Bio-ontologies

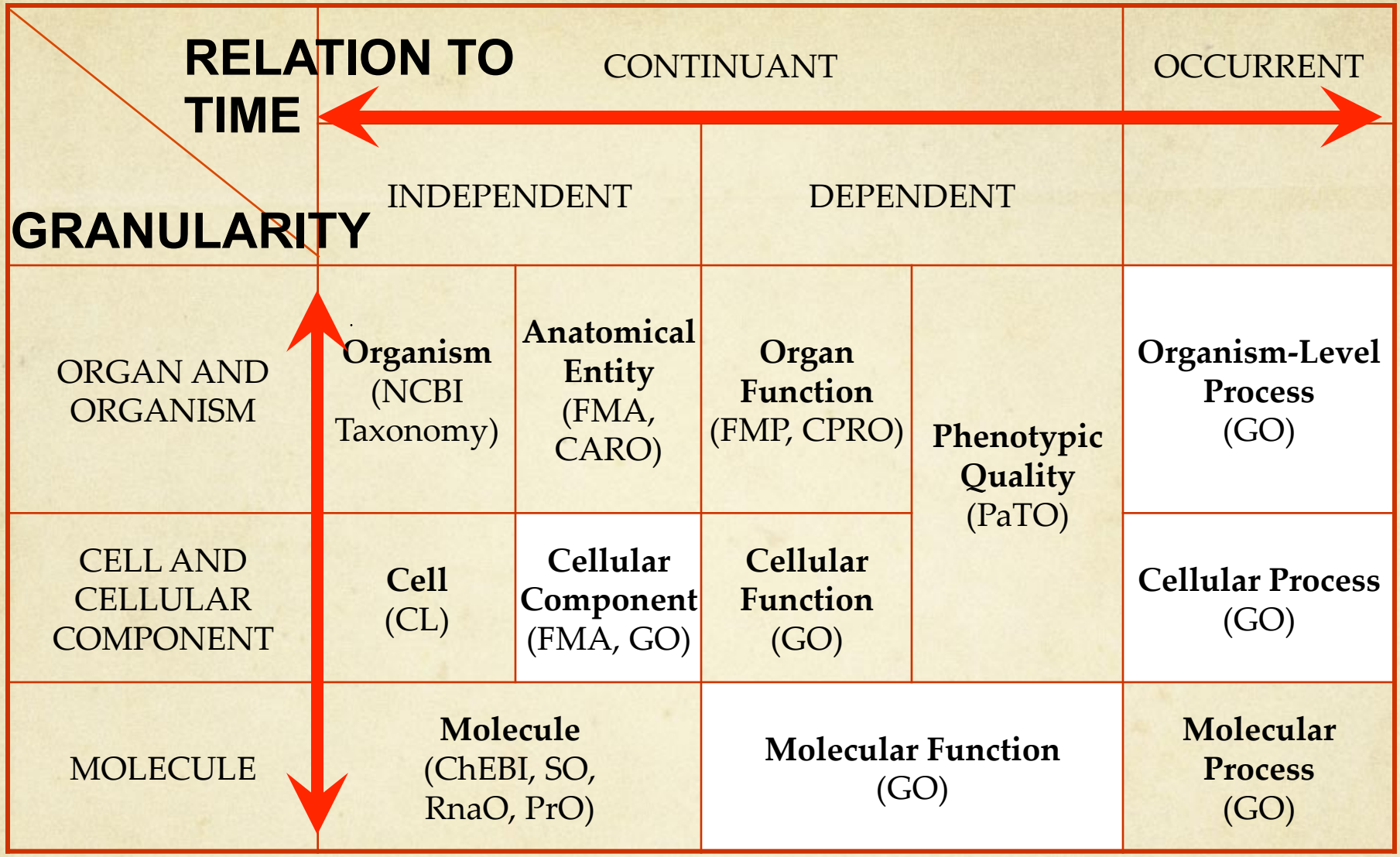
Use of Ontologies

The use of ontologies began in the biological sciences around 1998 with the development of the Gene Ontology (GO) [1].

By 2007, there was sufficient interest and activity in the area to merit national and international coordination efforts such as the Open Biomedical Ontologies (OBO) Foundry [2] or the National Center for Biomedical Ontologies [3].

OBO foundry

- Developers commit to working to ensure that, for each domain, there is community convergence on a single ontology
- and agree in advance to collaborate with developers of ontologies in adjacent domains.
- http://obofoundry.org



OBO Foundry coverage

NCBO - National Center for Biotechnology Information

Part of the National Library of Medicine at NIH

- Establish public databases
- Research in computational biology
- Develop software tools for sequence analysis
- Disseminate biomedical information

Ontology

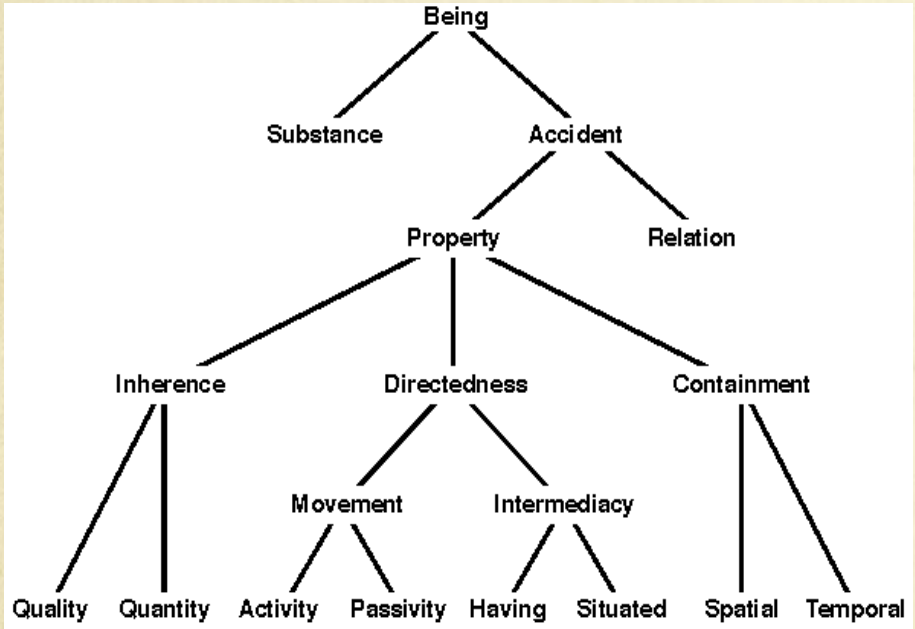
- First Philosophy (Aristotle Metaphysics Gamma)

õv ην õv

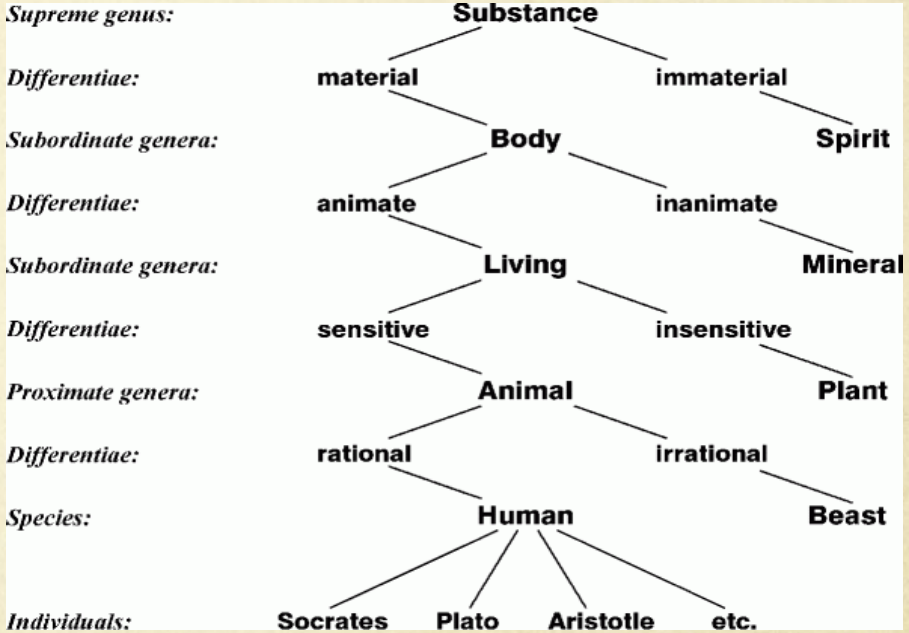
- The first philosophy seeks to provide a definitive and exhaustive classification of entities that exist
- The study of what exists - metaphysics

Ontology

Aims to find out what entities and types of entities exist



shows a version of the Tree of Porphyry drawn by the 13th century logician Peter of Spain.



Oldest known tree diagram was drawn in the 3rd century AD by the Greek philosopher Porphyry in his commentary on Aristotle's categories

Four established meanings of ontology

- Pure philosophical ontology

  which deals with questions such as what is meant by the concept of being, why does something exist rather than nothing

- Applied scientific ontology:
  - Discipline

    a method or activity of enquiry into philosophical problems about the concept or facts of existence
  - Domain

    the outcome or subject matter of ontology as a discipline.

Applied scientific ontology construed as an existence domain can be further subdivided as:

- the theoretical commitment to a preferred choice of existent entities
  - a description or inventory of the things that are supposed to exist according to a particular theory,
- to the real existent entities themselves, including the actual world considered as a whole, the extant domain
  - the actual world of all real existent entities, whatever these turn out to be, identified by a true complete applied ontological theory.

Ontology definition

Many definitions of 'ontology' have been proposed in the literature [4–10], and classifications of different types of vocabularies, thesauri, ontologies and knowledge bases have been proposed, based on criteria such as:

- their intended use,
- degree of formalization
- or philosophical interpretation [2, 11–15].

Ontology in Philosophy

- As a term is was created independently in 1613 from two philosophers, Rudolf Göckel and Jacob Lorhard
- Ontology as a branch of philosophy is the science of **what is**, of the kinds and structures of objects, properties, events, processes and relations.
- Ontology is the theory of objects and their ties.
- It provides criteria for distinguishing
  - various types of objects (concrete and abstract, existent and non-existent, real and ideal, independent and dependent) and
  - their ties (relations, dependences and predication).

Ontology in Information Science

Evolved from the interoperability requirement for system integration

- data heterogeneity
- semantic heterogeneity

An ontology is a **data model** that represents a domain and is used to reason about the **objects in that domain** and the relations between them

Ontology in Artificial Intelligence

- *"an explicit specification of a shared conceptualization"*

  Gruber, 1993

- an ontology is the description of the concepts and relationships that can exist within a given domain

The common ground...

Ontology = A specification of entities (or concepts), relations, instances and axioms in an area of study.

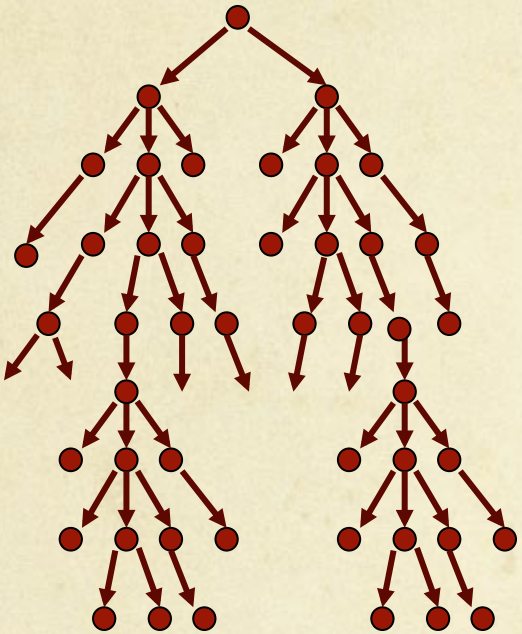
Irrespectively of these different but related definitions, ontologies play an important role with pragmatic implications, **forming specifications for sharing concepts**.

Building principles derived from each domain can prove very helpful in recognizing and avoiding potential logical ambiguities, ensuring that computational ontologies adhere to formalizations that help in better knowledge representation and organization

What do ontologies consist of ?

- Classes
  - Concepts:
    - Processes and Entities
    - Protein, Gene, DNA, glycolysis, neoplasia…
  - Terms: Labels for concepts
    - “Hexokinase”, “Shh”, “Protein”,
    - Identifiers
  - Uniquely identify a class / term
- Relationships (axioms) : Semantic links between concepts
  - is-a, part-of, has-quality,…

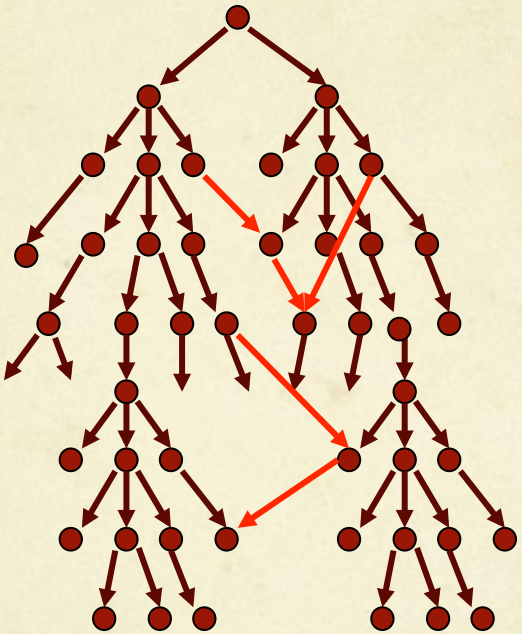
Hierarchies, graphs and DAGS



Rule e.g. “has part”

Directed rule: 1 parent

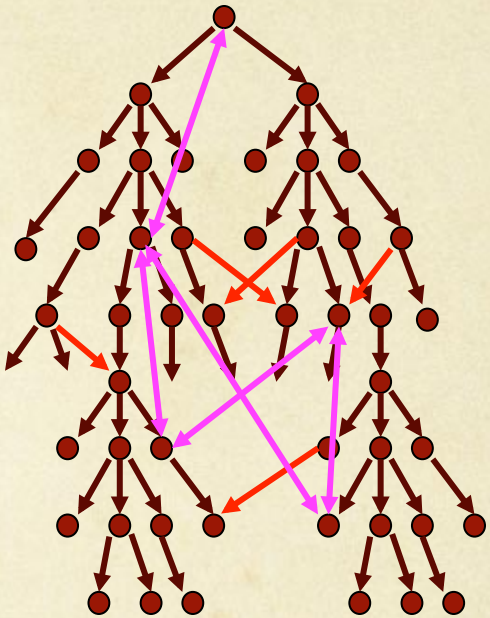
Simple hierarchy



Rule e.g.. “signals to”

Directed rule: >1 parent

Directed acyclic graph



Rule e.g. “is next to”

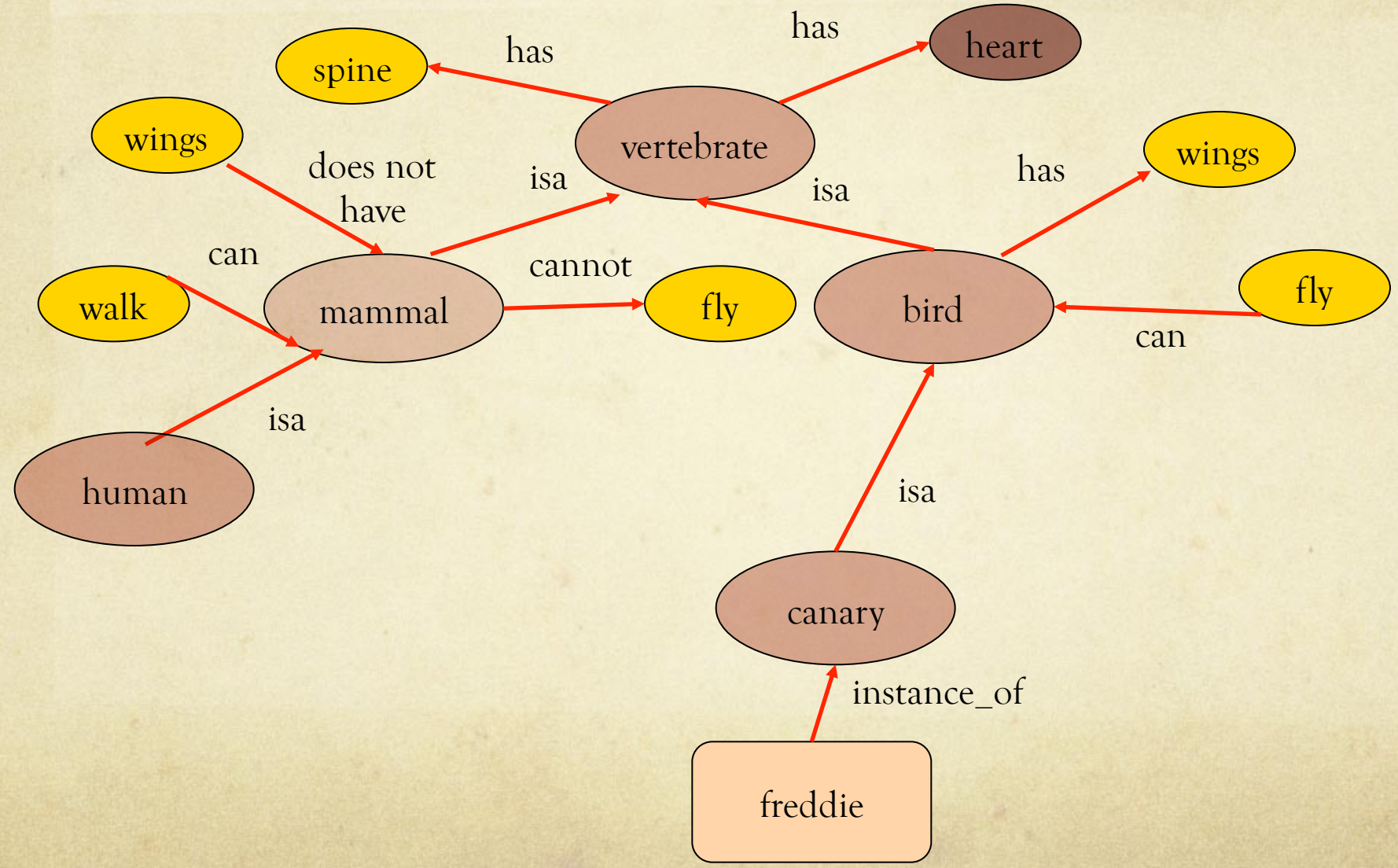
Undirected rule and parents ≡ children

Graph

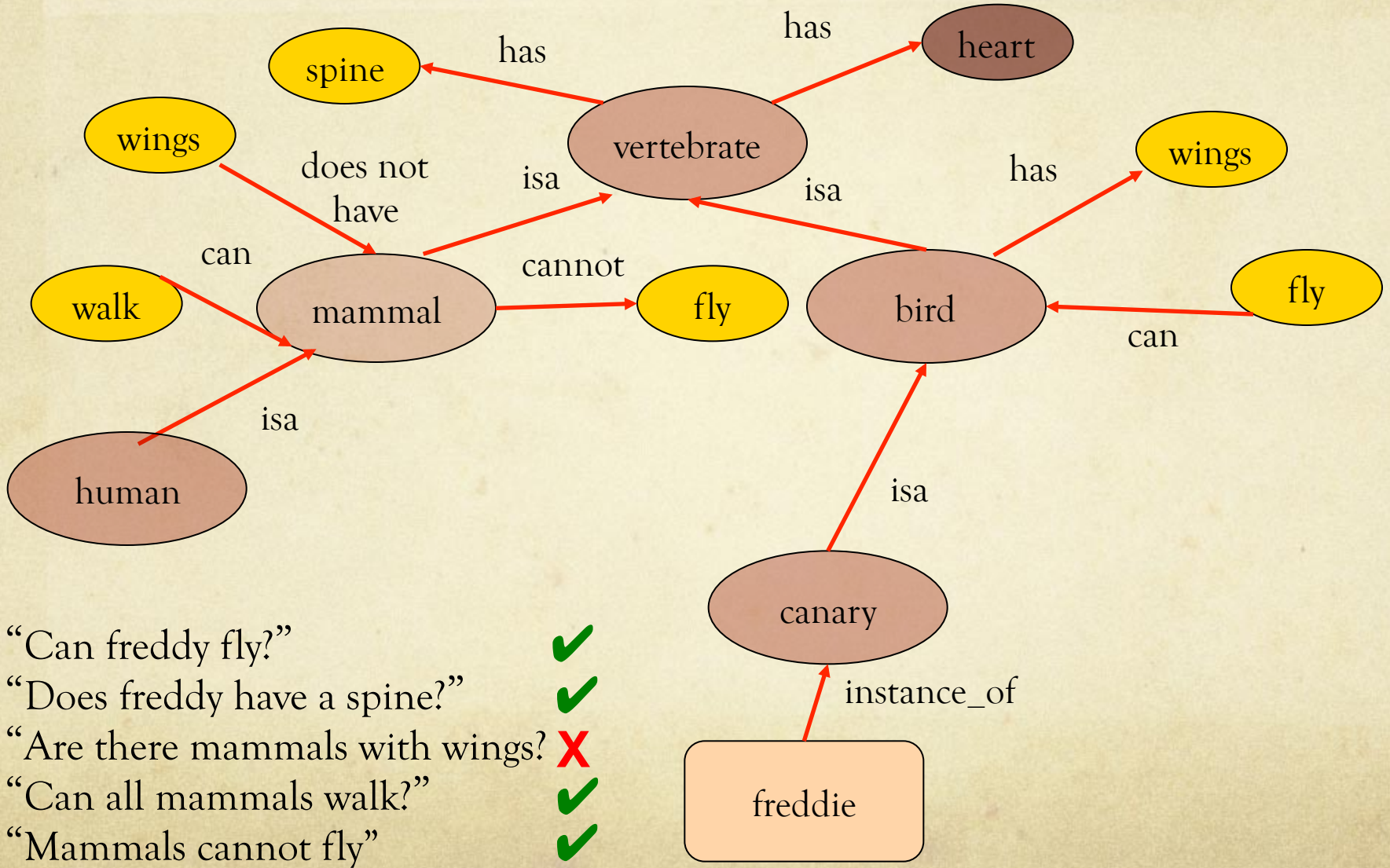
Freddy the canary

- “Can freddy fly?”
- “Does freddy have a spine?”
- “Are there mammals with wings?”
- “Can all mammals walk?”
- “Mammals cannot fly”

Semantic Networks

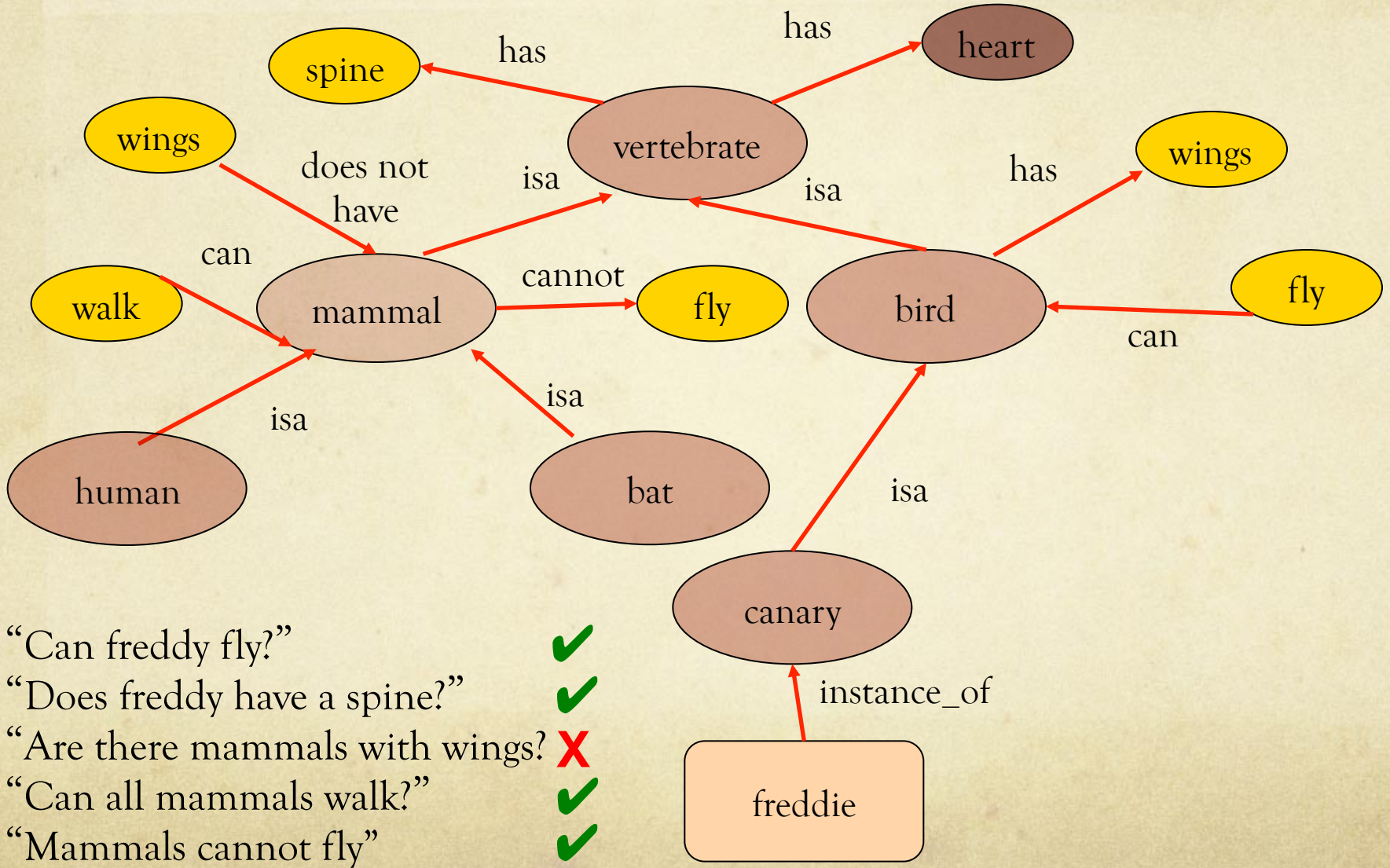


Semantic Networks



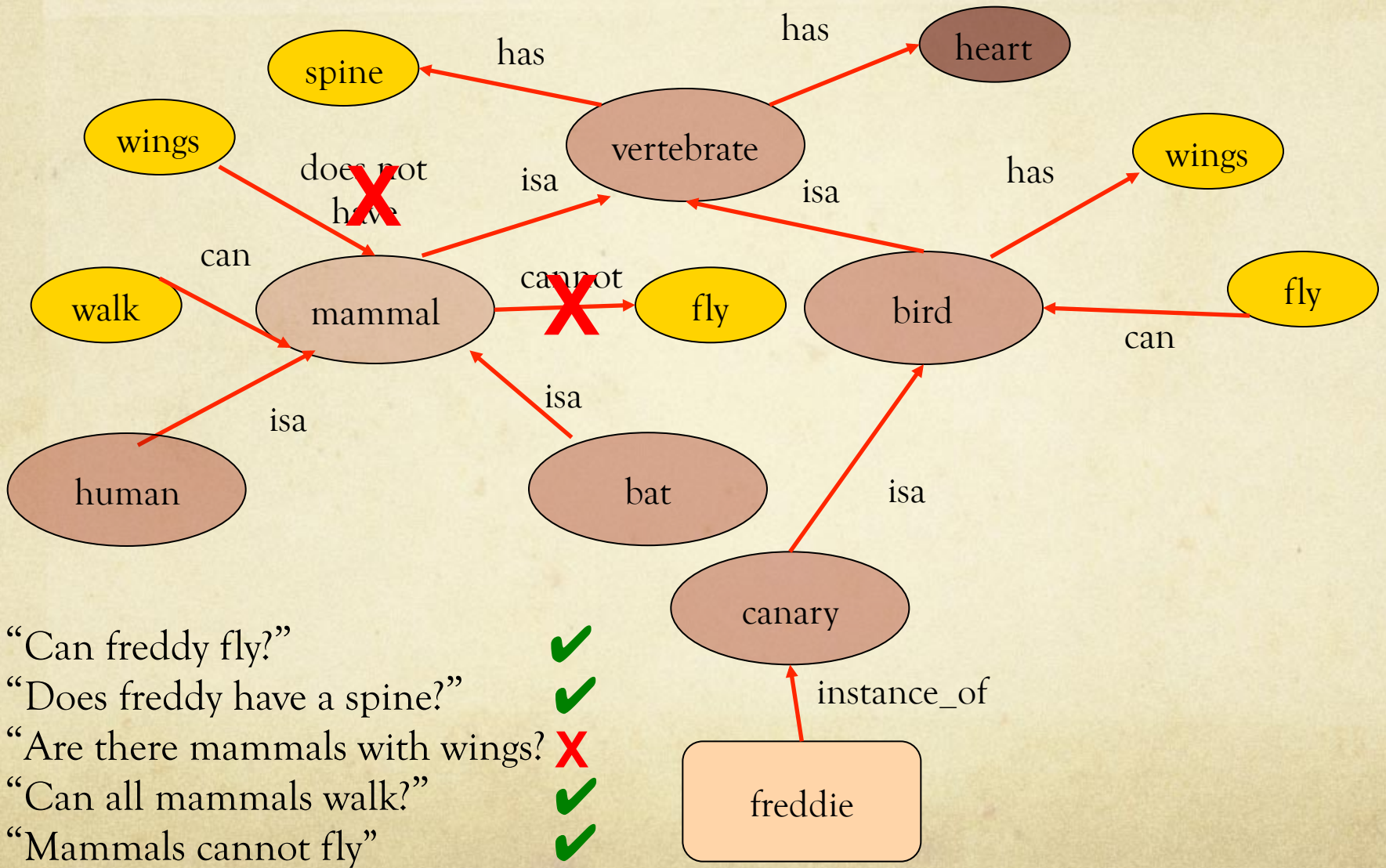
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? X
"Can all mammals walk?" ✔
"Mammals cannot fly" ✔

