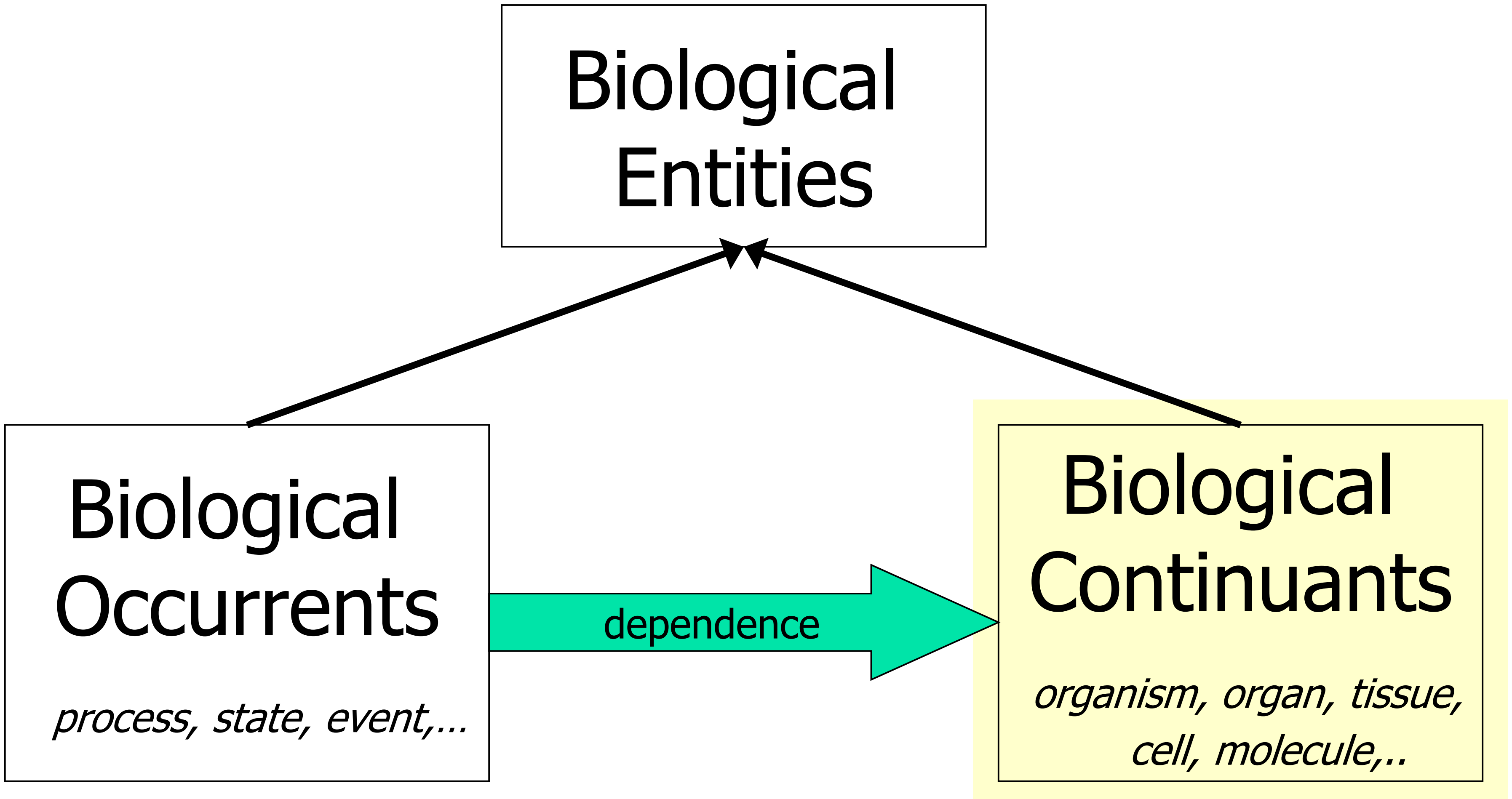


Structure of Talk

- Introduction
- Foundational Relations
- Foundational Attributes
- Theories
  - Granularity
  - Species
  - Development
  - “Canonicity”