Semantic Networks



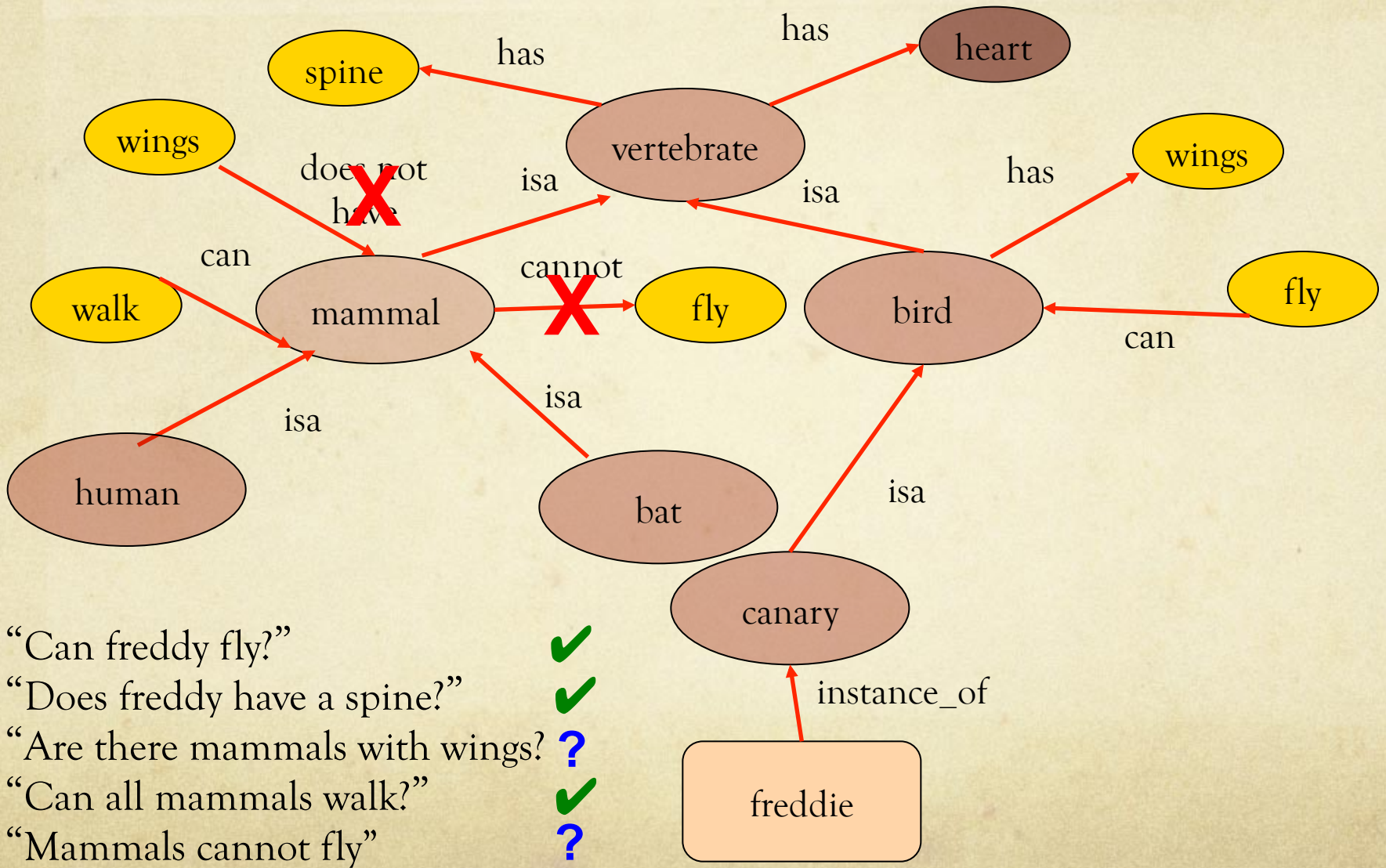
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? X
"Can all mammals walk?" ✔
"Mammals cannot fly" ✔

Semantic Networks



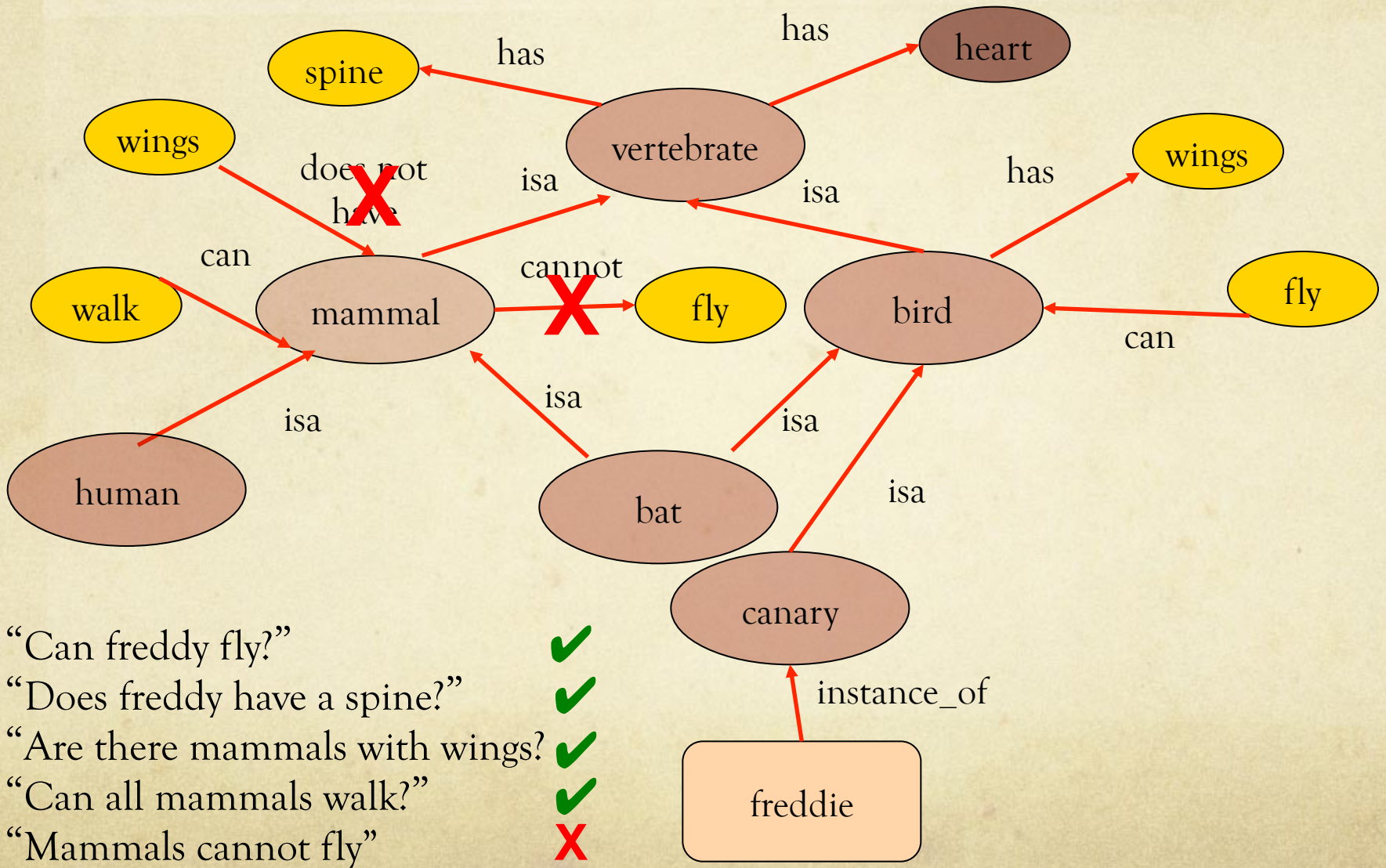
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? X
"Can all mammals walk?" ✔
"Mammals cannot fly" ✔

Semantic Networks



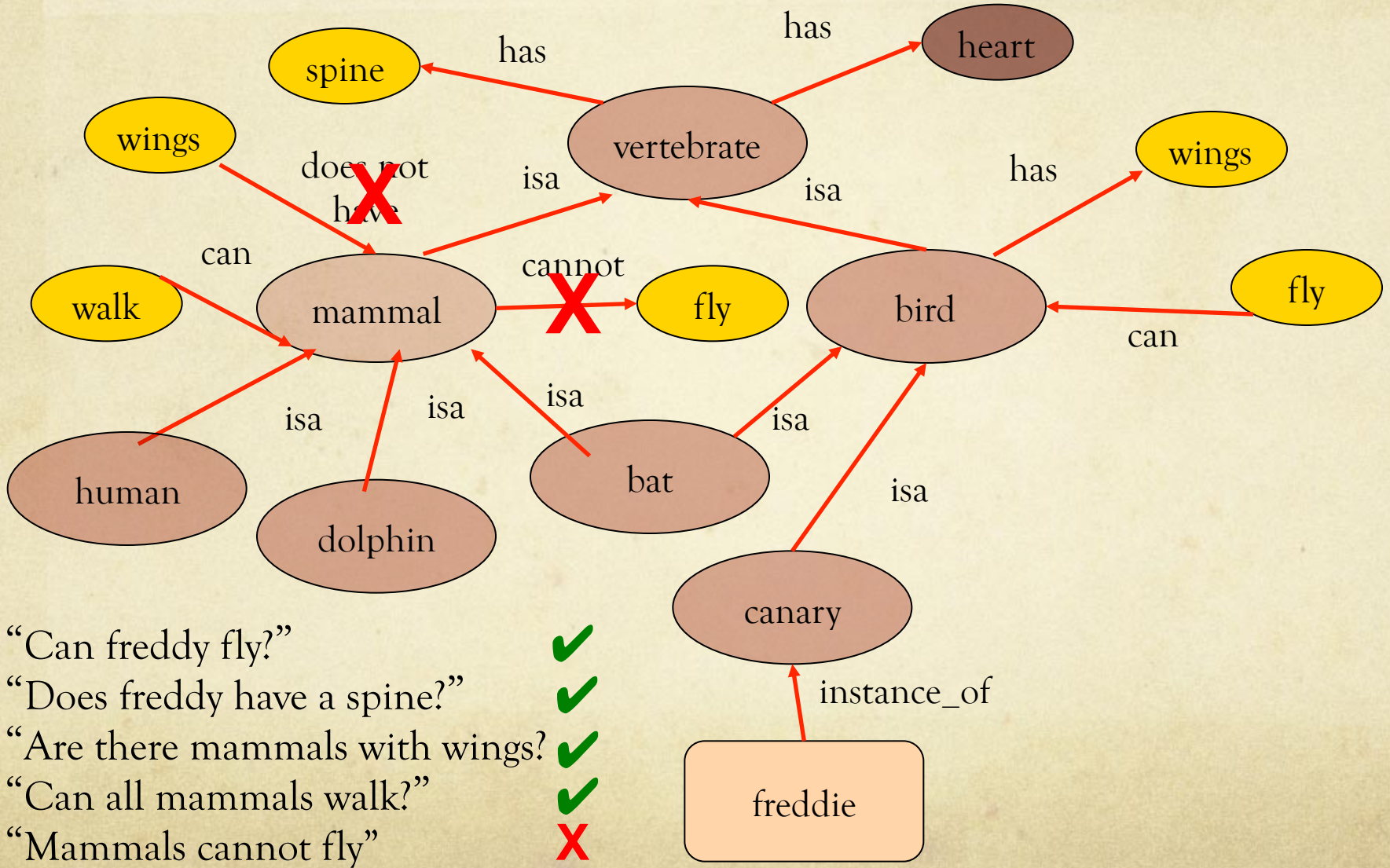
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? ?
"Can all mammals walk?" ✔
"Mammals cannot fly" ?

Semantic Networks



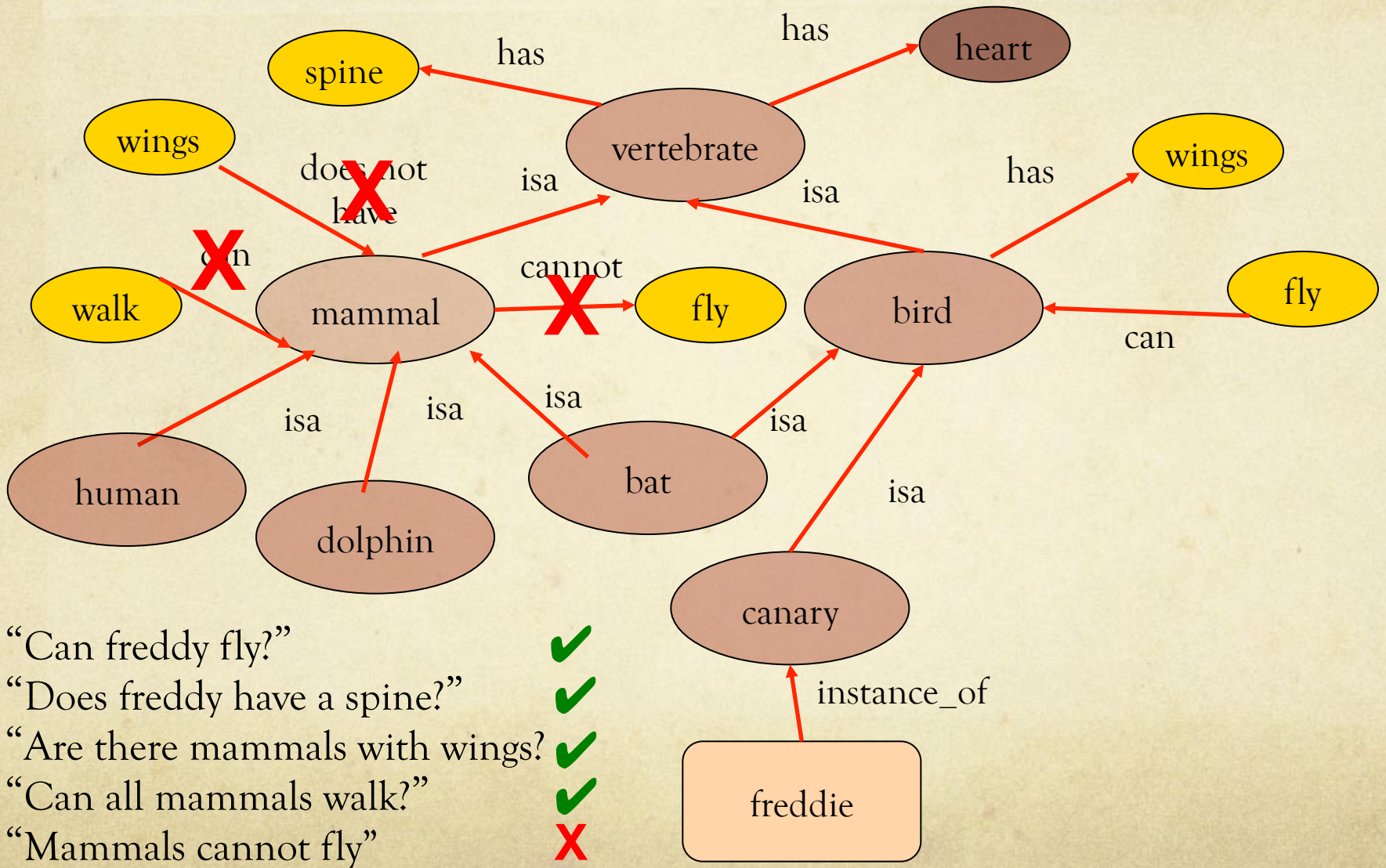
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? ✔
"Can all mammals walk?" ✔
"Mammals cannot fly" X

Semantic Networks



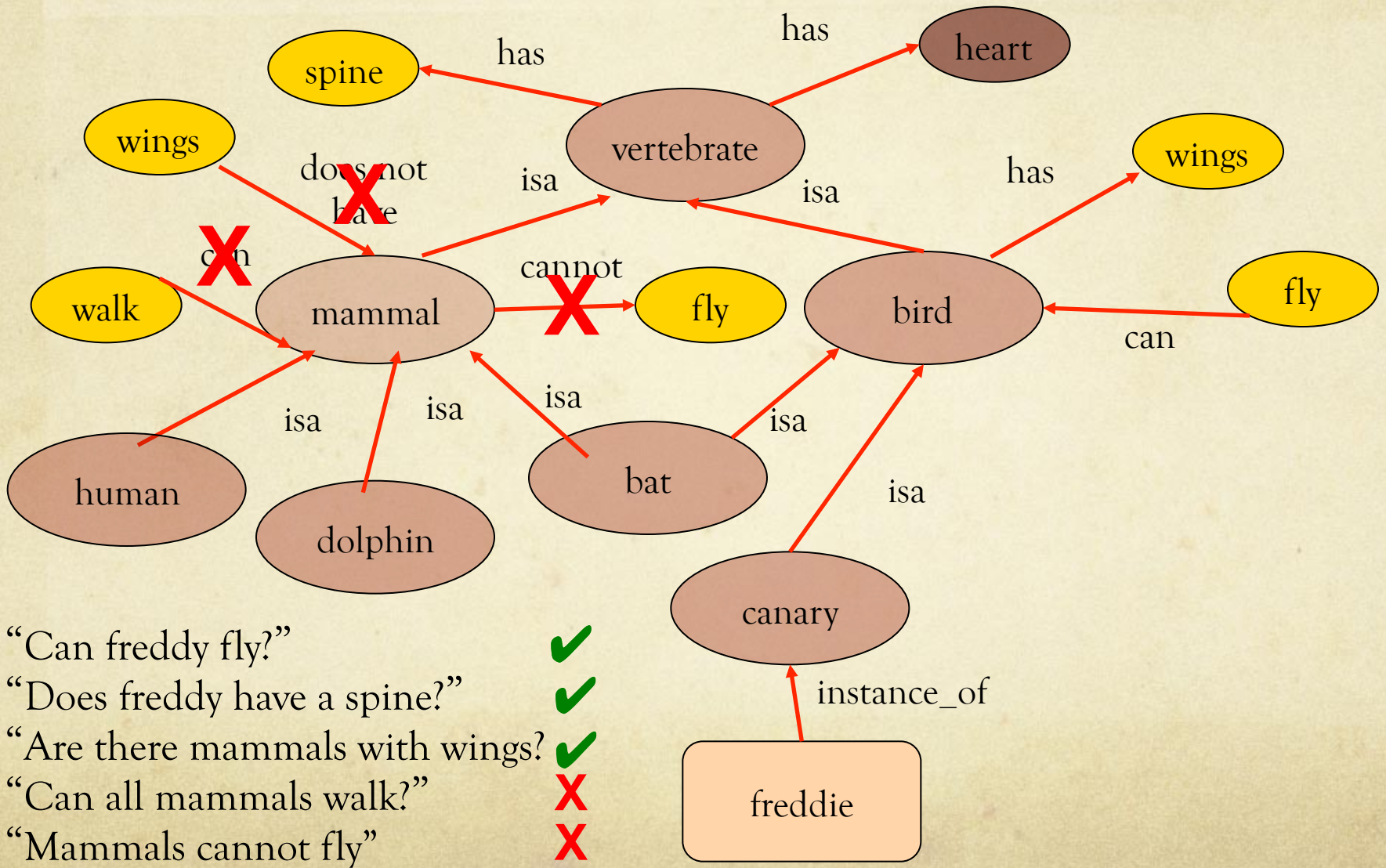
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? ✔
"Can all mammals walk?" ✔
"Mammals cannot fly" X

Semantic Networks



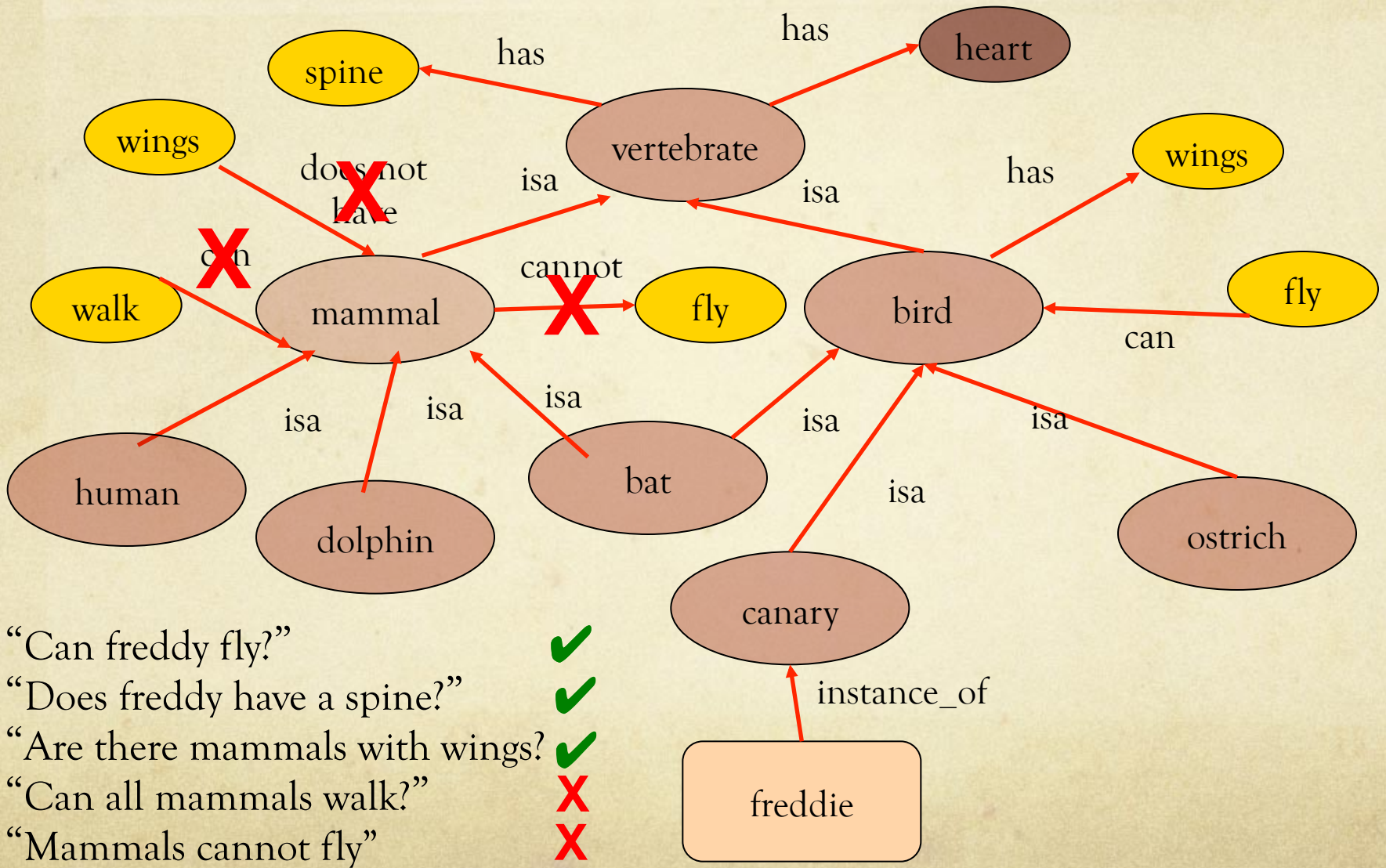
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? ✔
"Can all mammals walk?" ✔
"Mammals cannot fly" X

Semantic Networks



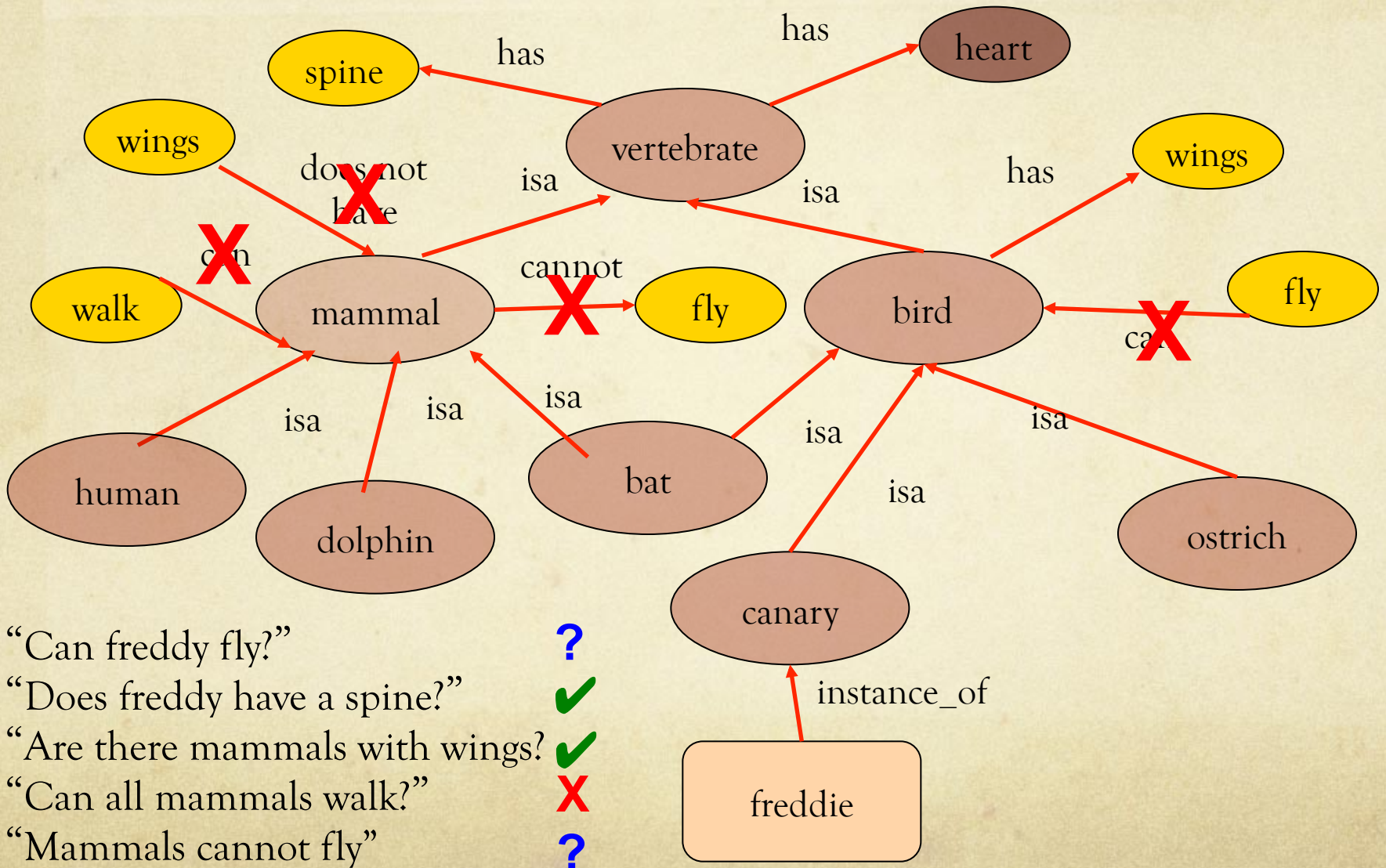
"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? ✔
"Can all mammals walk?" X
"Mammals cannot fly" X

Semantic Networks



"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? ✔
"Can all mammals walk?" X
"Mammals cannot fly" X

Semantic Networks



"Can freddy fly?" ?

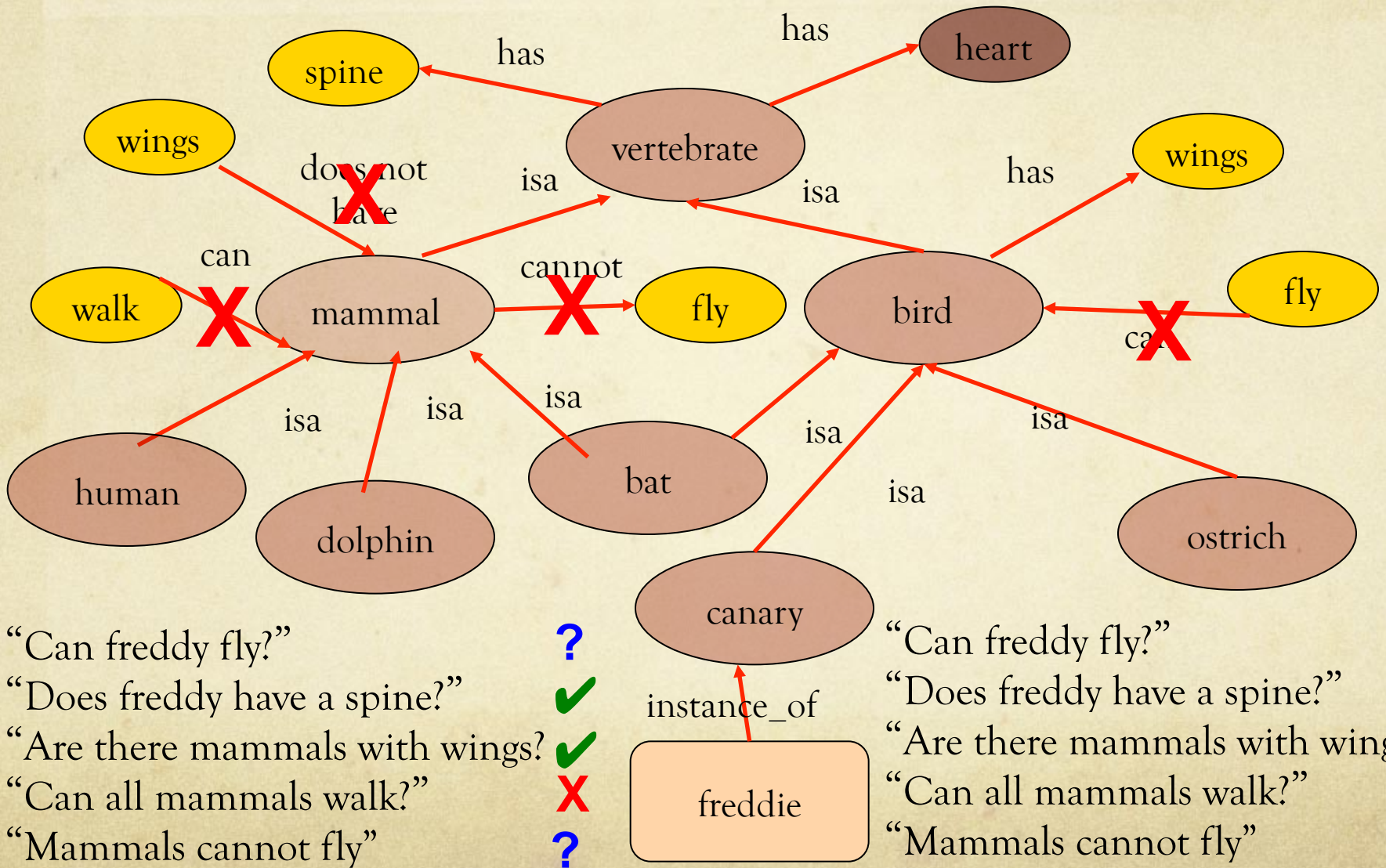
"Does freddy have a spine?" ✔

"Are there mammals with wings? ✔

"Can all mammals walk?" X

"Mammals cannot fly" ?

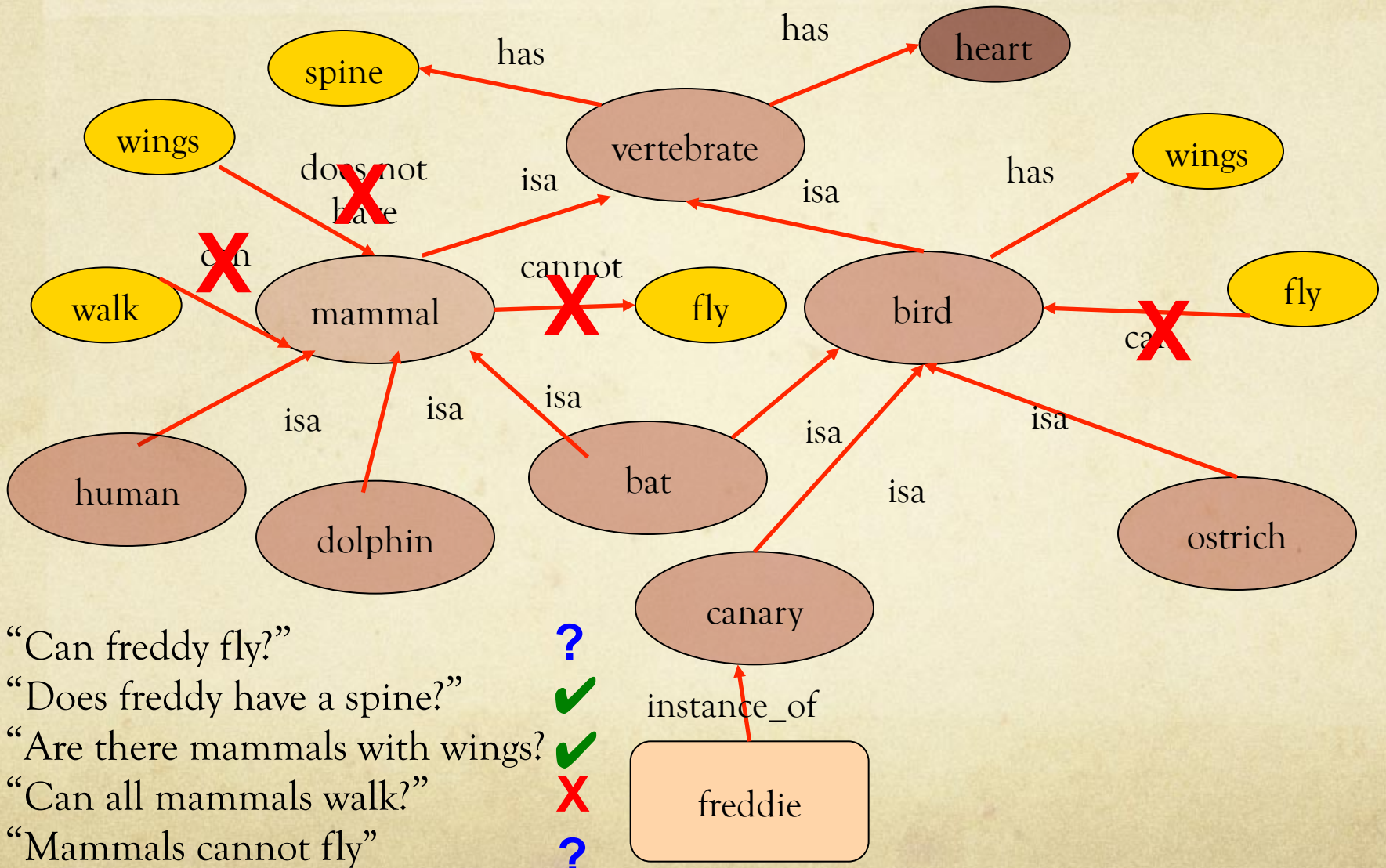
Semantic Networks



"Can freddy fly?" ?
"Does freddy have a spine?" ✔
"Are there mammals with wings? ✔
"Can all mammals walk?" X
"Mammals cannot fly" ?

"Can freddy fly?" ✔
"Does freddy have a spine?" ✔
"Are there mammals with wings? X
"Can all mammals walk?" ✔
"Mammals cannot fly" ✔

But what about Freddie ?



"Can freddy fly?" ?

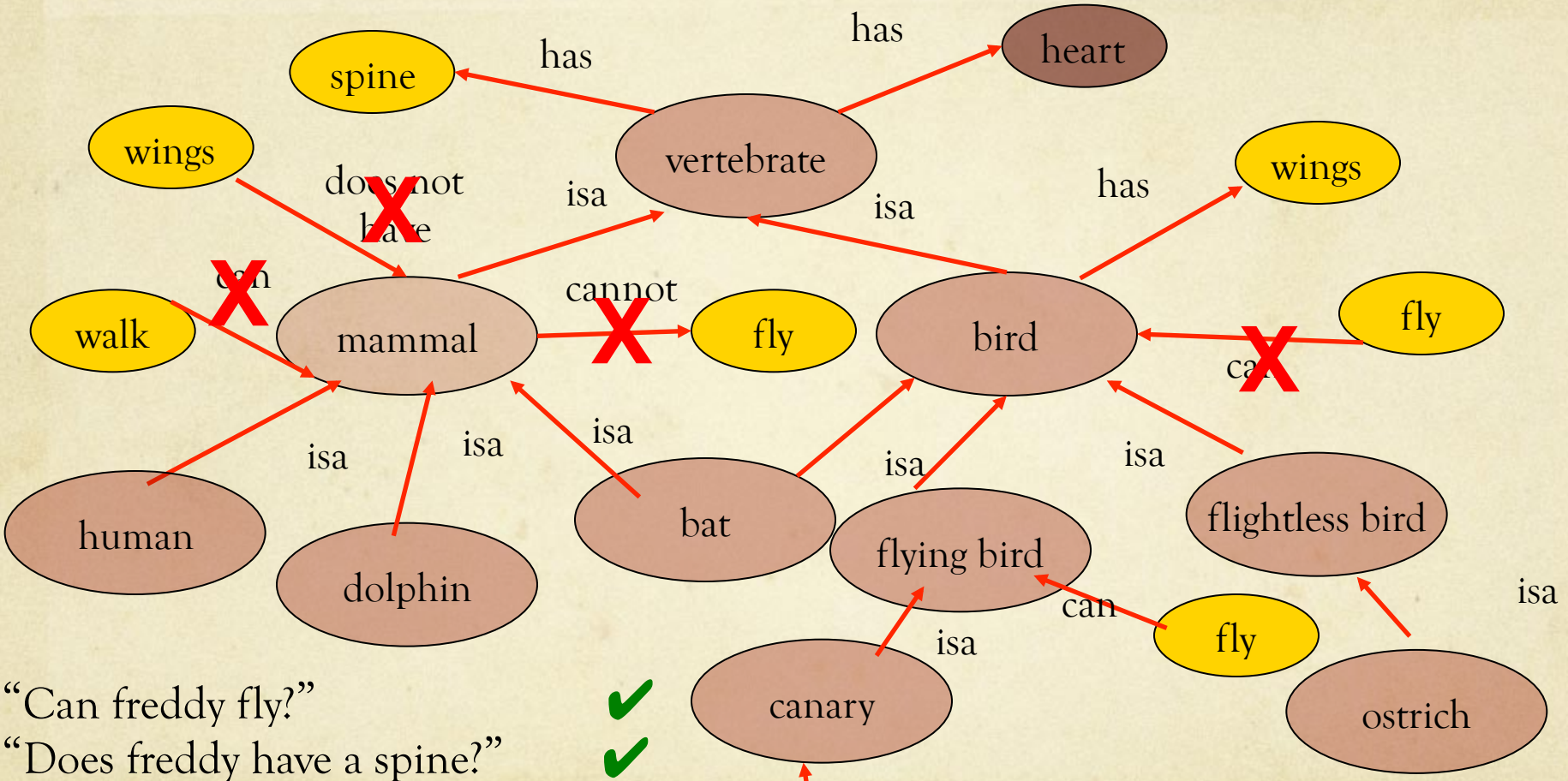
"Does freddy have a spine?" ✔

"Are there mammals with wings? ✔

"Can all mammals walk?" X

"Mammals cannot fly" ?

But what about Freddie ?



"Can freddy fly?" ✔

"Does freddy have a spine?" ✔

"Are there mammals with wings? ✔

"Can all mammals walk?" ✘

"Mammals cannot fly" ✘

Types of Ontologies

- Top (Upper) Level Ontologies
  - Sowa's top-level ontology
  - Cyc's upper ontology
  - etc
- Domain Ontologies (Domain/Application)
  - Biomedical Ontologies
  - Engineering Ontologies
  - Enterprise
  - Etc.

Tools for Ontology building

- **OilEd** from University of Manchester
  - http://oiled.man.ac.uk/
- **Ontolingua** from KSL (Stanford University)
  - http://www-ksl.stanford.edu
- **OntoSaurus** from ISI (USA)
  - http://www.isi.edu/isd/ontosaurus.html
- **OntoEdit** from Karlsrhue Univ.
  - http://ontoserver.aifb.unikarlsruhe.de/ontoedit/
- **Protégé** from SMI (Stanford University)
  - http://protege.stanford.edu/
- Karlsruhe
  - http://kaon.semanticweb.org/
- **WebOnto** from KMI (Open University)
  - http://kmi.open.ac.uk/projects/webonto/
- **WebODE** from UPM
  - http://webode.dia.fi.upm.es/webODE/
- **KAON** from AIFB and FZI at the University of Karlsruhe
  - http://kaon.semanticweb.org/
- OBO-Edit from OBO foundry
  - http://oboedit.org/

Upper level ontology

- Ontologies from different domains may be integrated by alignment to an upper level ontology.
- An upper-level ontology provides a common foundation for classes and relations [21].
- Typical classes found in upper-level ontologies include Process, Material object, Quality and Function.

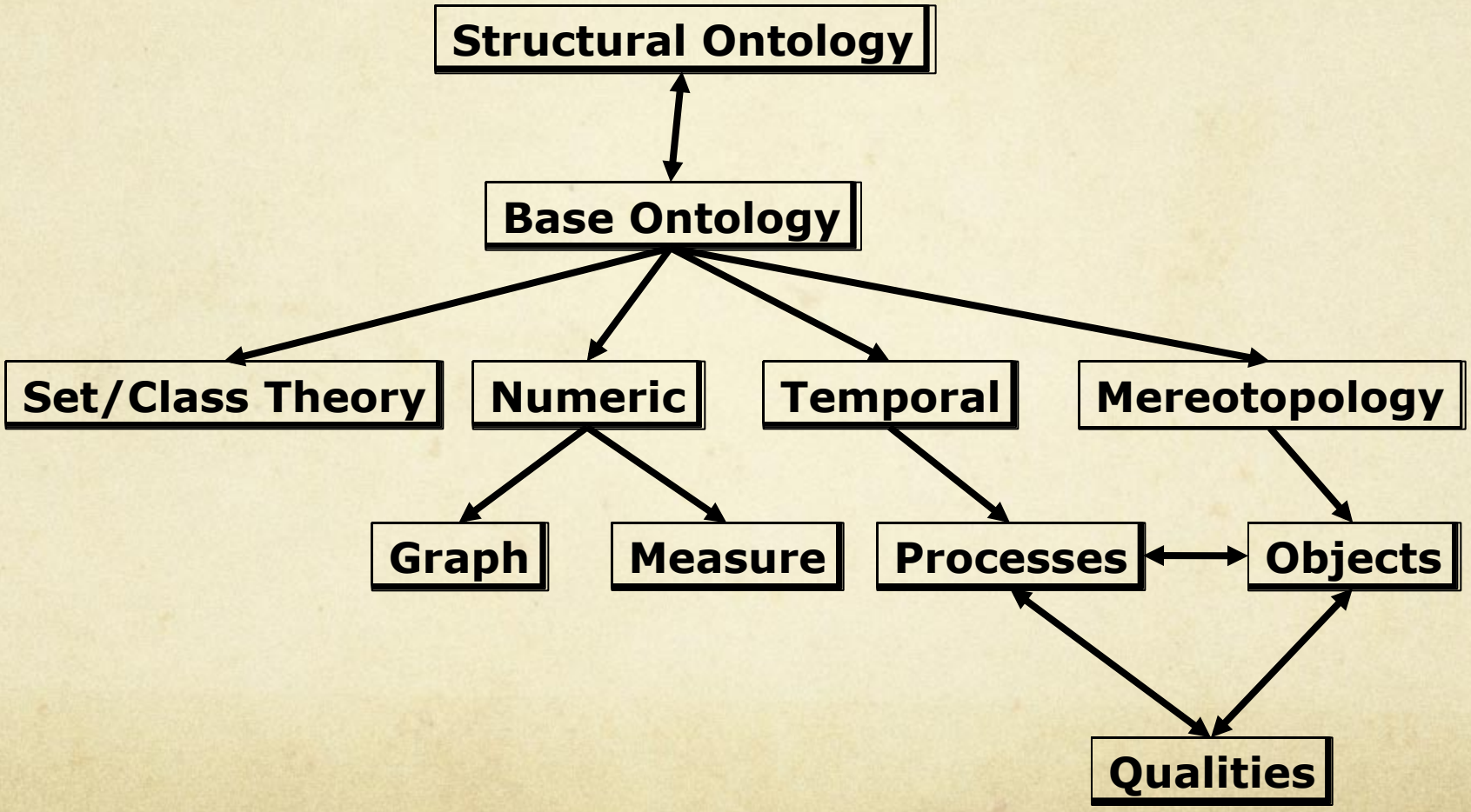
Upper level ontology

- Upper-level ontologies further provide relations that can hold between instances of their classes.
- Commonly included relations are has-part, has-participant and quality-of.
- Several upper-level ontologies are well established including the Basic Formal Ontology (BFO) [27], the Descriptive Ontology for Cognitive and Linguistic Engineering (DOLCE) [28], the SUMO ontology and the General Formal Ontology (GFO) [24].

Suggested Upper Merged Ontology (SUMO)

- 1000 terms, 4000 axioms, 750 rules
- Associated domain ontologies totalling 20,000 terms and 60,000 axioms
- http://www.ontologyportal.org

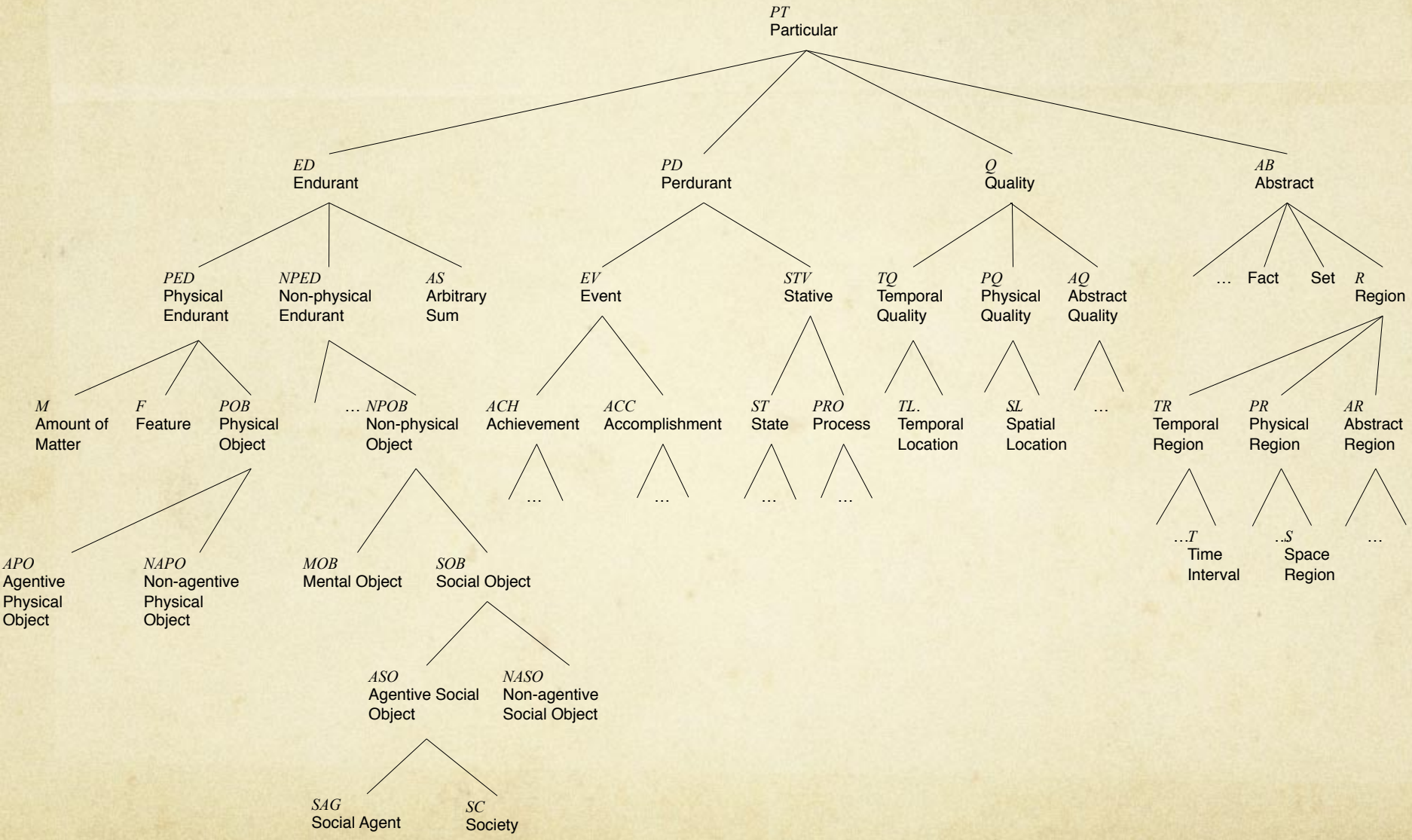
SUMO Taxonomy



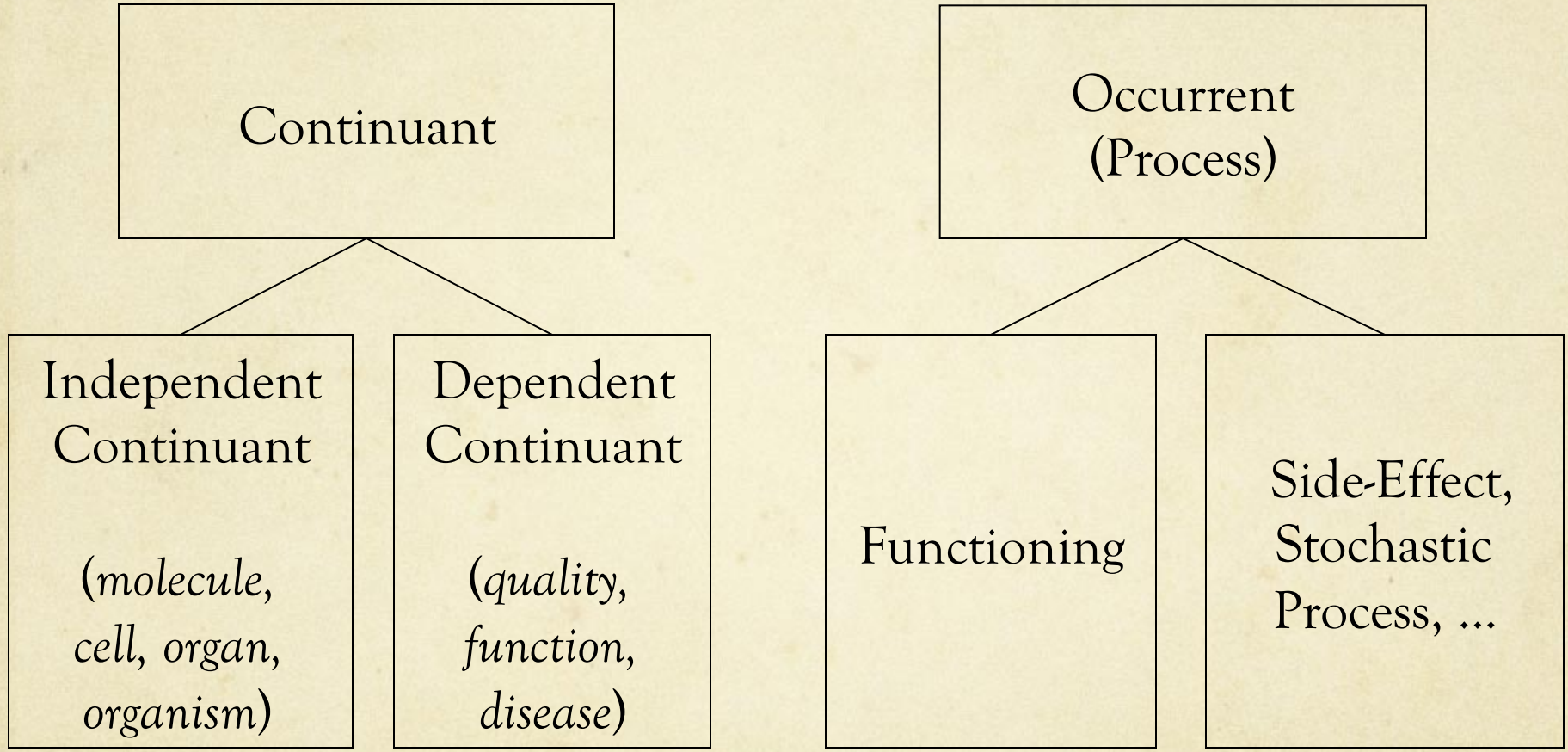
DOLCE: Descriptive Ontology for Linguistic and Cognitive Engineering

- Strong cognitive/linguistic bias:
  - *descriptive* (as opposite to *prescriptive*) attitude
  - Categories mirror cognition, common sense, and the lexical structure of natural language.
- Categories as *conceptual containers*: no "deep" metaphysical implications
- *Rich axiomatization*
  - *37 basic categories*
  - *7 basic relations*
  - *80 axioms, 100 definitions, 20 theorems*
- Rigorous quality criteria and extensive documentation

DOLCE taxonomy



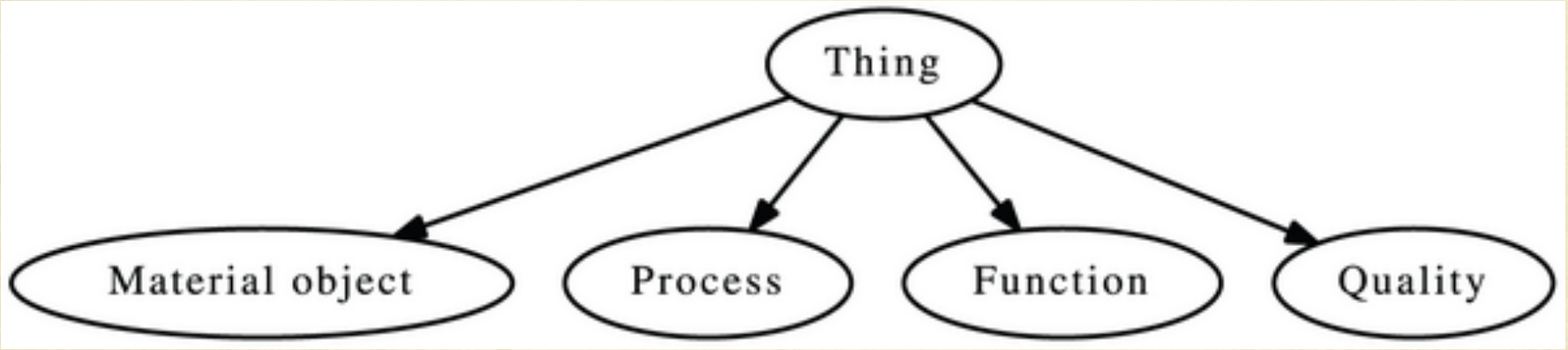
BFO Taxonomy



Basic Upper Level Ontology

- Consists of only four mutually disjoint classes: Material object, Process, Quality and Function.
- The instances of Material object exist with all their parts at a time point and need no other entity to exist.
- Processes, on the other hand, are temporally extended and cannot exist at a single time point.
- Functions are capabilities or potentials for the occurrence of processes [29] and depend on material objects.
- Qualities as attributes of other entities.