The World of Life Sciences…

Generalized Representation of Living Systems: Top Level

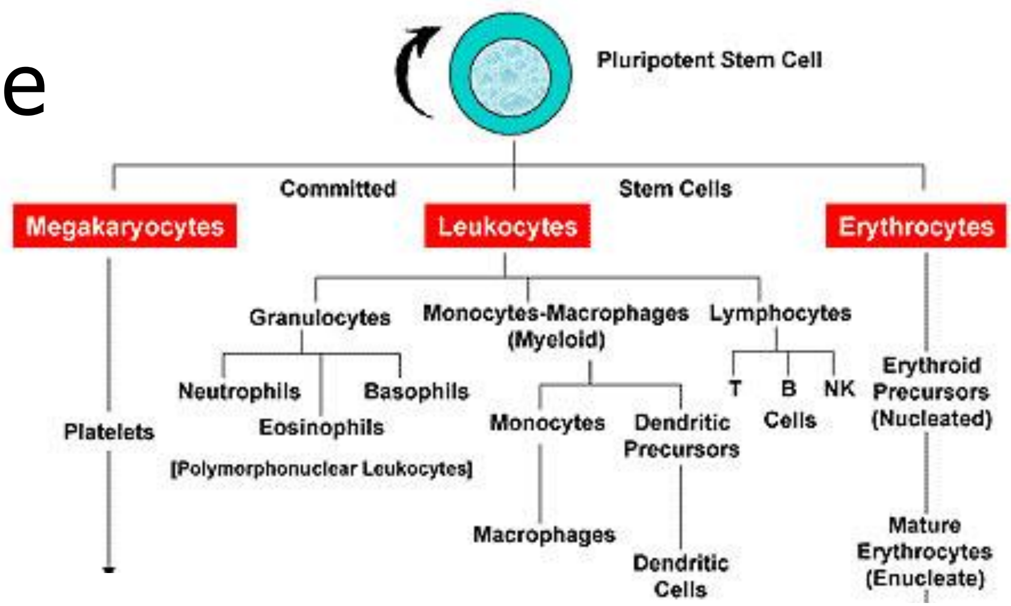


Ontological Account for Biological Continuants

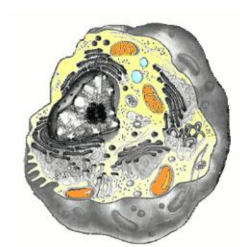
- Foundational Relations
- Foundational Attributes
- Theories
  - Granularity
  - Species
  - Development
  - “Canonicity”

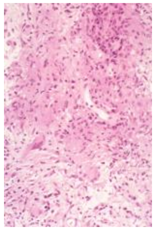
Granularity

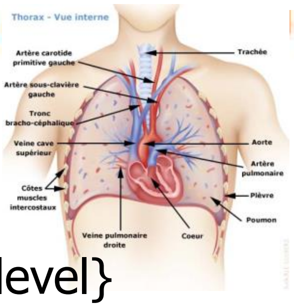
- Taxonomic: degree of specialization



- Mereologic: degree of dissection







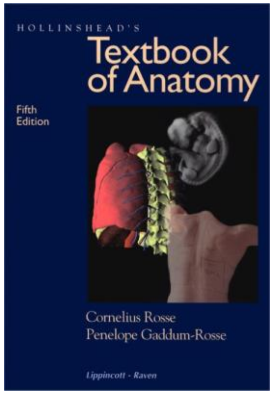
{molecular level, cellular level, tissue level, organ level, population level}

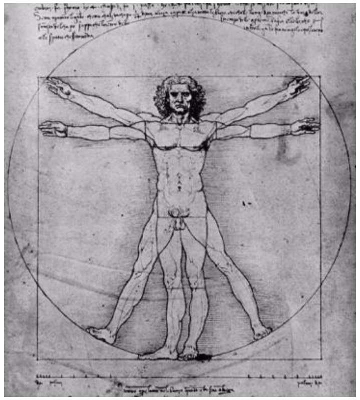
Change in Granularity level may be non-monotonous

- Change of sortal restrictions:
  - 3-D $\rightarrow$ 2-D boundary
  - Count concept $\rightarrow$ Mass concept
- Change of relational attributions:
  - disconnected $\rightarrow$ connected
-

Canonicity

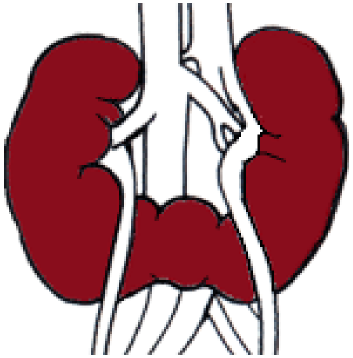
- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure





Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations



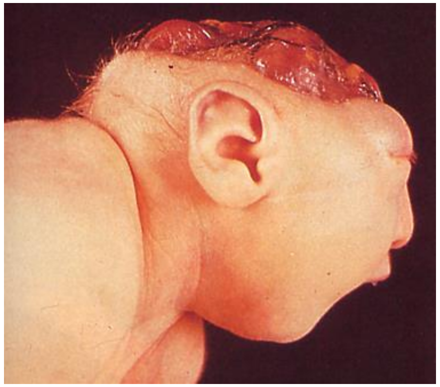
Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure

© Layla Aerts - HIB

Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure



Canonicity

- Degrees of “Wellformedness” of Biological Structure:
  - Canonic structure
  - Structural Variations
  - Pathological Structure
  - Lethal Structure
  - Derivates of biological structure