Taxonomy



Relations

inheres-in	*Quality*	*Thing*	**has-quality**
derives-from	*Material object*	*Material object*	
has-participant	*Process*	*Material object*	**participates-in**
has-input	*Process*	*Material object*	**input-of**
has-output	*Process*	*Material object*	**output-of**
has-central-participant	*Process*	*Material object*	**central-participant-of**
part-of	*Thing*	*Thing*	**has-part**
proper-part-of	*Thing*	*Thing*	**has-proper-part**
realized-by	*Function*	*Process*	**realizes**
results-in	*Process*	*Process*	

Gene Ontology

- Built for a very specific purpose:

"annotation of genes and proteins in genomic and protein databases"

- Applicable to all species
- Three disjoint axes:
  - molecular function
    - molecular role e.g. catalytic activity, binding
  - biological process
    - broad biological phenomena e.g. mitosis, growth, digestion
  - cellular component
    - sub-cellular location e.g nucleus, ribosome, origin recognition complex

Gene Ontology Project

- Started in 1998
- Primary Goals
  - Structured Vocabulary
  - Use to annotate genes and gene products
- 3 Model Organisms
  - FlyBase (Drosophila)
  - Saccharomyces Genome Database (SGD)
  - Mouse Genome Informatics (MGI) project

GENE ONTOLOGY™ CONSORTIUM

FlyBase

http://www.geneontology.org

European Bioinformatics Institute

UNIVERSITY OF CAMBRIDGE

InterPro

Harvard University

swissprot

The Wellcome Trust Sanger Institute

INDIANA UNIVERSITY

RGD RAT GENOME DATABASE

MGI Mouse Genome Informatics

IncyteGenomics For Life.

SGD

WormBase

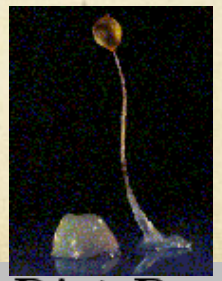
TIGR THE INSTITUTE FOR GENOMIC RESEARCH

compugen

tair

AstraZeneca

Genome KnowledgeBase



DictyBase

GO is wildly successful

- Biologists around the world contribute to GO on a regular basis
- The ontology is updated every 30 minutes!
- GO is now used in a variety of domains and application in computational biology
- The advent of GO gave raise to the generation of various biomedical ontologies that cover a wide range of biomedical domains and are employed in various applications

Anatomy and Development Anatomy Ontologies

- Examples of species specific
  - Foundational Model of Anatomy (FMA)
  - Mouse Anatomy Ontology
  - Plant Anatomy
  - Drosophila Gross Anatomy (FBBT)
  - Zebrafish anatomy and development (ZFA)
  - Human developmental anatomy (EHDAA2)
  - Mouse gross anatomy and development (EMAP)
- Species independent
  - UBERON

Foundational Model of Anatomy (FMA)

- A representation of classes or types and relationships necessary for the symbolic representation of the phenotypic structure of the human body in a form that is understandable to humans and is also navigable, parseable and interpretable by machine-based systems.
- Over 75,000 classes and over 120,000 terms; over 2.1 million relationship instances from over 168 relationship types
- One of the largest computer based resources in biomedical sciences

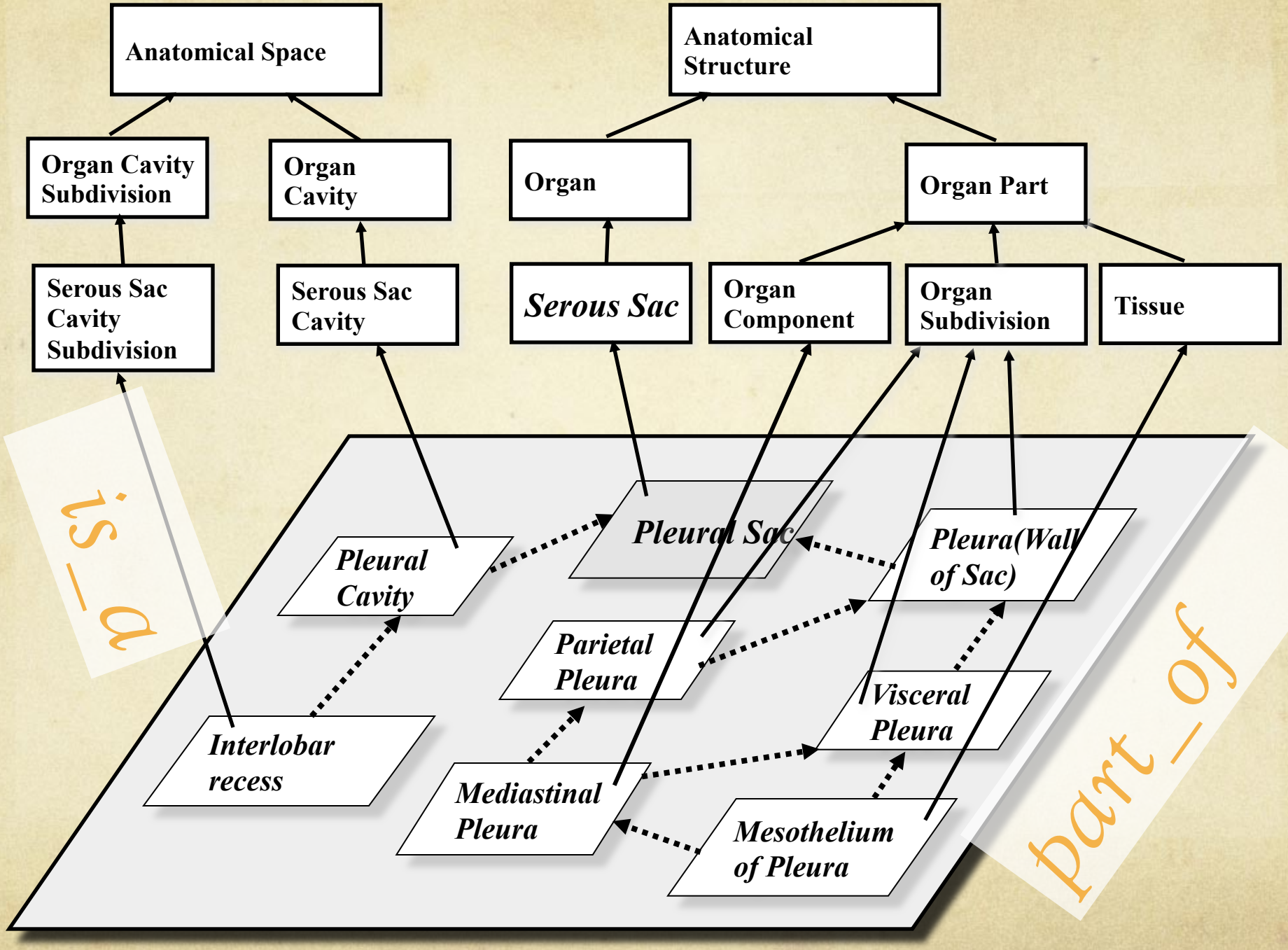
FMA organisation

- organized in a graph-theoretical structure involving two sorts of links or edges:
- **is-a** (= is a subtype of)

  pleural sac **is-a** serous sac

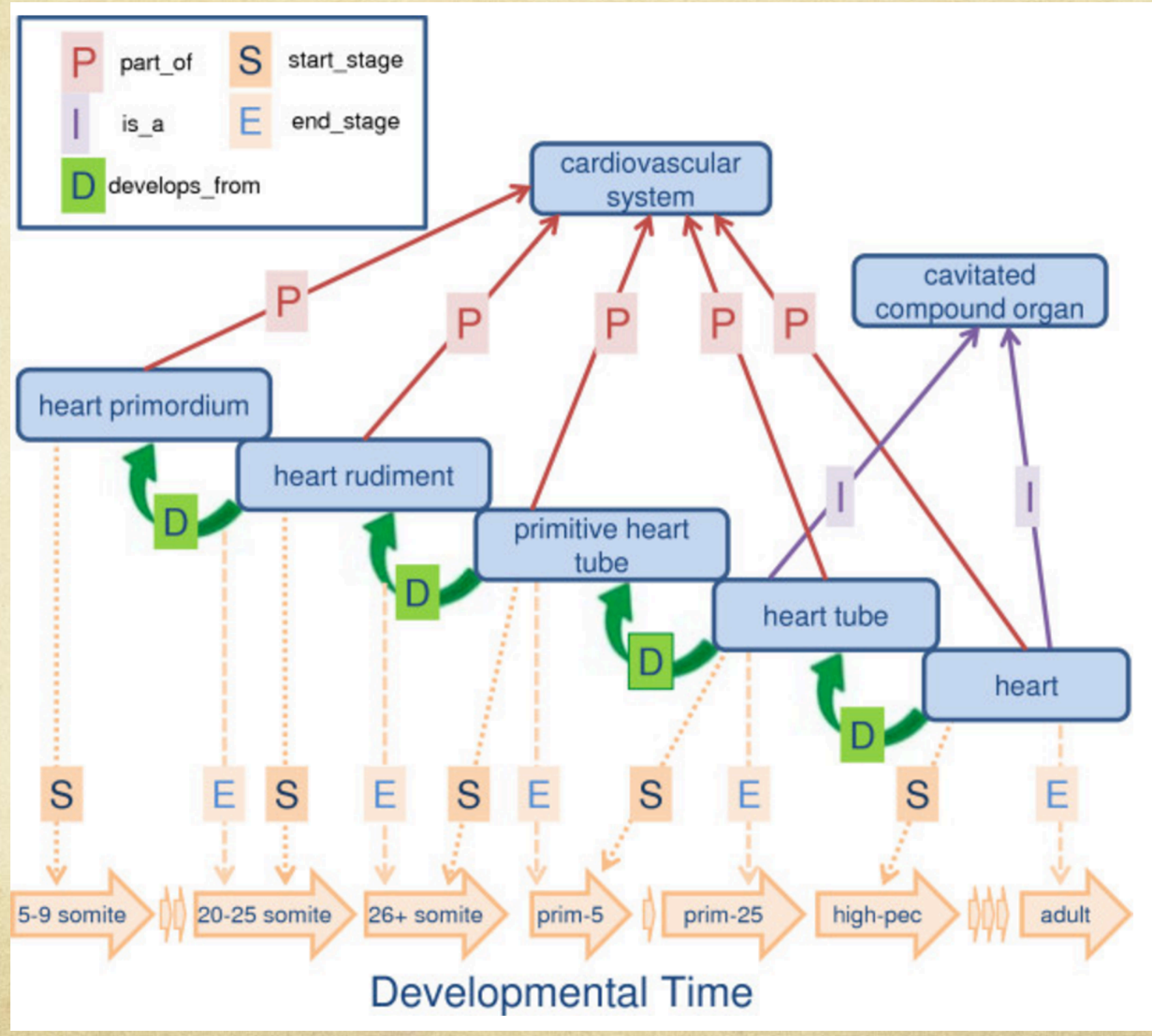
- **part-of**

  cervical vertebra **part-of** vertebral column

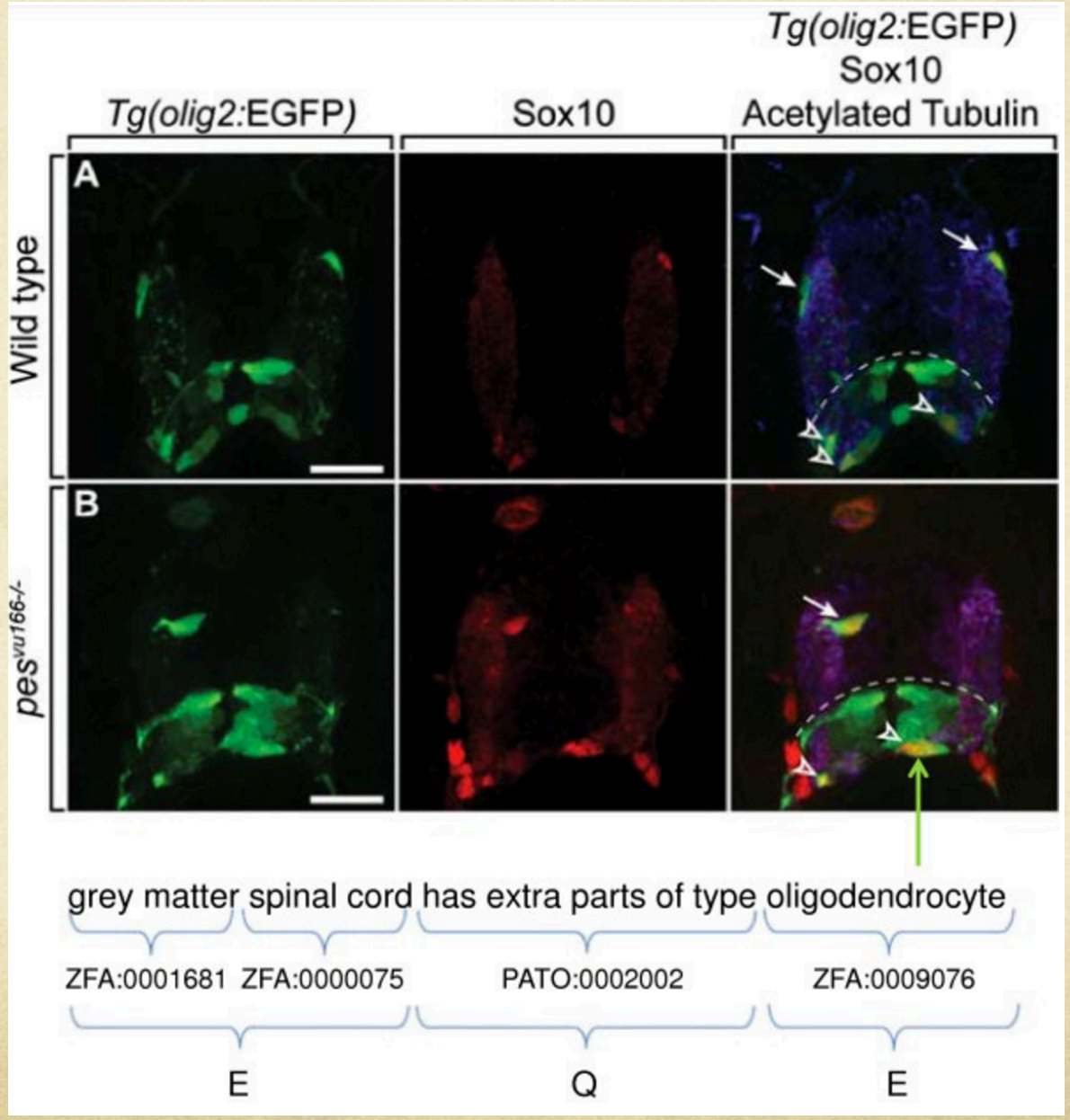

Anatomical Space
Anatomical Structure
Organ Cavity Subdivision
Organ Cavity
Organ
Organ Part
Serous Sac Cavity Subdivision
Serous Sac Cavity
Serous Sac
Organ Component
Organ Subdivision
Tissue
is_a
Pleural Cavity
Pleural Sac
Pleura(Wall of Sac)
Parietal Pleura
Visceral Pleura
Interlobar recess
Mediastinal Pleura
Mesothelium of Pleura
part_of

Zebrafish Anatomy Ontology (ZFA)

- ZFIN has developed the Zebrafish Anatomy Ontology (ZFA) [13] and the Zebrafish Stage Ontology (ZFS) to annotate gene expression and phenotypic data related to zebrafish
- ZFA is used in conjunction with the Zebrafish Stage Ontology (ZFS) to describe the gross and cellular anatomy and development of the zebrafish, Danio rerio, from single cell zygote to adult.


P part_of
S start_stage
I is_a
E end_stage
D develops_from
cardiovascular system
cavitated compound organ
heart primordium
heart rudiment
primitive heart tube
heart tube
heart
P
I
D
S
E
5-9 somite
20-25 somite
26+ somite
prim-5
prim-25
high-pec
adult
Developmental Time

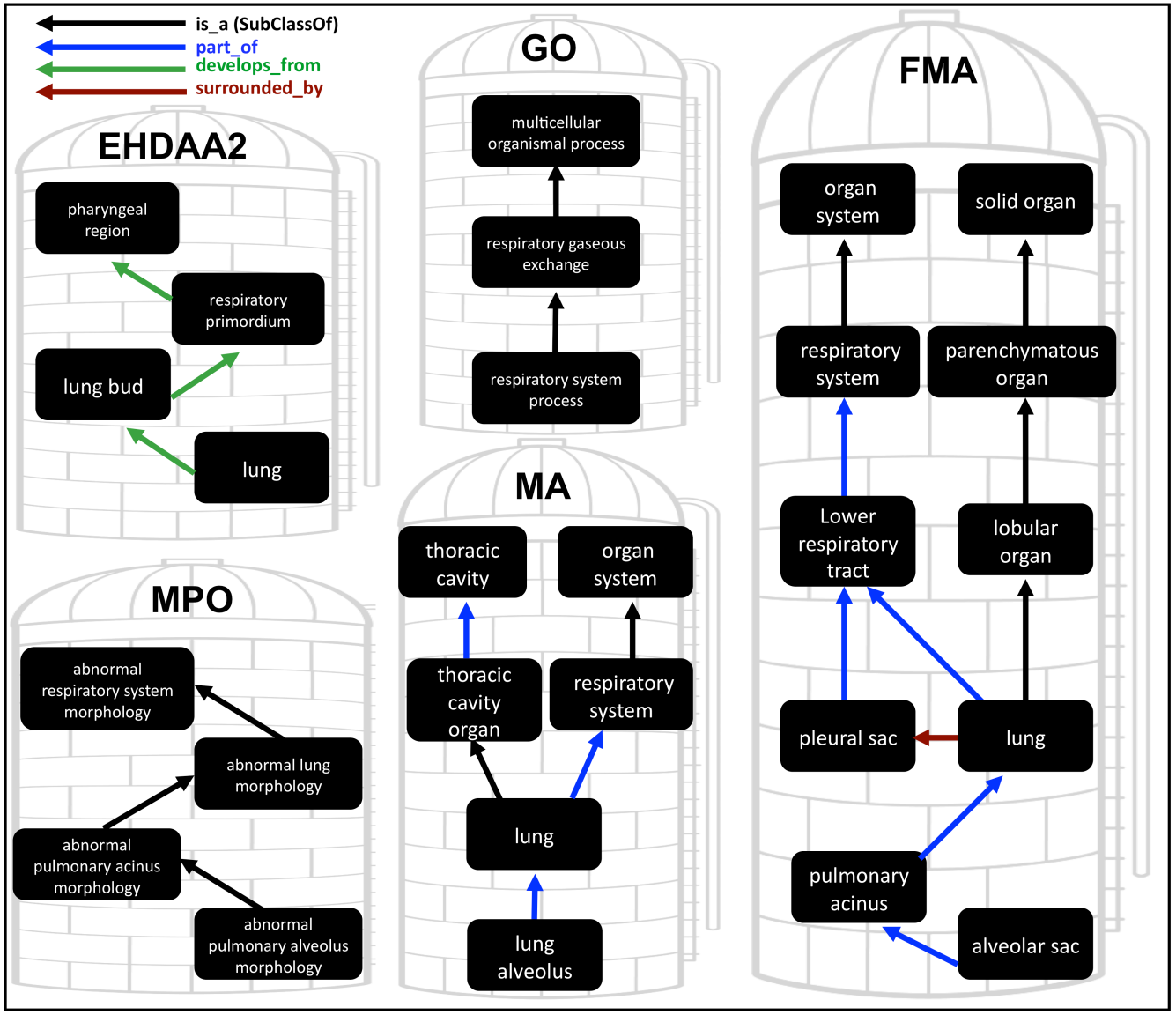
ZFA use for phenotype annotation



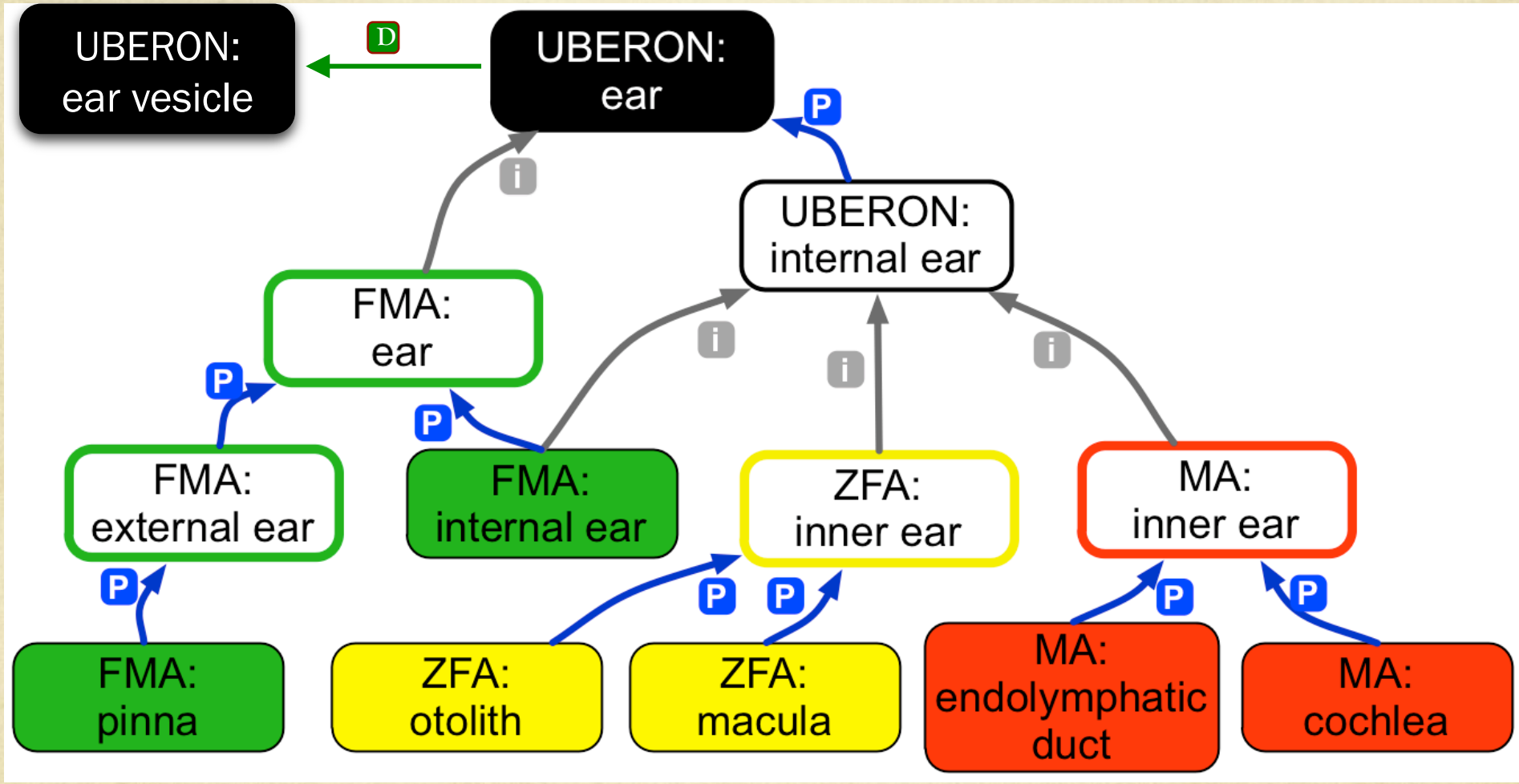
UBERON

- Existing anatomical ontologies are incomplete and poorly integrated, leading to Balkanization of data into separate silos.
- Integration of ontologies spanning multiple animal species is particularly important for translating data derived from model organisms.
- UBERON - a species-neutral resource for describing body parts.

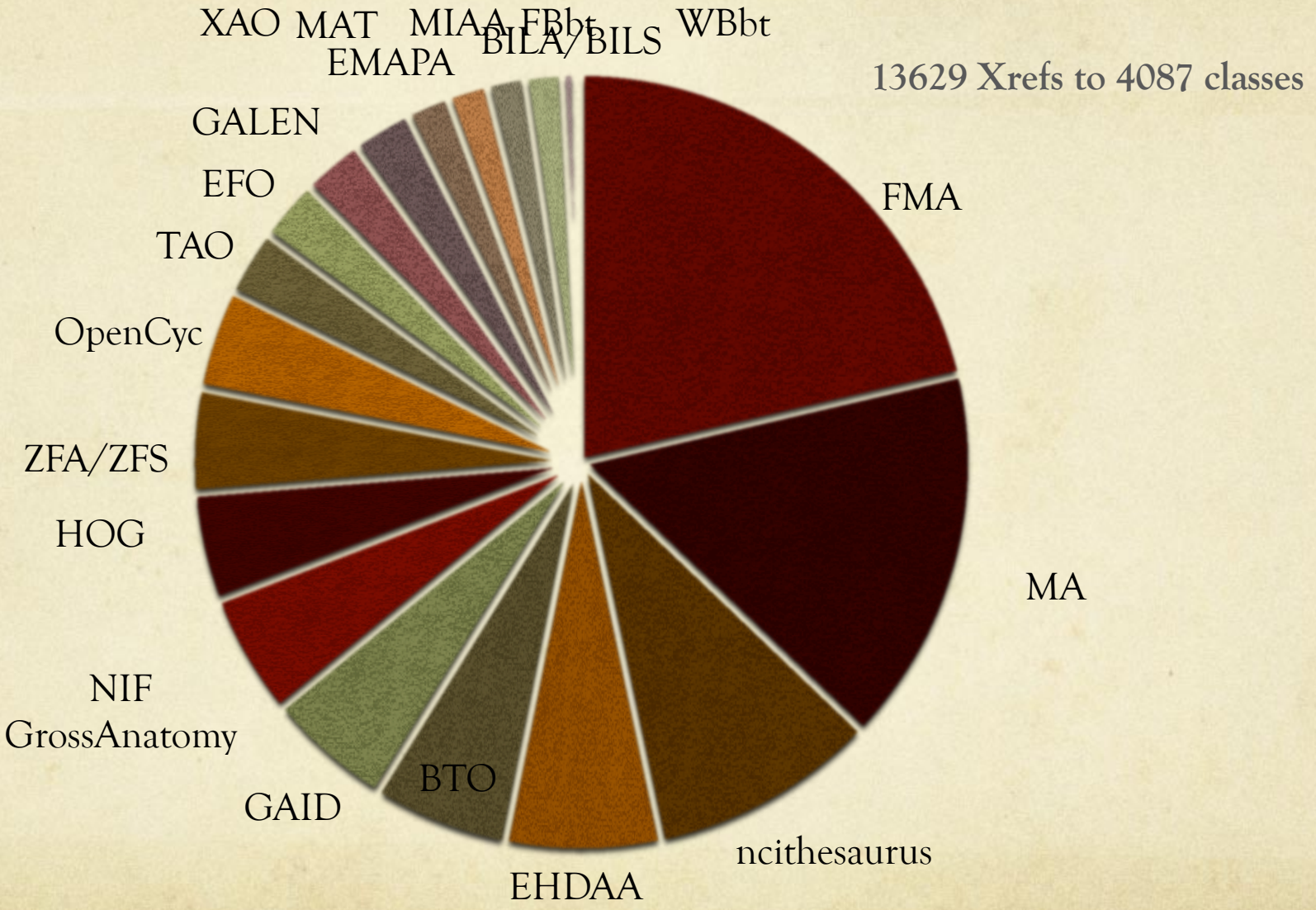
Unifying Species Centric Anatomies



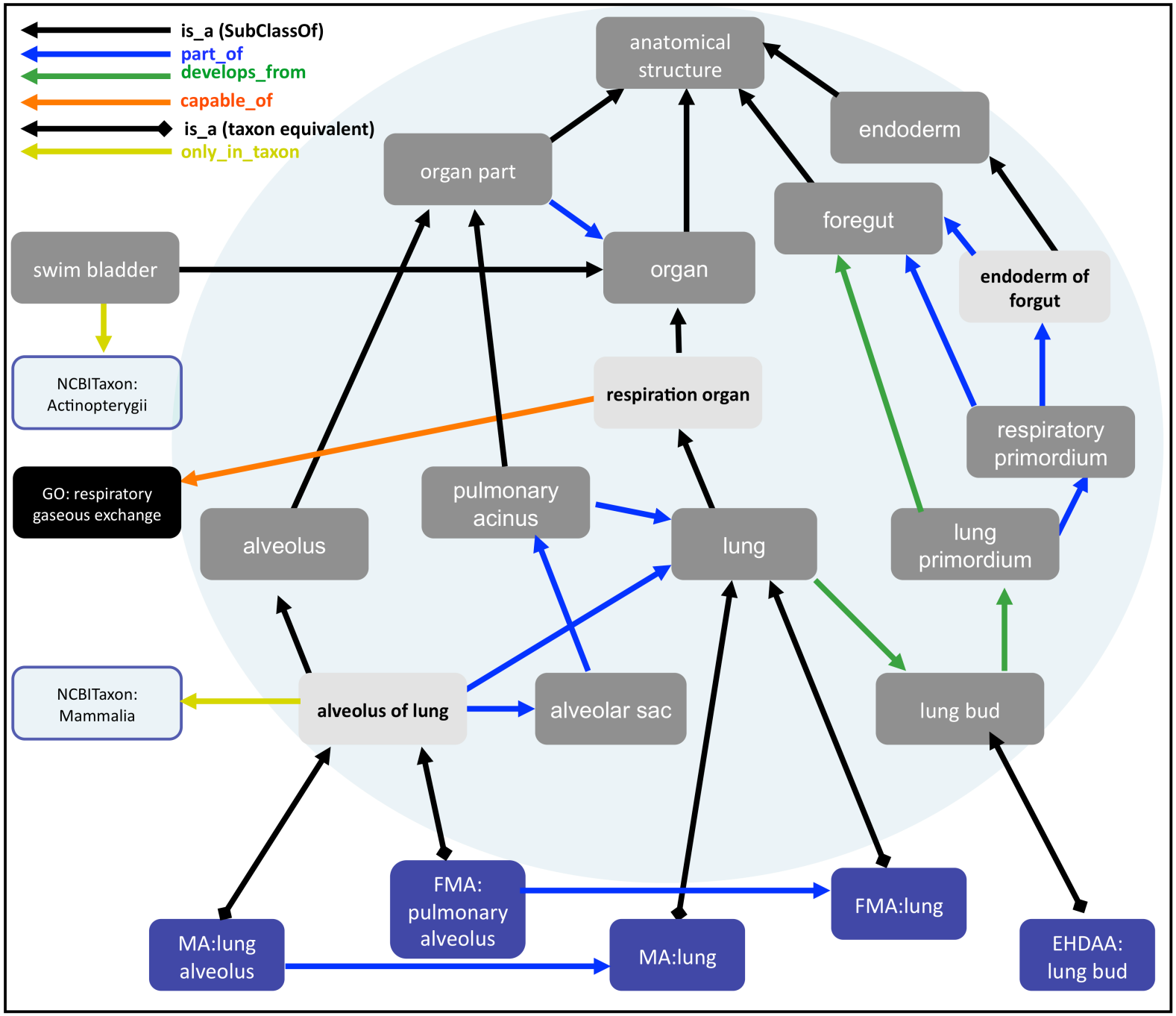
UBERON species neutral classes



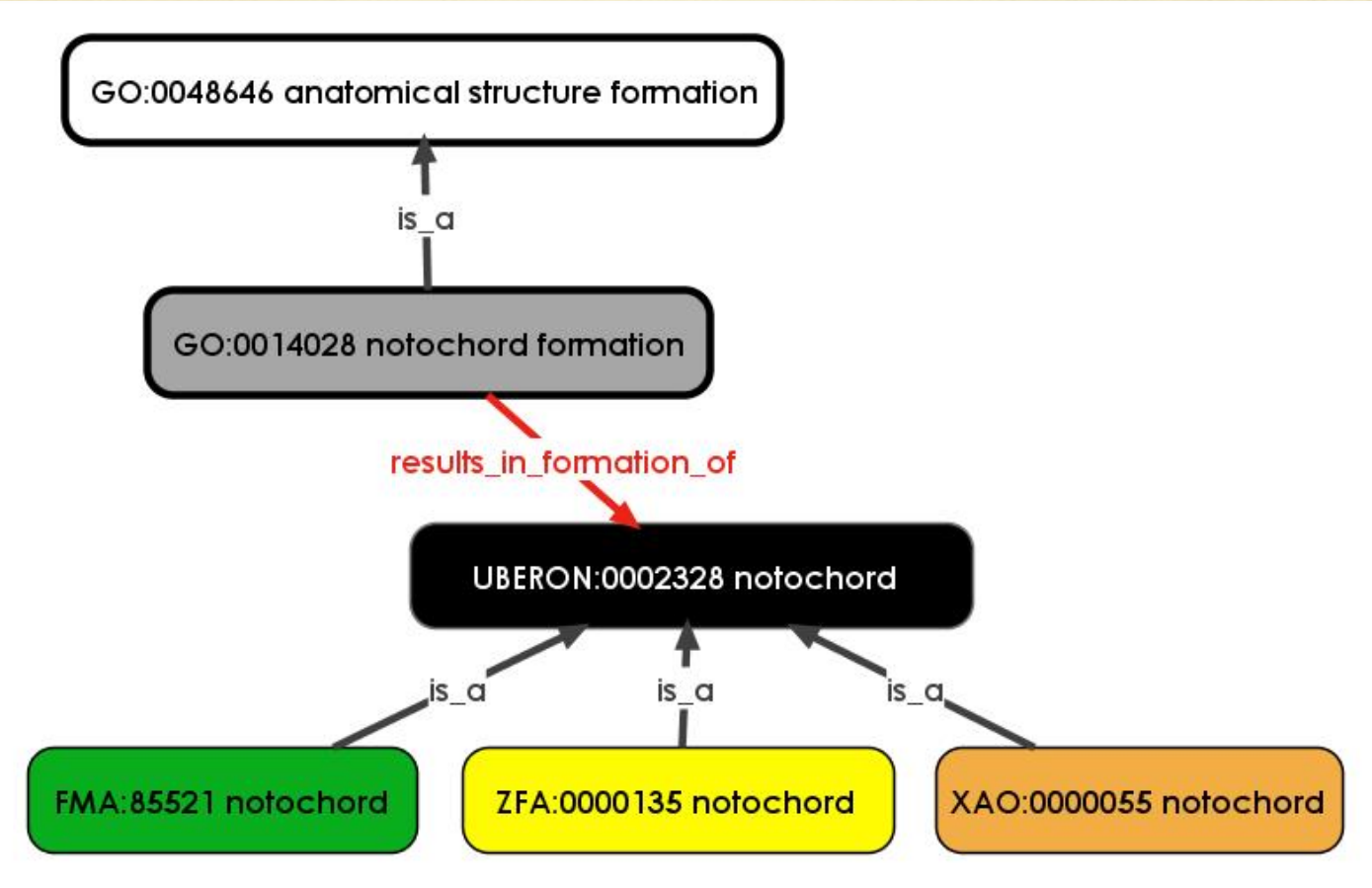
UBERON composition



Uberon classes subsume classes in contributing AOs

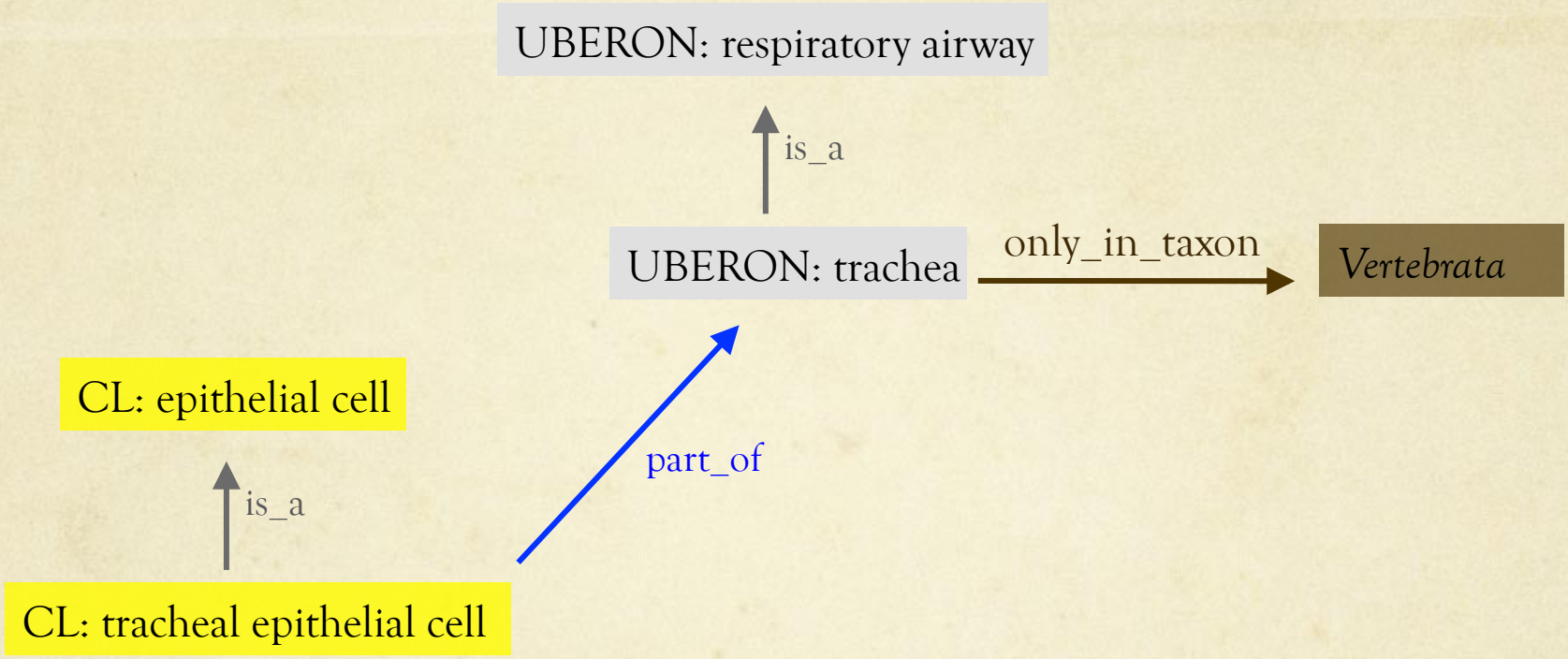

is_a (SubClassOf)
part_of
develops_from
capable_of
is_a (taxon equivalent)
only_in_taxon
anatomical structure
endoderm
organ part
foregut
swim bladder
organ
endoderm of forgut
NCBITaxon: Actinopterygii
respiration organ
respiratory primordium
GO: respiratory gaseous exchange
pulmonary acinus
alveolus
lung
lung primordium
NCBITaxon: Mammalia
alveolus of lung
alveolar sac
lung bud
FMA: pulmonary alveolus
FMA:lung
MA:lung alveolus
MA:lung
EHDAA: lung bud

Logical definitions in GO using UBERON



GO:notochord formation: The formation of the notochord from the chordamesoderm. The notochord is composed of large cells packed within a firm connective tissue sheath and is found in all chordates at the ventral surface of the neural tube. In vertebrates, the notochord contributes to the vertebral column.

Logical definitions in CL using UBERON



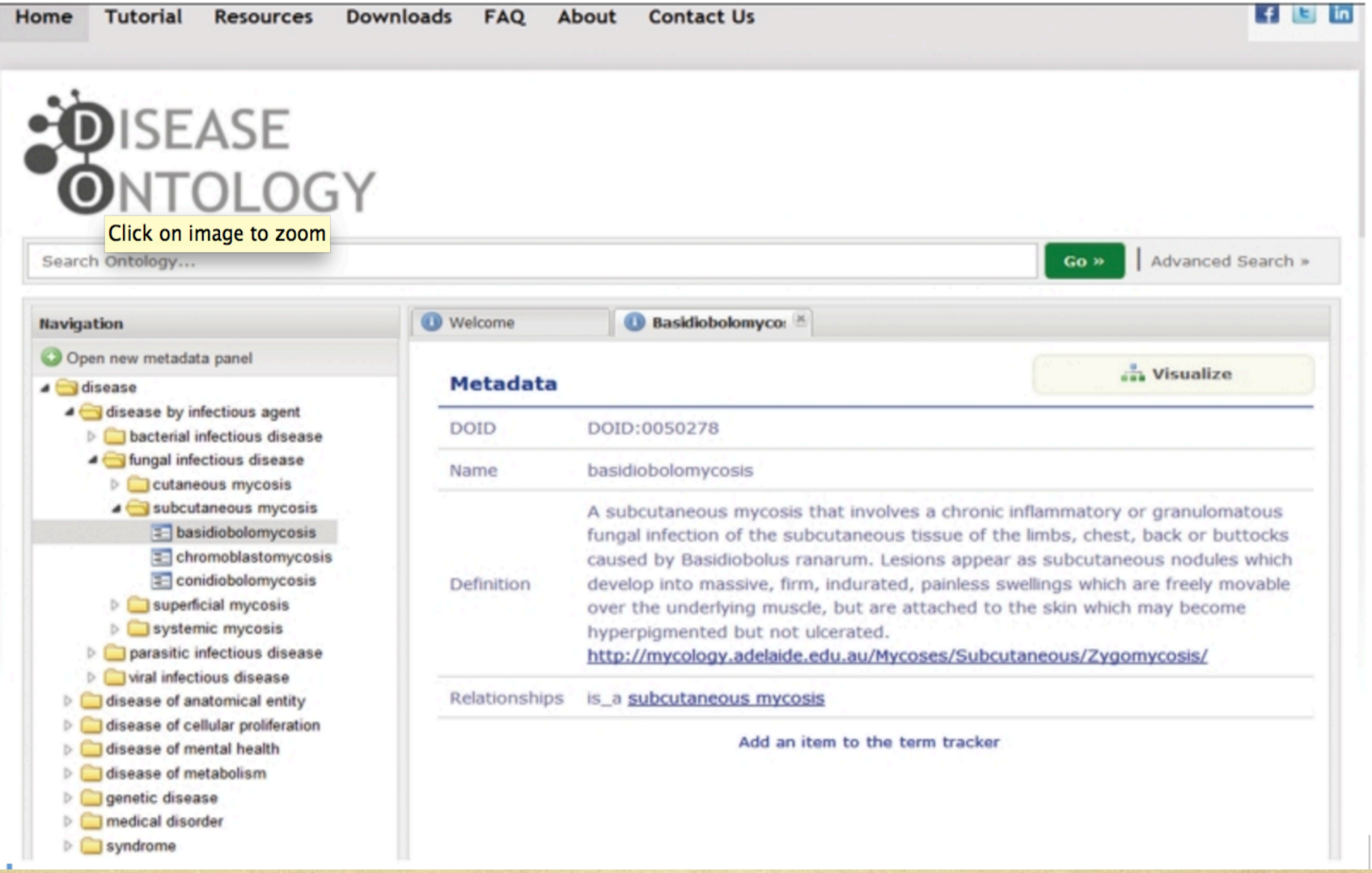
Uberon trachea: A trachea held open by up to 20 C-shaped rings of cartilage. The trachea is the portion of the airway that attaches to the bronchi as it branches. [GO:0060438, Wikipedia:Vertebrate_trachea]

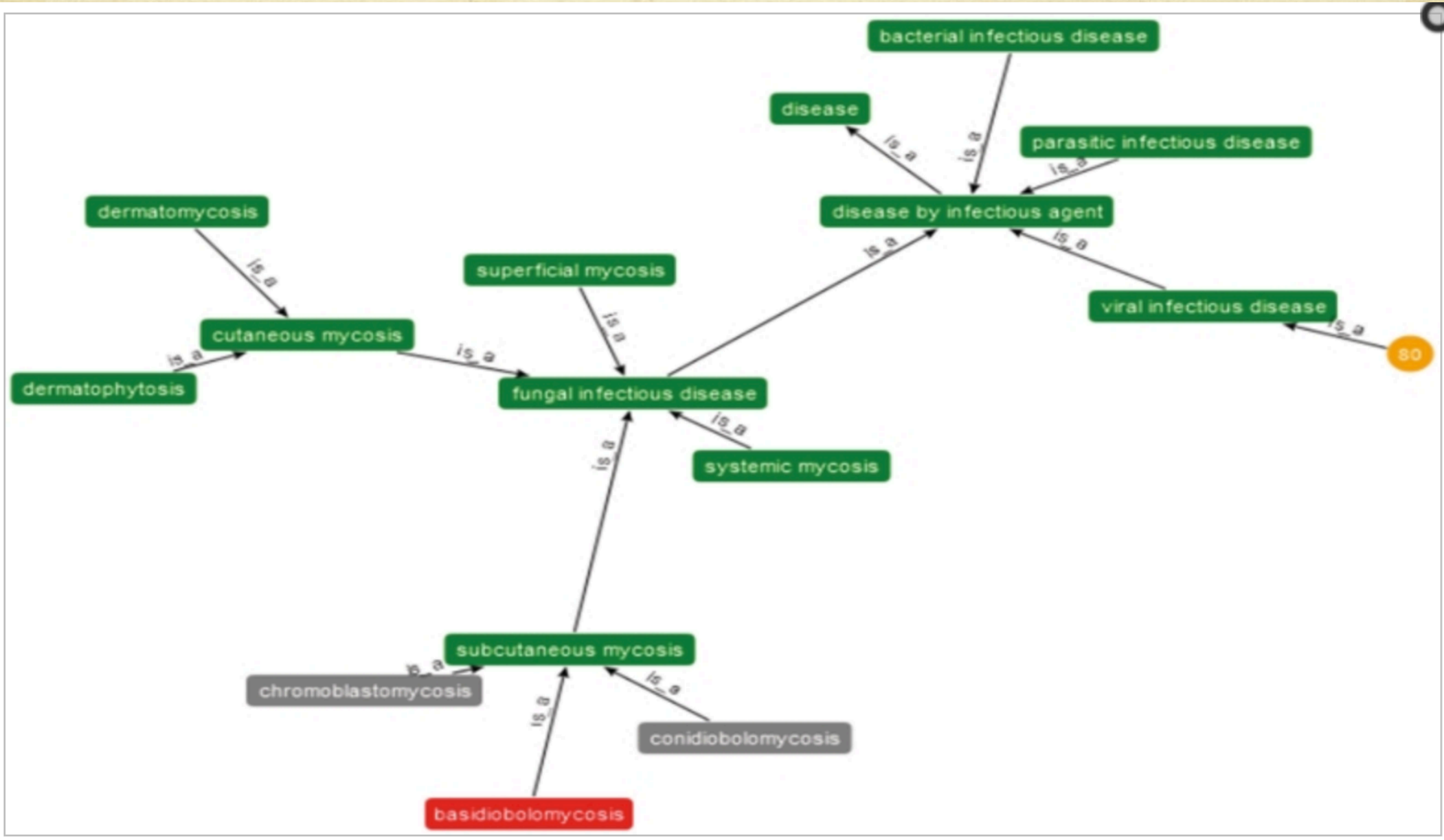
Examples of Disease Ontologies

- Disease Ontology (DO)
- Infectious Disease Ontology (IDO)
- Primary immunodeficiency disease Ontology (PIDO)
- Alzheimer Disease Ontology (ADO)

Disease Ontology (DO)

- The DO semantically integrates disease and medical vocabularies through extensive cross mapping and integration of MeSH, ICD, NCI's thesaurus, SNOMED CT and OMIM disease-specific terms and identifiers.
- Over 8000 inherited, developmental and acquired human diseases


Home
Tutorial
Resources
Downloads
FAQ
About
Contact Us
DISEASE ONTOLOGY
Click on image to zoom
Search Ontology...
Go »
Advanced Search »
Navigation
Open new metadata panel
disease
disease by infectious agent
bacterial infectious disease
fungal infectious disease
cutaneous mycosis
subcutaneous mycosis
basidiobolomycosis
chromoblastomycosis
conidiobolomycosis
superficial mycosis
systemic mycosis
parasitic infectious disease
viral infectious disease
disease of anatomical entity
disease of cellular proliferation
disease of mental health
disease of metabolism
genetic disease
medical disorder
syndrome
Welcome
Basidiobolomyco
Metadata
Visualize
DOID
DOID:0050278
Name
basidiobolomycosis
Definition
A subcutaneous mycosis that involves a chronic inflammatory or granulomatous fungal infection of the subcutaneous tissue of the limbs, chest, back or buttocks caused by Basidiobolus ranarum. Lesions appear as subcutaneous nodules which develop into massive, firm, indurated, painless swellings which are freely movable over the underlying muscle, but are attached to the skin which may become hyperpigmented but not ulcerated.
http://mycology.adelaide.edu.au/Mycoses/Subcutaneous/Zygomycosis/
Relationships
is_a subcutaneous mycosis
Add an item to the term tracker

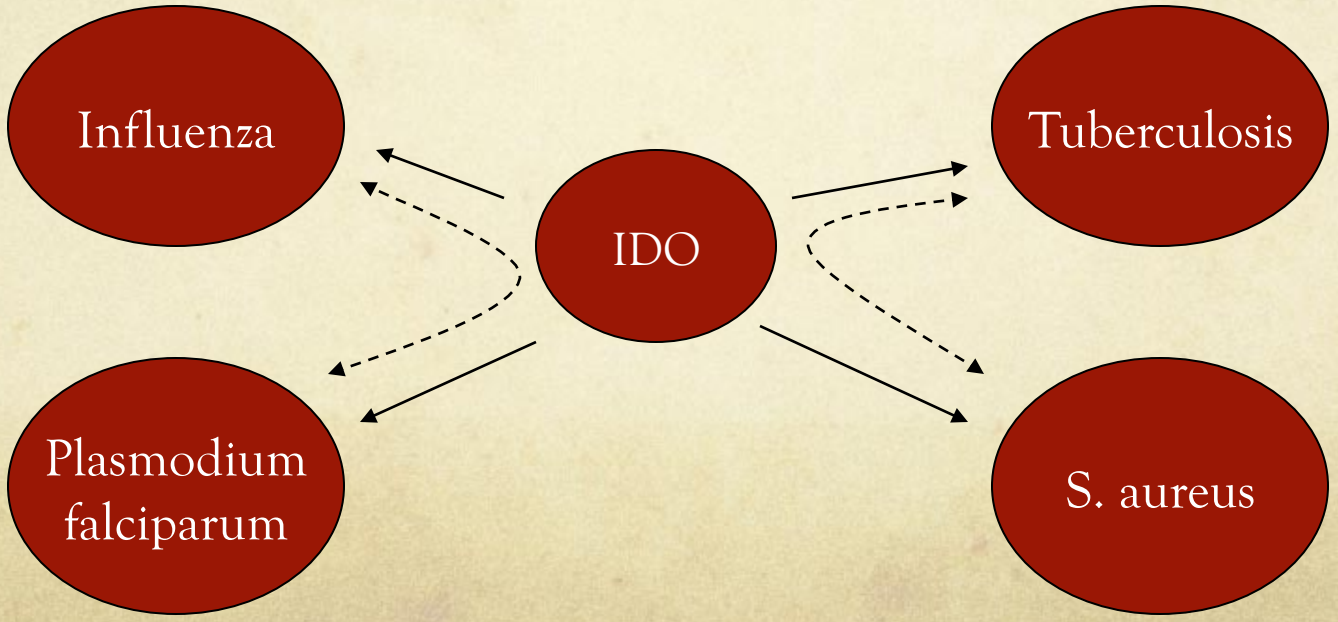

bacterial infectious disease
disease
parasitic infectious disease
disease by infectious agent
dermatomycosis
superficial mycosis
viral infectious disease
so
cutaneous mycosis
dermatophytosis
fungal infectious disease
systemic mycosis
subcutaneous mycosis
chromoblastomycosis
conidiobolomycosis
basidiobolomycosis
is_a

Infectious Disease Ontology (IDO)

- Biomedical Research (sequence data, cellular data, ...)
  - Pathogens, vectors, patients, model organisms
  - Microbiology, immunology, ...
- Vector Ecology Research
- Epidemiological Data for surveillance, prevention
- Clinical Care (case report data)
  - Clinical phenotypes, signs, symptoms
  - Treatments
  - Patient outcomes
- Clinical trial data for drugs, vaccines

Cross-domain Interoperability

- Disease- and organism-specific ontologies
- Built as refinements to a template infectious disease ontology with terms relevant to a large number of infectious diseases



IDO and other ontologies

- Anatomical entities in IDO
  - **Anatomical location: FMA**: e.g. lung, kidney
  - **Protein: PRO**: e.g virulence factors such as Eap
  - **Cell: CL**: e.g. macrophages

- Process in IDO
  - Imported from **GO BP** when possible

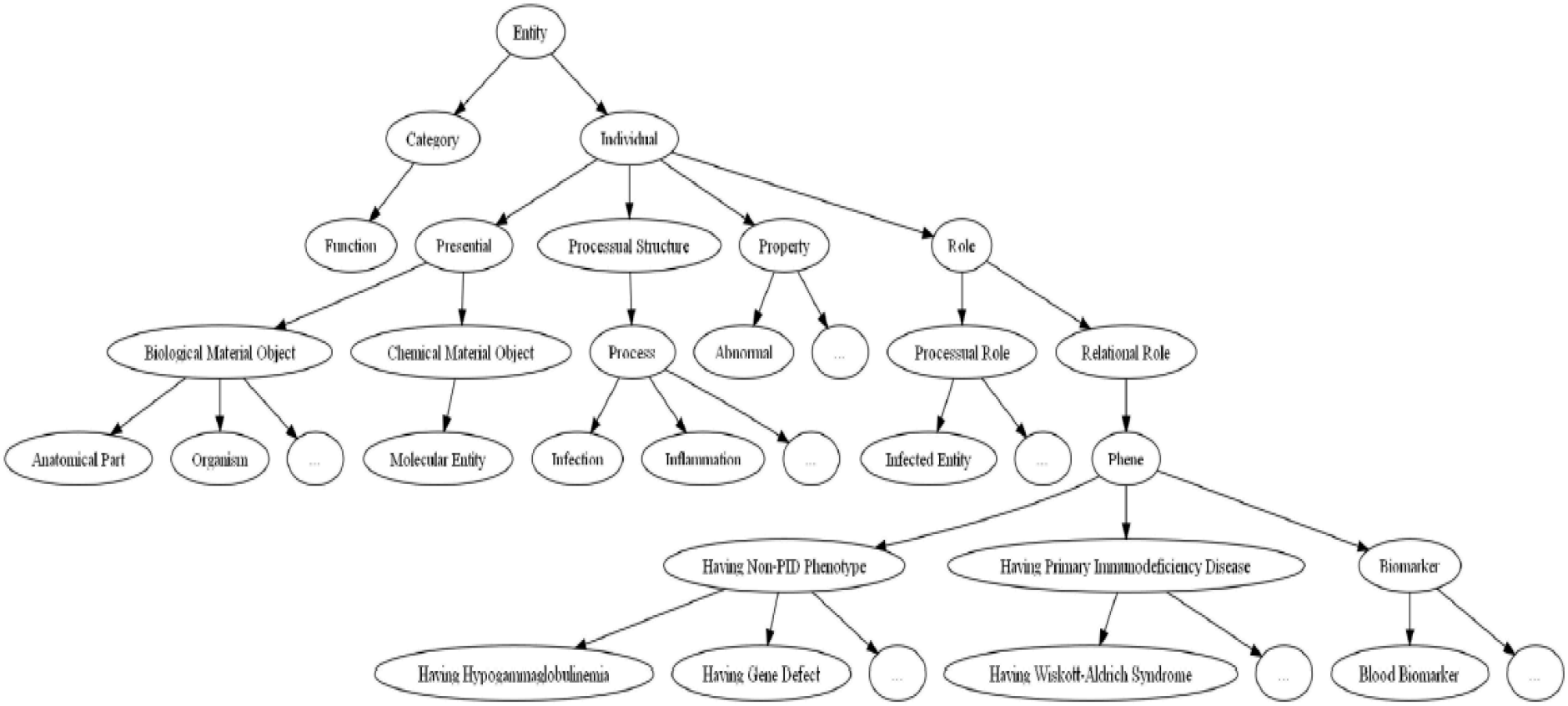
  e.g. GO:0044406 : adhesion to host

- Quality:
  - PATO: e.g. attenuated, susceptible, co-infected, immunocompromised, drug resistant, zoonotic

Primary Immunodeficiency Disease Ontology

- PIDO characterizes Primary immunodeficiency diseases in terms of the phenotypes commonly observed by clinicians during a diagnosis process
- Phenotype terms in PIDO are formally defined using complex definitions based on qualities, functions, processes, and structures.
- PIDO connects immunological knowledge across resources within a common framework and thereby enables translational research and the development of medical applications for the domain of immunology and primary immunodeficiency diseases.

PIDO organisation



PIDO definitions

- Elevated level of anti-DNA antibodies present in the serum

```
Phene and phene_of some (has_part some (Anti-
DNA_Antibody and (has_property some
Increased_Concentration)))
```

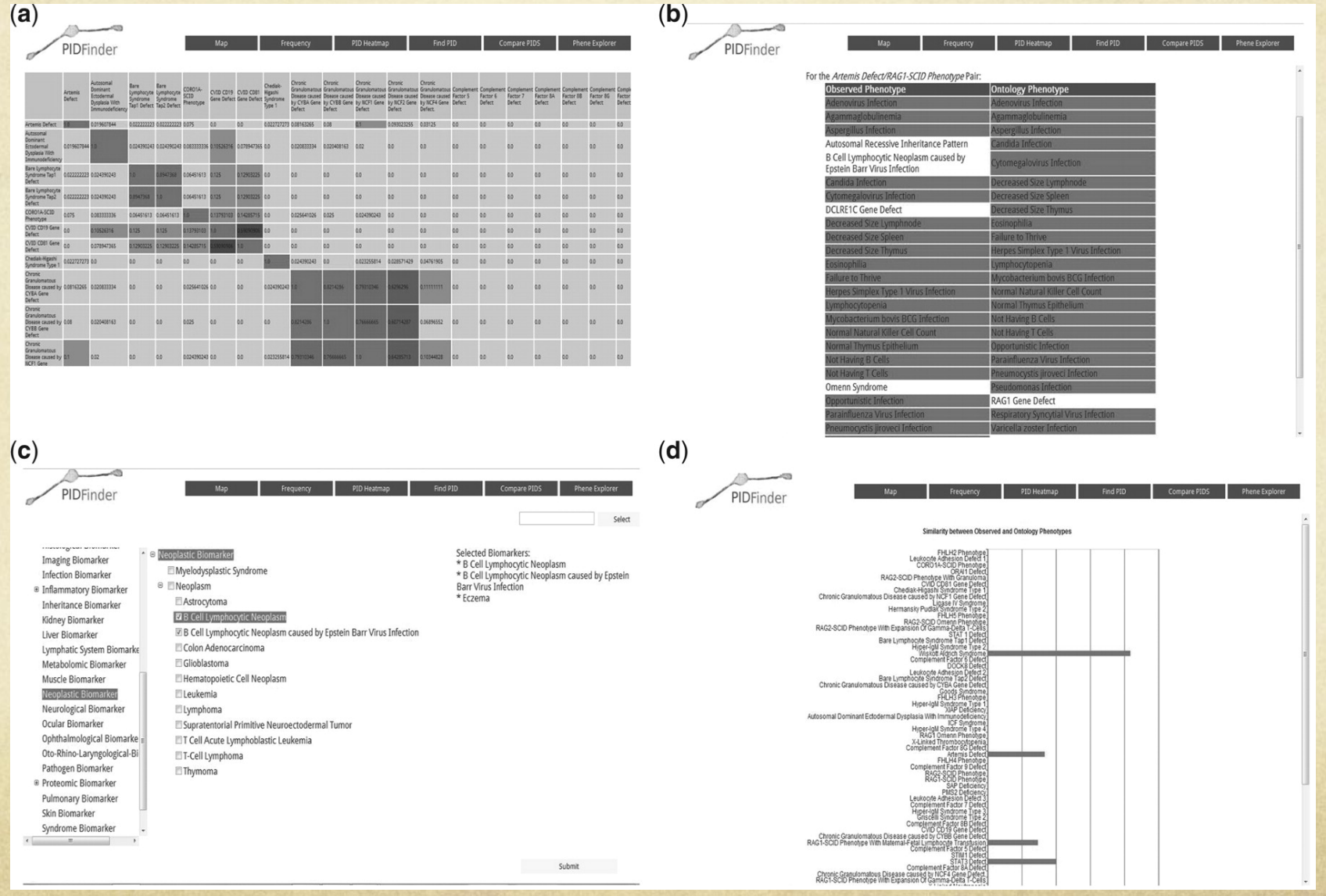
- Bacterial infection as an infection in which the role of the Infectious agent is played by a Bacterium

```
Bacterial_Infection EquivalentTo: (Infection and
(has_role some (Infectious_Entity and (played_by
some Bacterium))))
```

- Disease definition - Omenn Syndrome is commonly associated with primary immunodeficiencies arising due to defects of the RAG1 or RAG2 genes

```
Having_Omenn_Syndrome SubClassOf: Having_Alopecia
and Having_Hepatosplenomegaly and
Having_Lymphocytosis and (Having_Small_LymphNode or
Having_Large_LymphNode)
```

PIDFinder


(a)
PIDFinder
Map
Frequency
PID Heatmap
Find PID
Compare PIDS
Phene Explorer
(b)
PIDFinder
Map
Frequency
PID Heatmap
Find PID
Compare PIDS
Phene Explorer
For the Artemis Defect/RAG1-SCID Phenotype Pair:
Observed Phenotype
Ontology Phenotype
Adenovirus Infection
Adenovirus Infection
Agammaglobulinemia
Agammaglobulinemia
Aspergillus Infection
Aspergillus Infection
Autosomal Recessive Inheritance Pattern
Candida Infection
B Cell Lymphocytic Neoplasm caused by Epstein Barr Virus Infection
Cytomegalovirus Infection
Candida Infection
Decreased Size Lymphnode
Cytomegalovirus Infection
Decreased Size Spleen
DCLRE1C Gene Defect
Decreased Size Thymus
Decreased Size Lymphnode
Eosinophilia
Decreased Size Spleen
Failure to Thrive
Decreased Size Thymus
Herpes Simplex Type 1 Virus Infection
Eosinophilia
Lymphocytopenia
Failure to Thrive
Mycobacterium bovis BCG Infection
Herpes Simplex Type 1 Virus Infection
Normal Natural Killer Cell Count
Lymphocytopenia
Normal Thymus Epithelium
Mycobacterium bovis BCG Infection
Not Having B Cells
Normal Natural Killer Cell Count
Not Having T Cells
Normal Thymus Epithelium
Opportunistic Infection
Not Having B Cells
Parainfluenza Virus Infection
Not Having T Cells
RAG1 Gene Defect
Omenn Syndrome
Pseudomonas Infection
Opportunistic Infection
Respiratory Syncytial Virus Infection
Parainfluenza Virus Infection
Varicella zoster Infection
Pneumocystis jiroveci Infection
(c)
PIDFinder
Map
Frequency
PID Heatmap
Find PID
Compare PIDS
Phene Explorer
Select
Imaging Biomarker
Infection Biomarker
Inflammatory Biomarker
Inheritance Biomarker
Kidney Biomarker
Liver Biomarker
Lymphatic System Biomarke
Metabolomic Biomarker
Muscle Biomarker
Neoplastic Biomarker
Neurological Biomarker
Ocular Biomarker
Ophthalmological Biomarke
Oto-Rhino-Laryngological-Bi
Pathogen Biomarker
Proteomic Biomarker
Pulmonary Biomarker
Skin Biomarker
Syndrome Biomarker
Neoplastic Biomarker
Myelodysplastic Syndrome
Neoplasm
Astrocytoma
B Cell Lymphocytic Neoplasm
B Cell Lymphocytic Neoplasm caused by Epstein Barr Virus Infection
Colon Adenocarcinoma
Glioblastoma
Hematopoietic Cell Neoplasm
Leukemia
Lymphoma
Supratentorial Primitive Neuroectodermal Tumor
T Cell Acute Lymphoblastic Leukemia
T-Cell Lymphoma
Thymoma
Selected Biomarkers:
* B Cell Lymphocytic Neoplasm
* B Cell Lymphocytic Neoplasm caused by Epstein Barr Virus Infection
* Eczema
Submit
(d)
PIDFinder
Map
Frequency
PID Heatmap
Find PID
Compare PIDS
Phene Explorer
Similarity between Observed and Ontology Phenotypes
FHLH2 Phenotype
Leukocyte Adhesion Defect 1
CORO1A-SCID Phenotype
ORA1 Defect
RAG2-SCID Phenotype With Granuloma
CVID CD81 Gene Defect
Chediak-Higashi Syndrome Type 1
Chronic Granulomatous Disease caused by NCF1 Gene Defect
Ligase IV Syndrome
Hermansky Pudlak Syndrome Type 2
FHLH5 Phenotype
RAG2-SCID Omenn Phenotype
RAG2-SCID Phenotype With Expansion Of Gamma-Delta T-Cells
STAT 1 Defect
Bare Lymphocyte Syndrome Tap 1 Defect
Hiper-IgM Syndrome Type 2
Wiskott Aldrich Syndrome
Complement Factor 8 Defect
DOCK8 Defect
Leukocyte Adhesion Defect 2
Bare Lymphocyte Syndrome Tap2 Defect
Chronic Granulomatous Disease caused by CYBA Gene Defect
Coccle Syndrome
FHLH3 Phenotype
Hiper-IgM Syndrome Type 1
XIAP Deficiency
Autosomal Dominant Ectodermal Dysplasia With Immunodeficiency
ICF Syndrome
Hiper-IgM Syndrome Type 4
RAG1 Omenn Phenotype
X-Linked Thrombocytopenia
Complement Factor 8G Defect
Artemis Defect
FHLH4 Phenotype
Complement Factor 9 Defect
RAG2-SCID Phenotype
RAG1-SCID Phenotype
SAP Deficiency
PMS2 Deficiency
Leukocyte Adhesion Defect 3
Complement Factor 7 Defect
Hiper-IgM Syndrome Type 3
Griscelli Syndrome Type 2
Complement Factor 8B Defect
CVID CD19 Gene Defect
Chronic Granulomatous Disease caused by CYBB Gene Defect
RAG1-SCID Phenotype With Maternal-Fetal Lymphocyte Transfusion
Complement Factor 5 Defect
STIM1 Defect
STAT2 Defect
Complement Factor 8A Defect
Chronic Granulomatous Disease caused by NCF4 Gene Defect
RAG1-SCID Phenotype With Expansion Of Gamma-Delta T-Cells

Phenotype Ontologies

- Examples of species specific
  - Mammalian Phenotype Ontology (MP)
  - Human Phenotype Ontology (HPO)
  - Drosophila Phenotype Ontology
- Examples of domain specific
  - Cell Phenotype Ontology
  - Mammalian Pathology Ontology (MPATH)
  - Plant and Trait Phenotype Ontology (PO)
- Species and domain independent
  - PATO

Mammalian Phenotype Ontology (MP)

- The Mammalian Phenotype Ontology (MP) is build with the intention of classifying and organizing phenotypic information related to the mouse and other mammalian species.
- MP is used for the annotation of mouse phenotype descriptions in the Mouse Genome Informatics Database (MGI, http://www.informatics.jax.org/), the Rat Genome Database (RGD, http://rgd.mcw.edu), the Online Mendelian Inheritance in Animals (OMIA, http://omia.angis.org.au/)

Mammalian Phenotype Browser

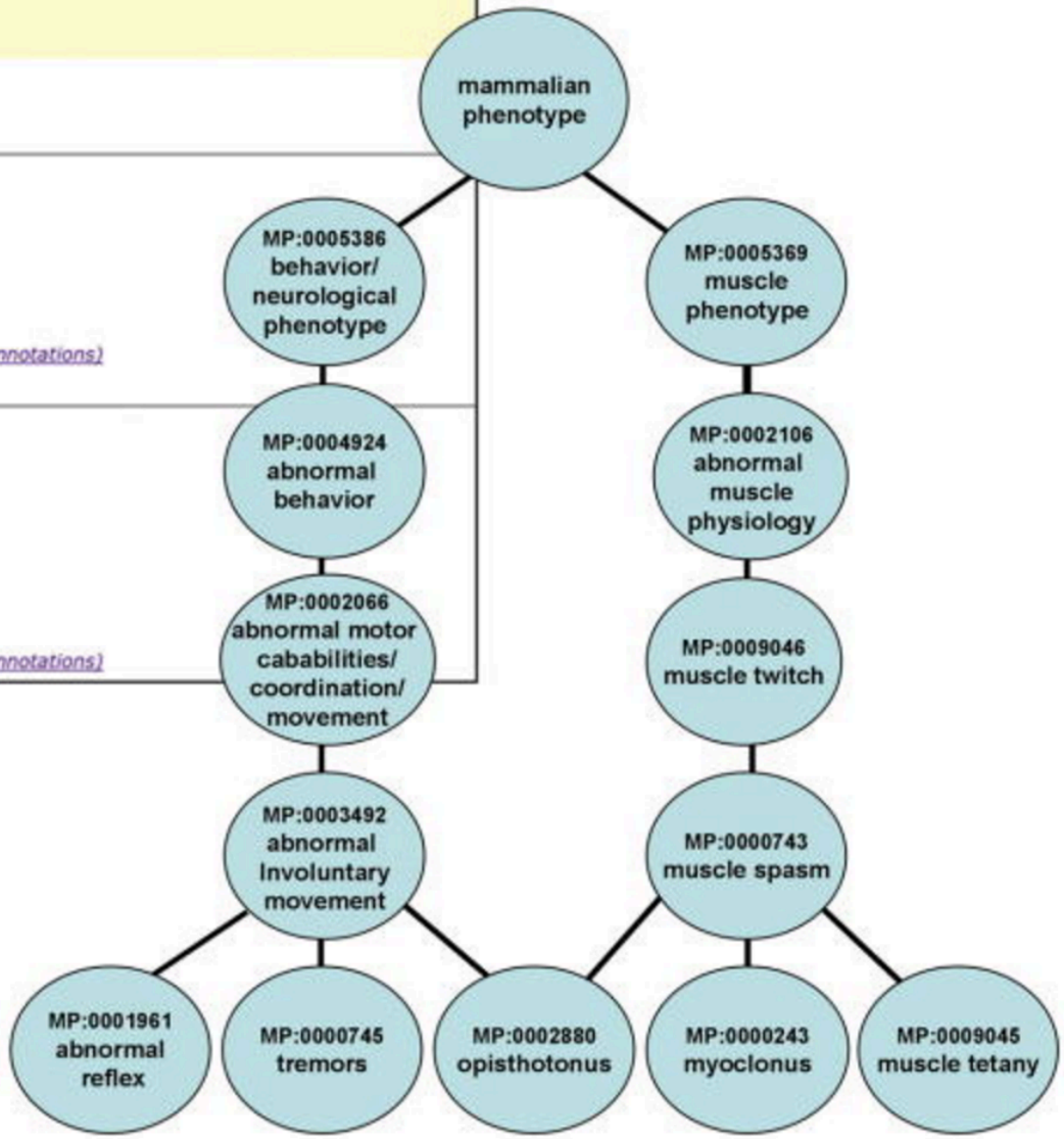
Term Detail

MP term:	**opisthotonus**
Synonym:	**opisthotonos**
MP id:	**MP:0002880**
Definition:	**a form of tetanic spasm in which the head, neck and spine are bent backward and the body is bowed forward**
Number of paths to term:	**2**

denotes an 'is-a' relationship
denotes a 'part-of' relationship

mammalian phenotype
- behavior/neurological phenotype
  - abnormal behavior
    - abnormal motor capabilities/coordination/movement
      - abnormal involuntary movement
        - abnormal reflex +
        - opisthotonus [MP:0002880] *(12 genotypes, 12 annotations)*
        - tremors

mammalian phenotype
- muscle phenotype
  - abnormal muscle physiology
    - muscle twitch
      - muscle spasm
        - muscle tetany
        - myoclonus
        - opisthotonus [MP:0002880] *(12 genotypes, 12 annotations)*



Eng^{tm1Mie} Targeted Allele Detail

Allele
Symbol: Eng^{tm1Mie}
Name: targeted mutation 1, Michelle Letarte
ID: MGI:2177624

Synonyms
End -, Eng-

Allele details
Allele Type: Targeted (Reporter)
Strain of Origin: 129P2/OlaHsd
ES Cell Line: E14
Mutation: Disruption caused by insertion of vector
Gene Expression in Eng^{tm1Mie} mutants (7 Assay Results)
International Mouse Strain Resource: (Search for IMSR strains with Eng mutations)
Mouse Models of Human Disease: (See Below)
References and Additional Notes: (See Below)

Allele Status: Approved
Submitted by: csmith
Approval Date: 2002-04-18 16:01:22.603

Gene information
Symbol: Eng | Human Ortholog: ENG
Name: endoglin
Chromosome: 2: 32502115-32538178 bp, + strand | Genetic Position: 21.4 cM

Phenotype summary
Phenotype Summary by Mammalian Phenotype terms
(show or hide all annotated terms)
Genotypes are listed in the next section.
Key:

Affected Systems	Genotypes:	hm1	ht2	ht3	ht4
cardiovascular system	▼	√	√	√	N
cellular	▶	√			
craniofacial	▶	√			
embryogenesis	▶	√			
growth/size	▶	√			
homeostasis/metabolism	▶	√			
lethality/prenatal-perinatal	▶	√			
life span/aging	▶				
respiratory system	▶	√			
skin/coat/nails	▶				
Disease Models					

Phenotypic data by genotype
Phenotypic Data by Genotype
(show or hide all phenotypic details)

Genotype	Allele
hm1	Eng
ht2 Disease Model	Eng
ht3 Disease Model	Eng
ht4	Eng

Disease models
Mouse Models of Human Disease
Models with phenotypic similarity to human diseases associated with human ENG.
Telangiectasia, Hereditary Hemorrhagic, of Rendu, Osler, and Weber;
OMIM ID: 187300

References
(Original) J:58502 Bourdeau A et al., "A murine mo

Affected Systems	Genotypes:	hm1	ht2	ht3	ht4
cardiovascular system	▼	√	√	√	N
cardiovascular system phenotype					N
abnormal heart development		√			
abnormal endocardium morphology		√			
enlarged heart		√			
pericardial edema		√			
abnormal vasculature		√			
absent heartbeat		√			
hemorrhage			√	√	
internal hemorrhage				√	
hemoperitoneum		√			
ear telangiectases			√	√	
tail telangiectases			√	√	
neck telangiectases				√	
cellular	▶	√			
craniofacial	▶	√			
embryogenesis	▶	√			
growth/size	▶	√		√	
homeostasis/metabolism	▶	√			
lethality/prenatal-perinatal	▶	√			
life span/aging	▶		√		
respiratory system	▶	√		√	
skin/coat/nails	▶			√	

Genotype: ▼ ht2 *Disease Model*
Allelic Composition: Eng^{tm1Mie}/Eng^{+}
Genetic Background: 129P2/OlaHsd-Eng^{tm1Mie}

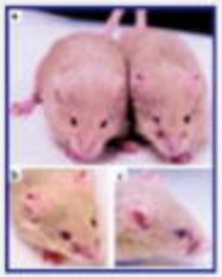
cardiovascular system
hemorrhage (J:58502)
- 3 of 10 have nosebleeds when observed for over a period of 7 months and some show tail bleeding

ear telangiectases (J:58502)
- 5 of 10 show visible telangiectases on the ears; exhibit large dilated vessels at the periphery of the ear lobe
- these small vascular lesions last for 2-3 days and then rapture and bleed, often leading to partial ear necrosis

tail telangiectases (J:58502)
- seen in some heterozygotes

life span/aging
premature death (J:58502)
- a female died at 7 weeks

Mouse Models of Human Disease	Note	Ref(s)
Models with phenotypic similarity to human diseases associated with human ENG.		
Telangiectasia, Hereditary Hemorrhagic, of Rendu, Osler, and Weber;		J:58502

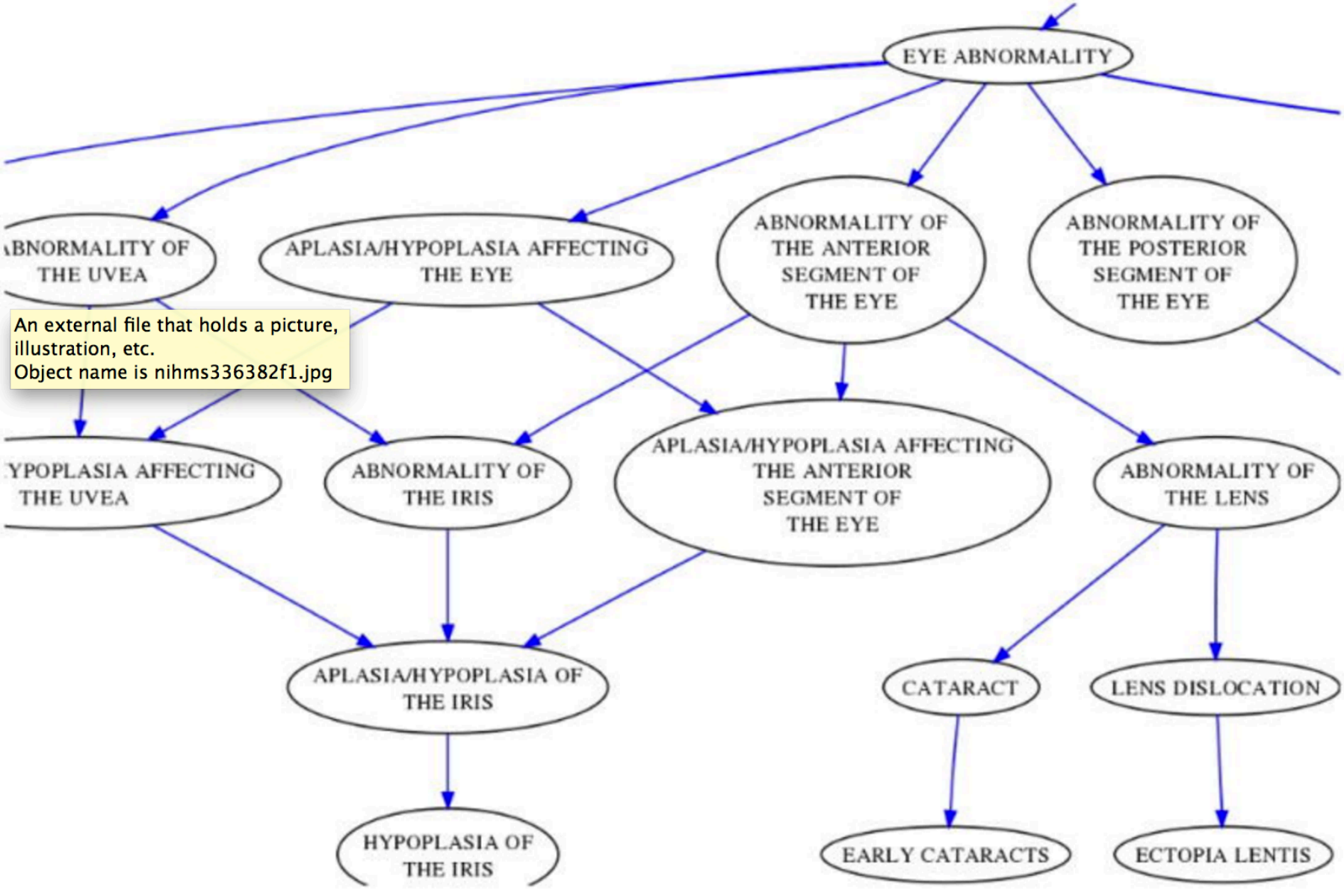


Eng^{tm1Mie}/Eng^{+} mice have nosebleeds and telangiectases characteristic of heriditary hemorrhagic telangiectasia

Human Phenotype Ontology (HPO)

- 10,088 classes describing human phenotypic abnormalities
- Used for annotations of 7,278 human hereditary syndromes listed in OMIM, Orphanet and DECIPHER to classes of the HPO.

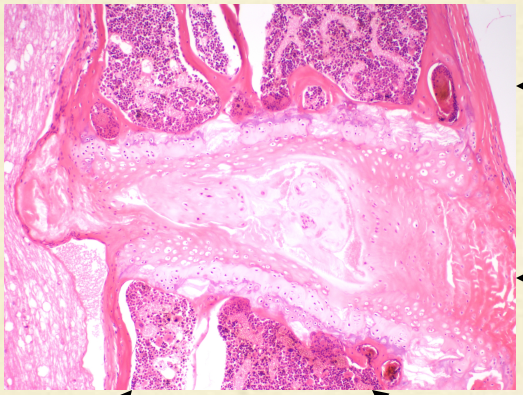
HPO structure



Mammalian Pathology Ontology (MPATH)

- Covers Pathological Process (General Pathology) and Specific Pathology
- Relatively high granularity but still reasonably shallow (9 deep)
- Linked to a database of pathology related images

Pathbase



Mutant strain name

Embryonic stage and tissue: EMAP

Common allele name

Free text

Gene affected

Pathology: MPATH

Literature reference

Location: MA

Genotype status

Cell type primarily affected:CL

GO terms for affected gene: GO

Class of mutation

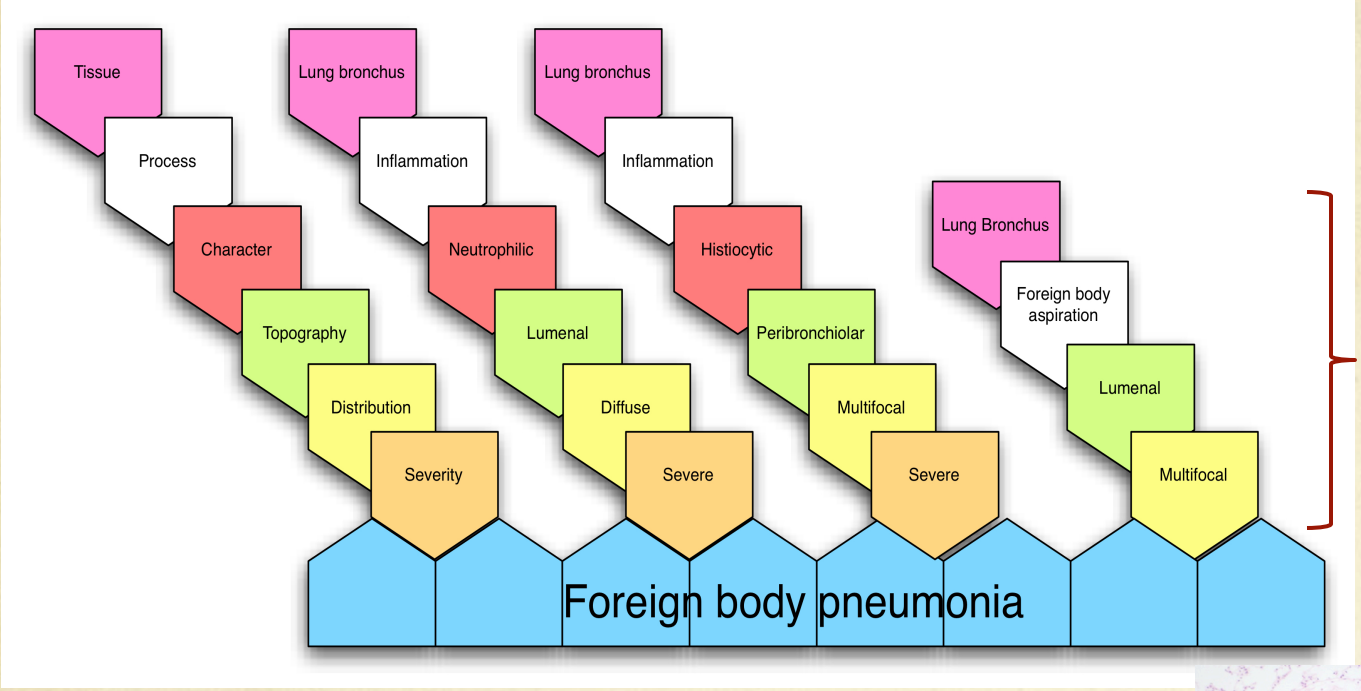
Adult age and stage

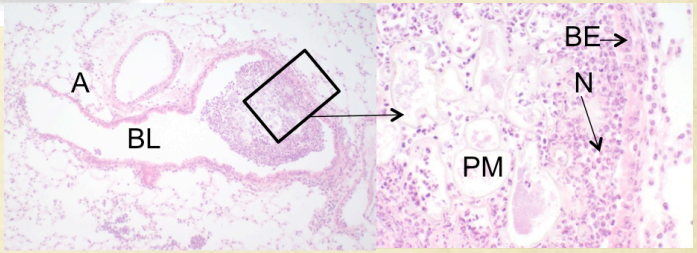
Constrained free text

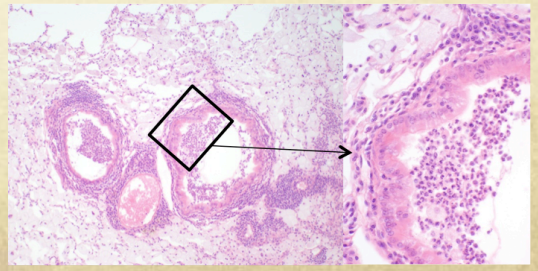
Simple CV

Ontology

Building up diagnostic components with anatomy, pathology and qualifiers







INHAND

Mann, P.C. et al (2012). International harmonization of toxicologic pathology nomenclature:
an overview and review of basic principles. *Toxicol Pathol*, **40,** 7S-13S.

Phenotype And Trait Ontology (PATO)

The meaningful cross species and across domain translation of phenotype is essential → phenotype-driven gene function discovery and comparative pathobiology

Goal - "A platform for facilitating mutual understanding and interoperability of phenotype information *across*

- **species,**
- **domains of knowledge,**

and amongst people and machines" .....

Phenotype Descriptions

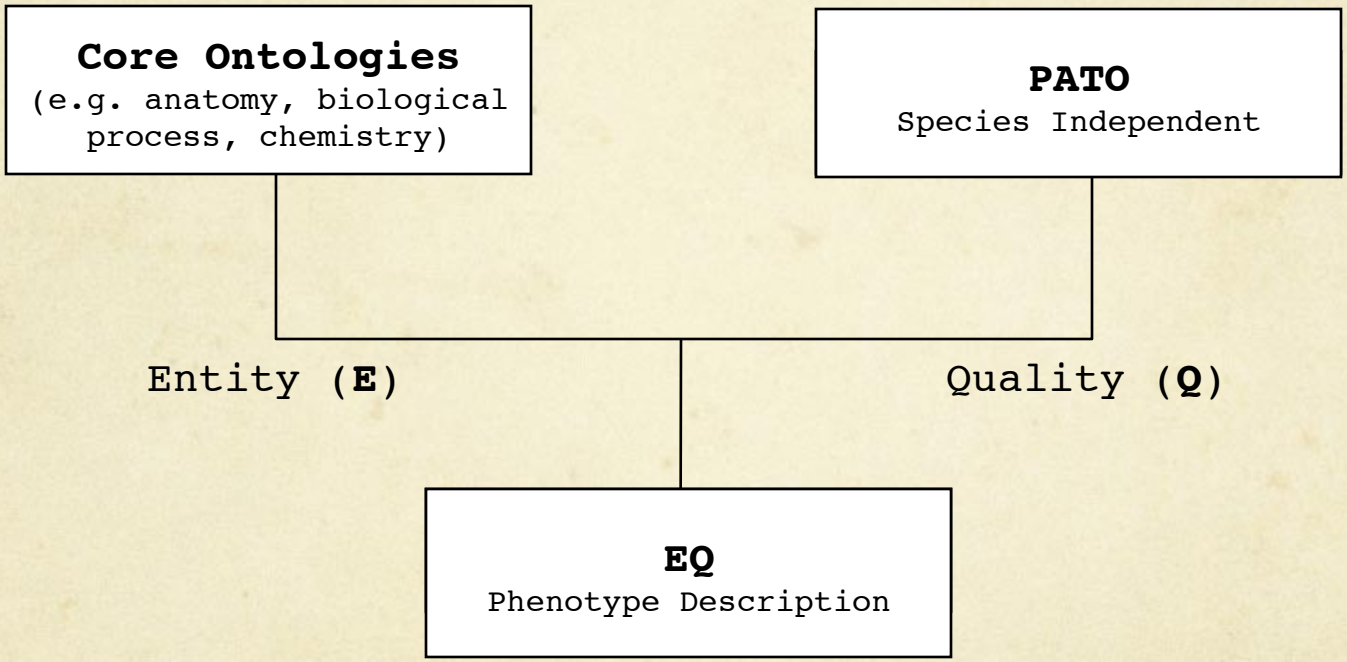
- phenotypes may be described in many different dimensions, e.g.
  - the biochemical ('alcohol dehydrogenase null')
  - the cellular ('cell division arrested at metaphase')
  - the anatomical ('eye absent')
  - the behavioral ('hyperactive')
  - etc.

- in whatever dimension and granularity, however, there is a commonality so that phenotypic descriptions can be decomposed into two parts

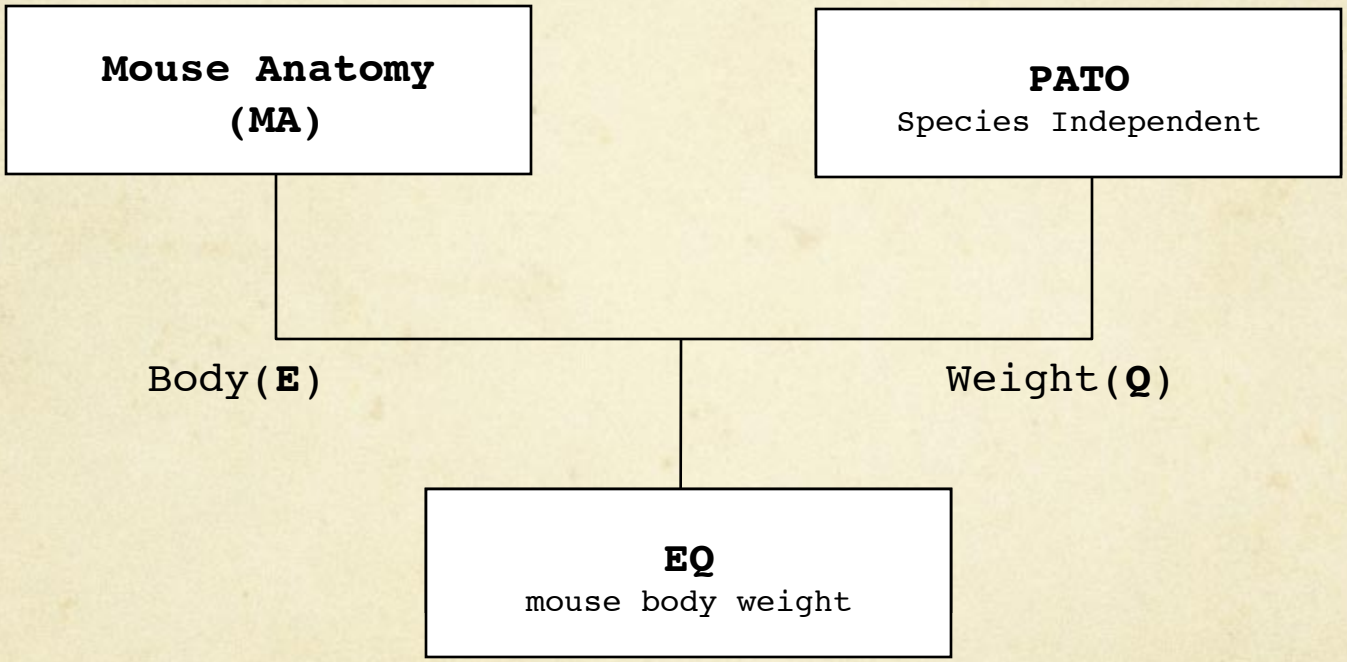
  - An **entity** that is affected. This entity may be an enzyme, an anatomical structure or a complex biological process.

  - The **qualities** of that entity.

EQ descriptors



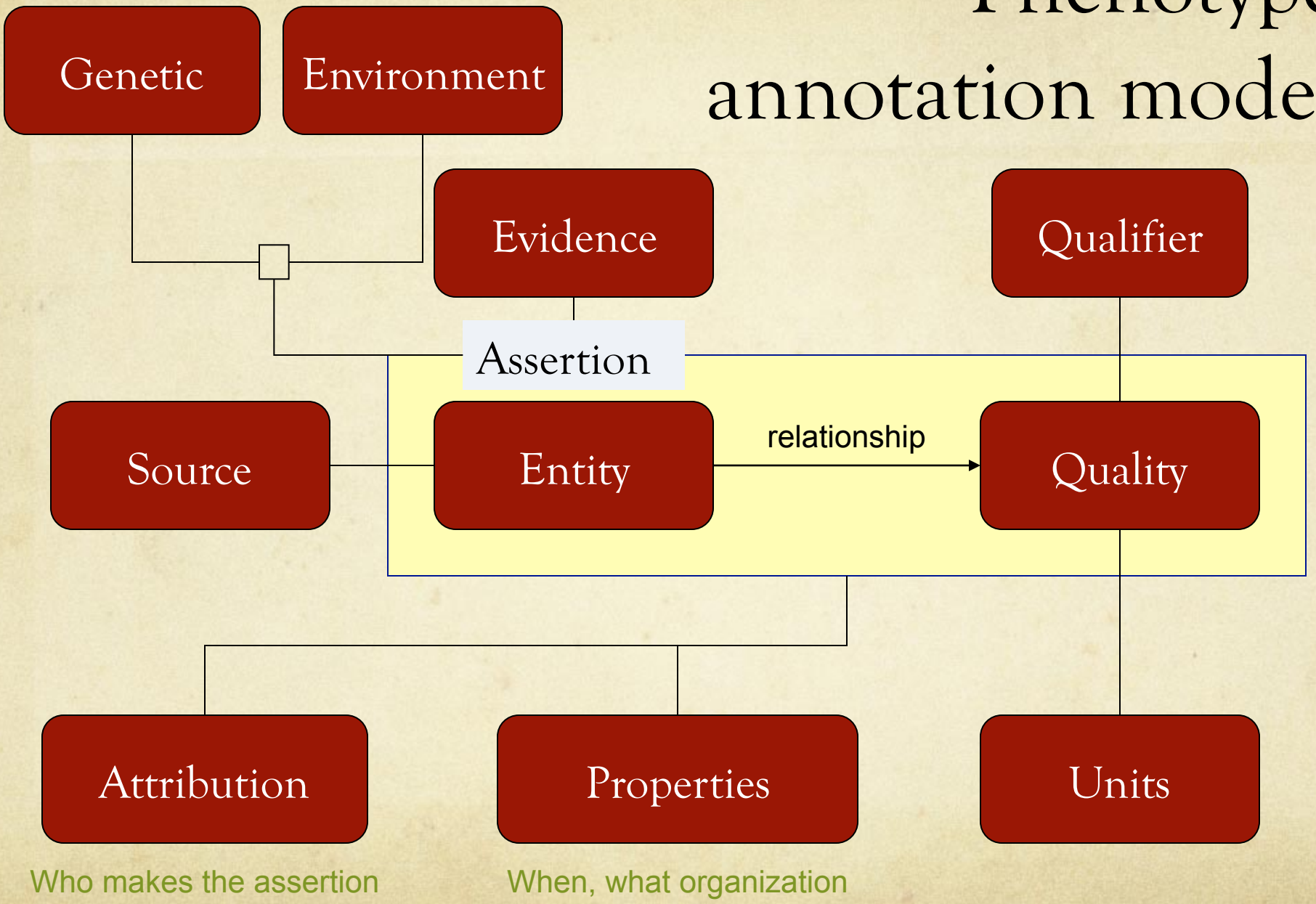
Mouse Body weight



PATO

- An ontology of phenotypic qualities, which can be shared across different species and domains of knowledge.
- Qualities are the basic entities that we can perceive and/or measure:
  - colors, sizes, masses, lengths etc.
- Qualities inhere to entities: every entity comes with certain qualities, which exist as long as the entity exist.
- Qualities belong in a finite set of quality types (i.e. color, size etc) and inhere in specific individuals. No two individuals can have the same quality, and each quality is specifically constantly dependent on the entity it inheres in.

Phenotype annotation model



In contrast, eye discs of *eya*[1] animals reveal a dramatic increase in cell death during the third instar larval stage. Cells are present in the *eya*[1] eye disc that appear condensed and refractile by light microscopy, reminiscent of

Annotation:
Phenotypes
in literature

Source:
PMID:8431945

Evidence:
light microscopy

Assertion

eya[1]

influences

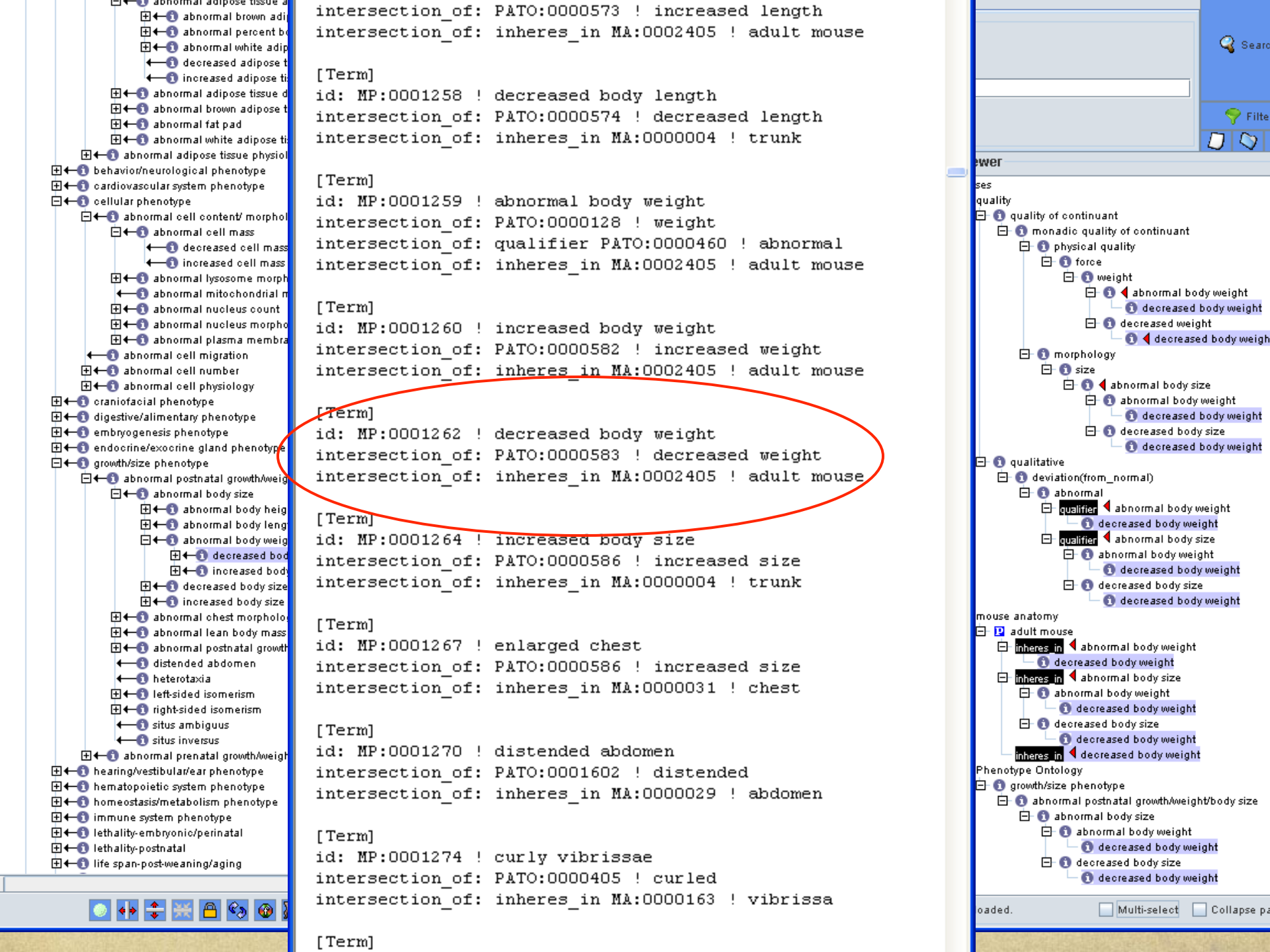
E=eye disc
(FBbt:00001768)

appears

Q=condensed
(PATO:0001485)

M. Ashburner

Date: 10/26/2007
Organization: FlyBase
Version: 1


intersection_of: PATO:0000573 ! increased length
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001258 ! decreased body length
intersection_of: PATO:0000574 ! decreased length
intersection_of: inheres_in MA:0000004 ! trunk
[Term]
id: MP:0001259 ! abnormal body weight
intersection_of: PATO:0000128 ! weight
intersection_of: qualifier PATO:0000460 ! abnormal
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001260 ! increased body weight
intersection_of: PATO:0000582 ! increased weight
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001262 ! decreased body weight
intersection_of: PATO:0000583 ! decreased weight
intersection_of: inheres_in MA:0002405 ! adult mouse
[Term]
id: MP:0001264 ! increased body size
intersection_of: PATO:0000586 ! increased size
intersection_of: inheres_in MA:0000004 ! trunk
[Term]
id: MP:0001267 ! enlarged chest
intersection_of: PATO:0000586 ! increased size
intersection_of: inheres_in MA:0000031 ! chest
[Term]
id: MP:0001270 ! distended abdomen
intersection_of: PATO:0001602 ! distended
intersection_of: inheres_in MA:0000029 ! abdomen
[Term]
id: MP:0001274 ! curly vibrissae
intersection_of: PATO:0000405 ! curled
intersection_of: inheres_in MA:0000163 ! vibrissa
[Term]
abnormal adipose tissue physiol
behavior/neurological phenotype
cardiovascular system phenotype
cellular phenotype
abnormal cell content/ morphol
abnormal cell mass
decreased cell mass
increased cell mass
abnormal lysosome morph
abnormal mitochondrial m
abnormal nucleus count
abnormal nucleus morpho
abnormal plasma membra
abnormal cell migration
abnormal cell number
abnormal cell physiology
craniofacial phenotype
digestive/alimentary phenotype
embryogenesis phenotype
endocrine/exocrine gland phenotype
growth/size phenotype
abnormal postnatal growth/weig
abnormal body size
abnormal body heig
abnormal body leng
abnormal body weig
decreased bod
increased body
decreased body size
increased body size
abnormal chest morpholo
abnormal lean body mass
abnormal postnatal growth
distended abdomen
heterotaxia
left-sided isomerism
right-sided isomerism
situs ambiguus
situs inversus
abnormal prenatal growth/weigh
hearing/vestibular/ear phenotype
hematopoietic system phenotype
homeostasis/metabolism phenotype
immune system phenotype
lethality-embryonic/perinatal
lethality-postnatal
life span-post-weaning/aging
Search
Filte
ewer
ses
quality
quality of continuant
monadic quality of continuant
physical quality
force
weight
abnormal body weight
decreased body weight
decreased weight
decreased body weigh
morphology
size
abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
qualitative
deviation(from_normal)
abnormal
qualifier abnormal body weight
decreased body weight
qualifier abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
mouse anatomy
P adult mouse
inheres_in abnormal body weight
decreased body weight
inheres_in abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
inheres_in decreased body weight
Phenotype Ontology
growth/size phenotype
abnormal postnatal growth/weight/body size
abnormal body size
abnormal body weight
decreased body weight
decreased body size
decreased body weight
oaded.
Multi-select
Collapse pa

MP - PATO based definitions

MP term	MP Definition	Entity	Quality
cataract MP:0001304	complete or partial opacity of the lens	lens MA:0000275 FMA:58241	opaque PATO:0000963
jaundice MP:0000611	clinical manifestation of hyperbilirubinemia, with deposition of bile pigments in the skin, resulting in yellowish staining of the skin and mucous membranes	skin MA:0000151 FMA:7163	yellow PATO:0000324
		skin mucous gland MA:0000148 mucous gland FMA:62888	yellow PATO:0000324
		pigment accumulation in tissues GO:0043480	yellow PATO:0000324
		pigment accumulation in tissues GO:0043480	mislocalized PATO:0000628

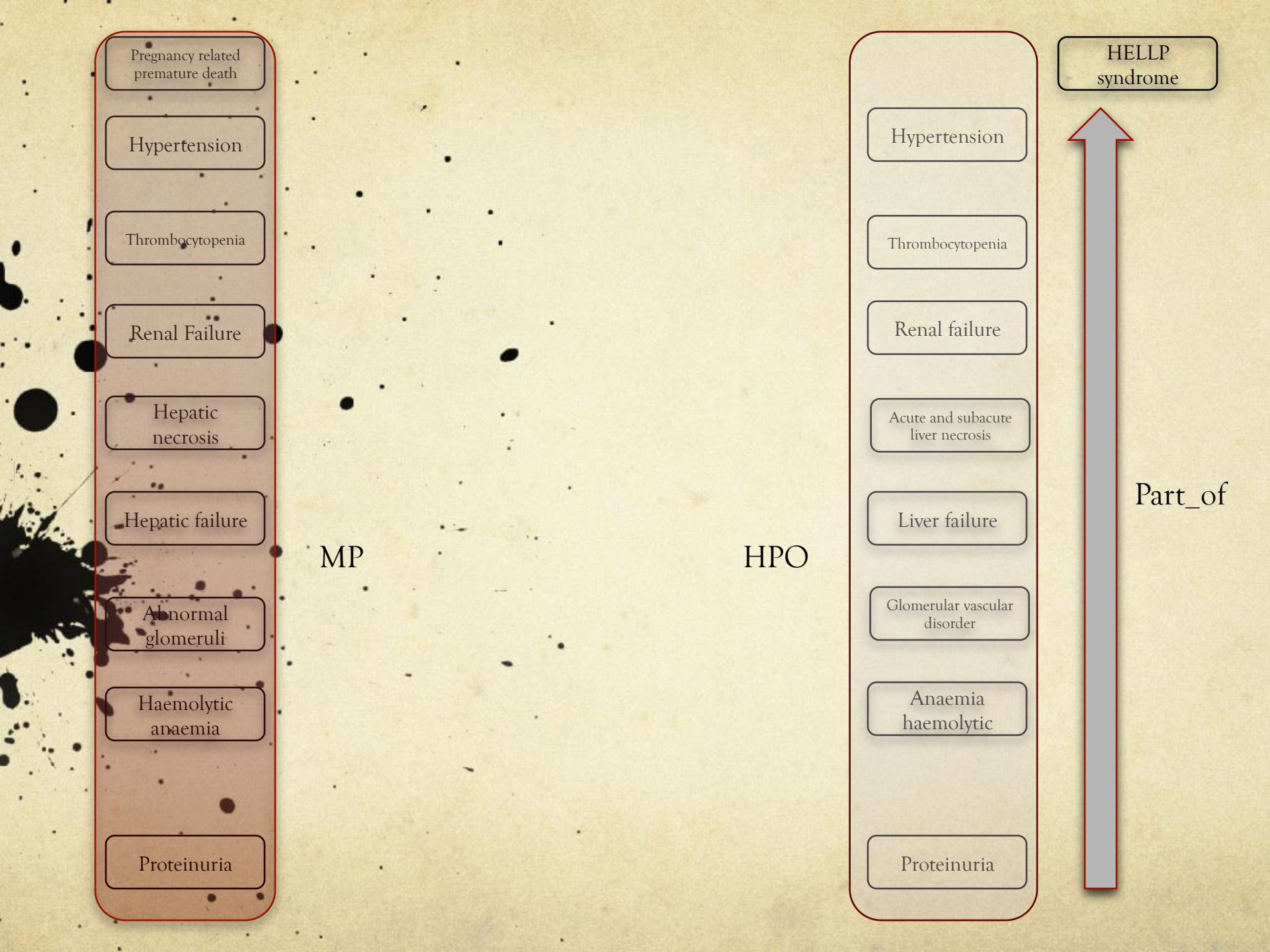
HPO-PATO based definitions

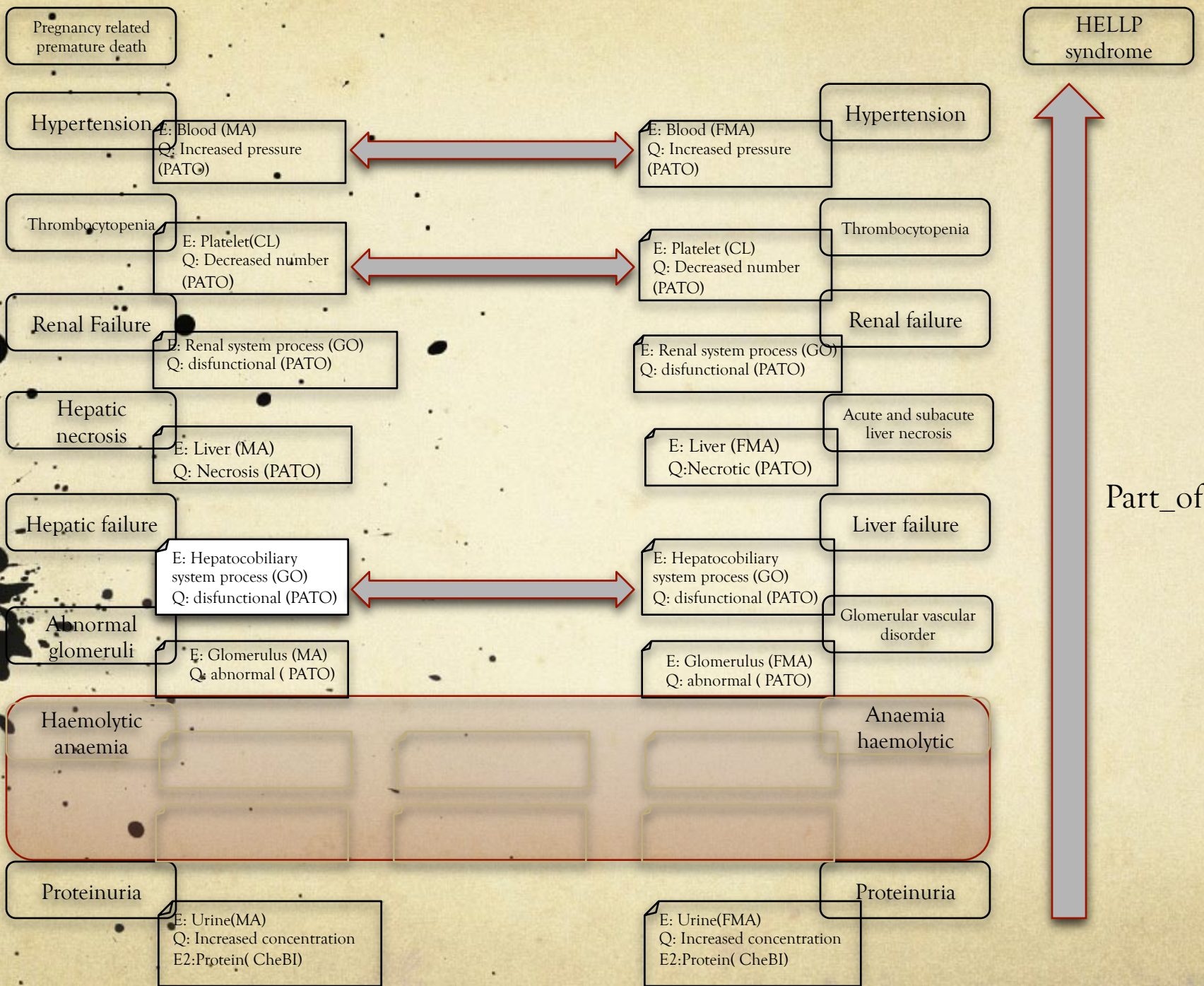
- OBO format

```
[Term]
id: HP:0004349 ! Reduced bone mineral density
intersection_of: PATO:0001790 ! decreased density
intersection_of: inheres_in FMA:30317 ! bone
```

- OWL format

```
Class: Hypoglycemia
EquivalentTo:
'decreased concentration'
and towards some 'glucose'
and inheres_in some 'portion of blood'
and qualifier some 'abnormal'
```


Pregnancy related premature death
Hypertension
Thrombocytopenia
Renal Failure
Hepatic necrosis
Hepatic failure
Abnormal glomeruli
Haemolytic anaemia
Proteinuria
MP
HPO
Hypertension
Thrombocytopenia
Renal failure
Acute and subacute liver necrosis
Liver failure
Glomerular vascular disorder
Anaemia haemolytic
Proteinuria
HELLP syndrome
Part_of


Pregnancy related premature death
HELLP syndrome
Hypertension
E: Blood (MA)
Q: Increased pressure (PATO)
Hypertension
E: Blood (FMA)
Q: Increased pressure (PATO)
Thrombocytopenia
E: Platelet(CL)
Q: Decreased number (PATO)
Thrombocytopenia
E: Platelet (CL)
Q: Decreased number (PATO)
Renal Failure
E: Renal system process (GO)
Q: disfunctional (PATO)
Renal failure
E: Renal system process (GO)
Q: disfunctional (PATO)
Hepatic necrosis
E: Liver (MA)
Q: Necrosis (PATO)
Acute and subacute liver necrosis
E: Liver (FMA)
Q:Necrotic (PATO)
Hepatic failure
E: Hepatocobiliary system process (GO)
Q: disfunctional (PATO)
Liver failure
E: Hepatocobiliary system process (GO)
Q: disfunctional (PATO)
Abnormal glomeruli
E: Glomerulus (MA)
Q: abnormal (PATO)
Glomerular vascular disorder
E: Glomerulus (FMA)
Q: abnormal (PATO)
Haemolytic anaemia
Anaemia haemolytic
Proteinuria
E: Urine(MA)
Q: Increased concentration
E2:Protein(CheBI)
Proteinuria
E: Urine(FMA)
Q: Increased concentration
E2:Protein(CheBI)
Part_of

Examples of Species Independent Ontologies

- CHeBI
- Cell Ontology
- Protein Ontology (PRO)
- NBO
- Unit Ontology (UO)

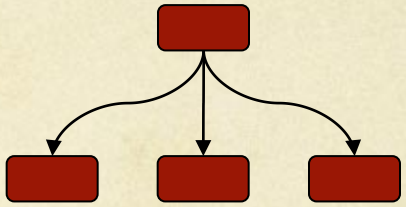
ChEBI

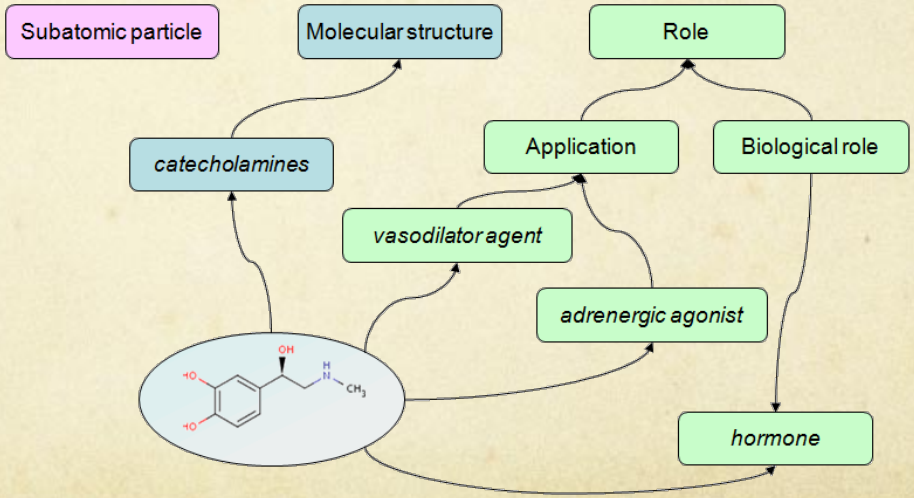
- Chemical Entities of Biological Interest
- Focused on ‘small’ chemical entities (no proteins or nucleic acids)
- Illustrated dictionary of chemical nomenclature
- High quality, manually annotated

ChEBI organisation

Organised into three sub-ontologies, namely

- Molecular structure ontology
- Subatomic particle ontology
- Role ontology





(*R*)-adrenaline

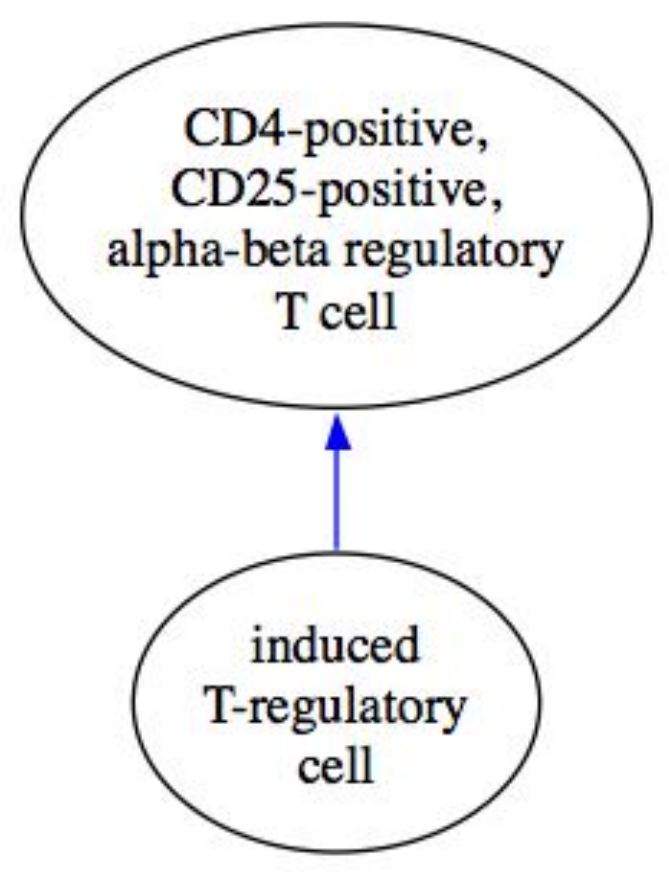
ChEBI terms

- A unique, unambiguous, recommended ChEBI name and an associated stable unique identifier
- An illustration where appropriate (compounds and groups, but generally not classes)
- A definition where appropriate (mostly classes)
- A collection of synonyms, including the IUPAC recommended name for the entity where appropriate
- A collection of cross-references to other databases

Cell Ontology (CL)

- An ontology of cell types built by biologists for the needs of data annotation and analysis.
- The Cell Ontology covers in vivo cell types from all of biology.
- The CL has over 1500 cell type terms, over 500 of which have logical definitions.

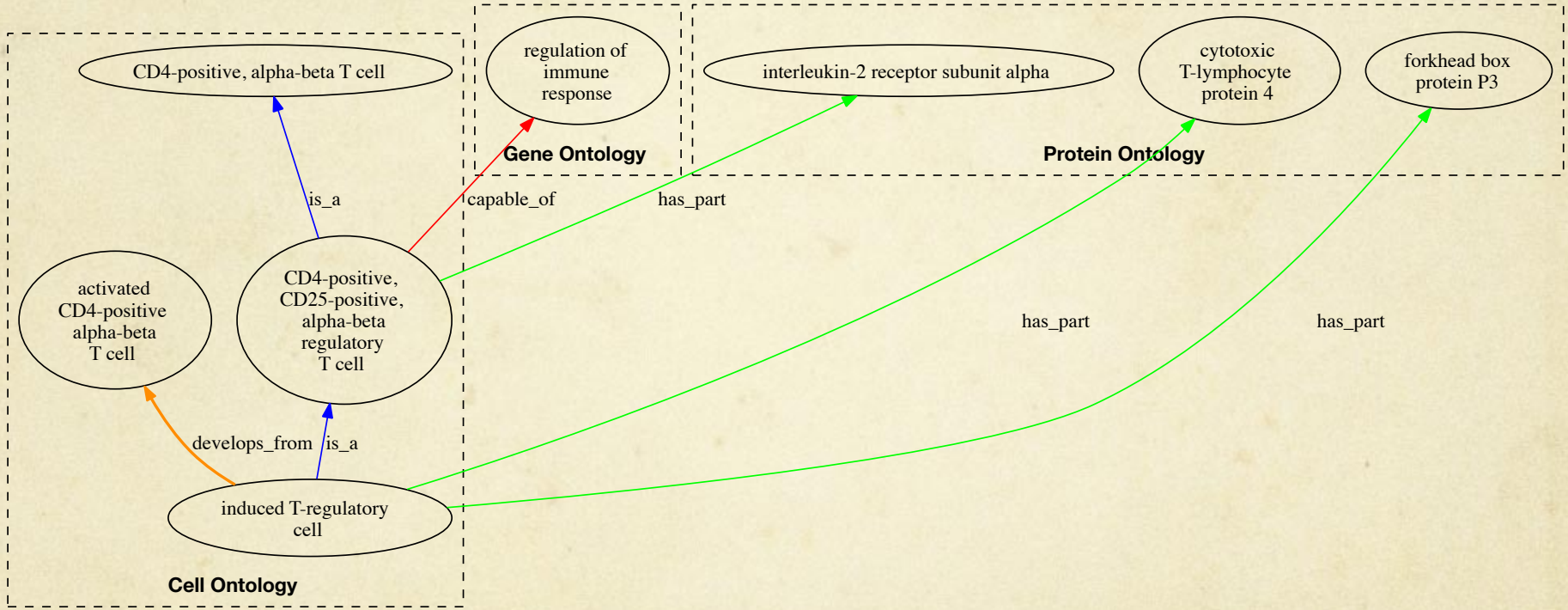
Textual Definitions



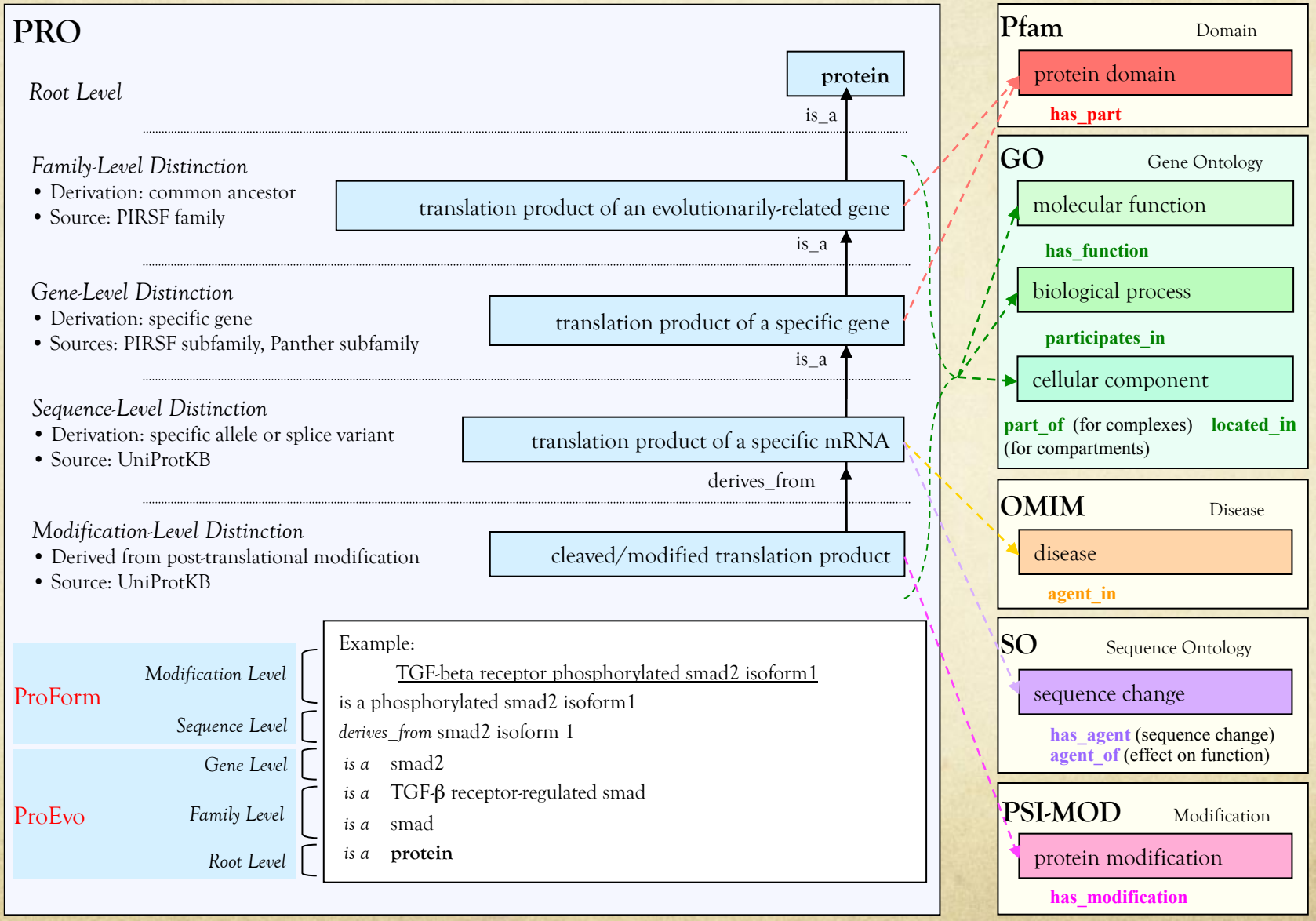
- name: CD4-positive, CD25-positive, alpha-beta regulatory T cell
- def: "A CD4-positive, CD25-positive, alpha-beta T cell that regulates overall immune responses as well as the responses of other T cell subsets through direct cell-cell contact and cytokine release."

- name: induced T-regulatory cell
- def: "CD4-positive alpha-beta T cell with the phenotype CD25-positive, CTLA-4-positive, and FoxP3-positive with regulatory function."

Logical Definitions



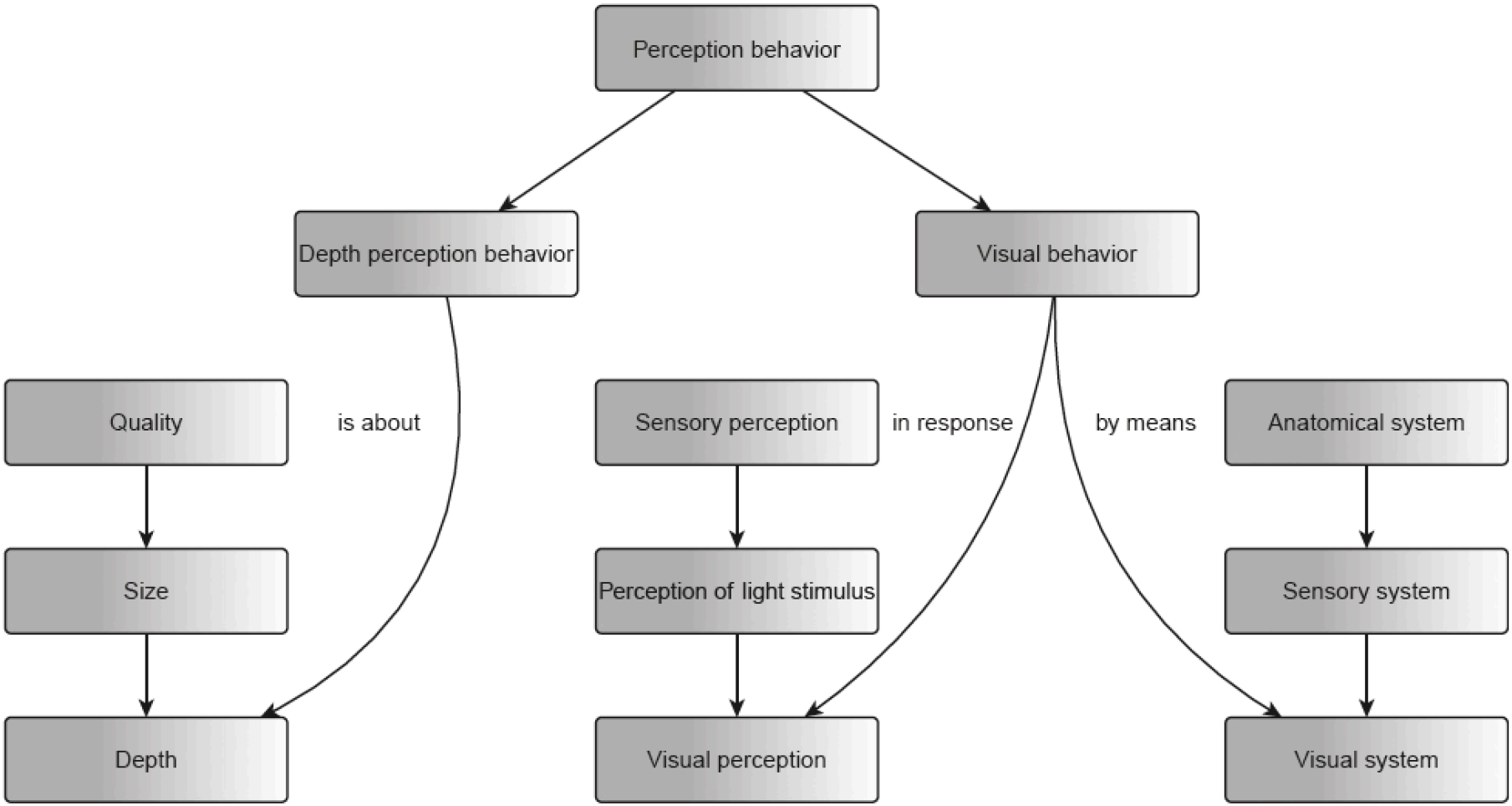
Protein Ontology (PRO)



NeuroBehavior Ontology (NBO)

- The neurobehavior ontology facilitates the systematic representation of behavior and behavioral phenotypes, thereby improving the unification and integration behavioral data in neuroscience research.
- Behavioral process ontology

NBO's axes of classification



Important NBO relations

Relation	Definition	Example
In-response-to	The relation **in-response-to** holds between a process *x* and a process *y* if and only if *x* occurs in response to *y*.	A *perception of visual stimulus* process occurs in response to the reception of light in the eye.
By-means-of	A process *x* occurs **by-means-of** a material structure *y* if and only if *x* occurs by means of *y*.	A perception of visual stimulus process occurs by means of the visual system.
Is-about	A process *x* **is-about** some entity *y* if and only if *x* is about or directed toward *y*.	A depth perception of process is about depth.

Complex phenotypes

- Increased amount of liquid in a single drinking act

'participates in' some ((has-input some ('liquid substance' and (has_quality some 'increased mass'))) and (regulates some 'drinking behavior'))

- Hyperdipsia

'participates in' some ((regulates some 'drinking behavior') and (has_quality some ('increased frequency' and (towards some 'drinking behavior') and (owl:qualifier some 'temporally extended'))))

- Dipsosis

'participates in' some ((regulates some 'drinking behavior') and (has_quality some ('increased frequency' and (towards some 'drinking behavior') and (owl:qualifier some 'chronic'))))

UO – an ontology of unit

- UO's top-level division is between primary base units of a particular measure and units that are derived from base units
- mapping between the various scalar qualities (such as weight, height, concentration etc.) and the corresponding units used to measure those qualities
- UO includes 264 terms, all of which are defined

Measurements

- Ontologies provide *qualitative* partitions on the kinds of entities we find in nature

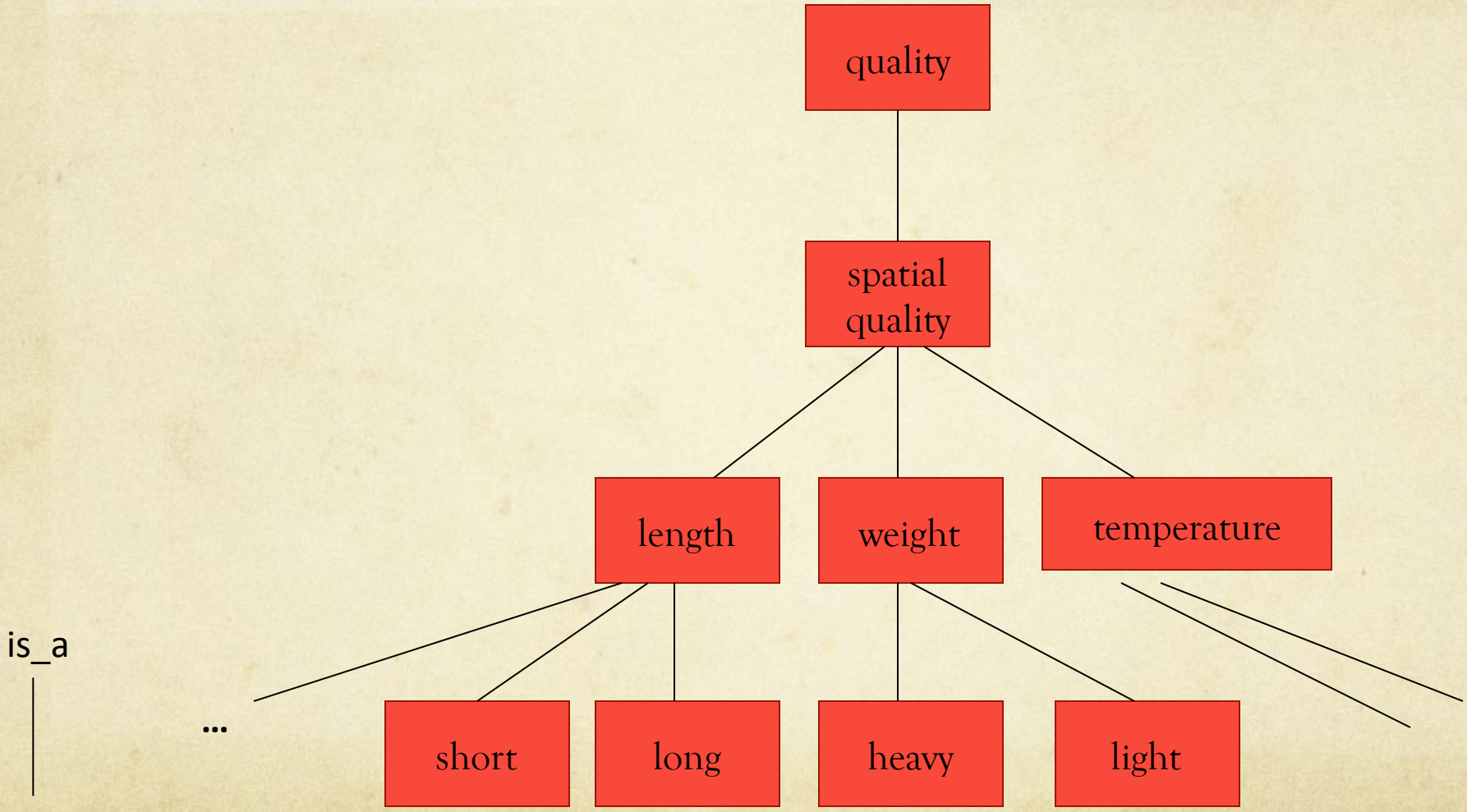
- We may also want to record *quantitative* information
  - Comes from *measurements* of qualities
  - The measurement is not the phenotype
    - Phenotypes exist independently of our measurements of them

Scalar qualities

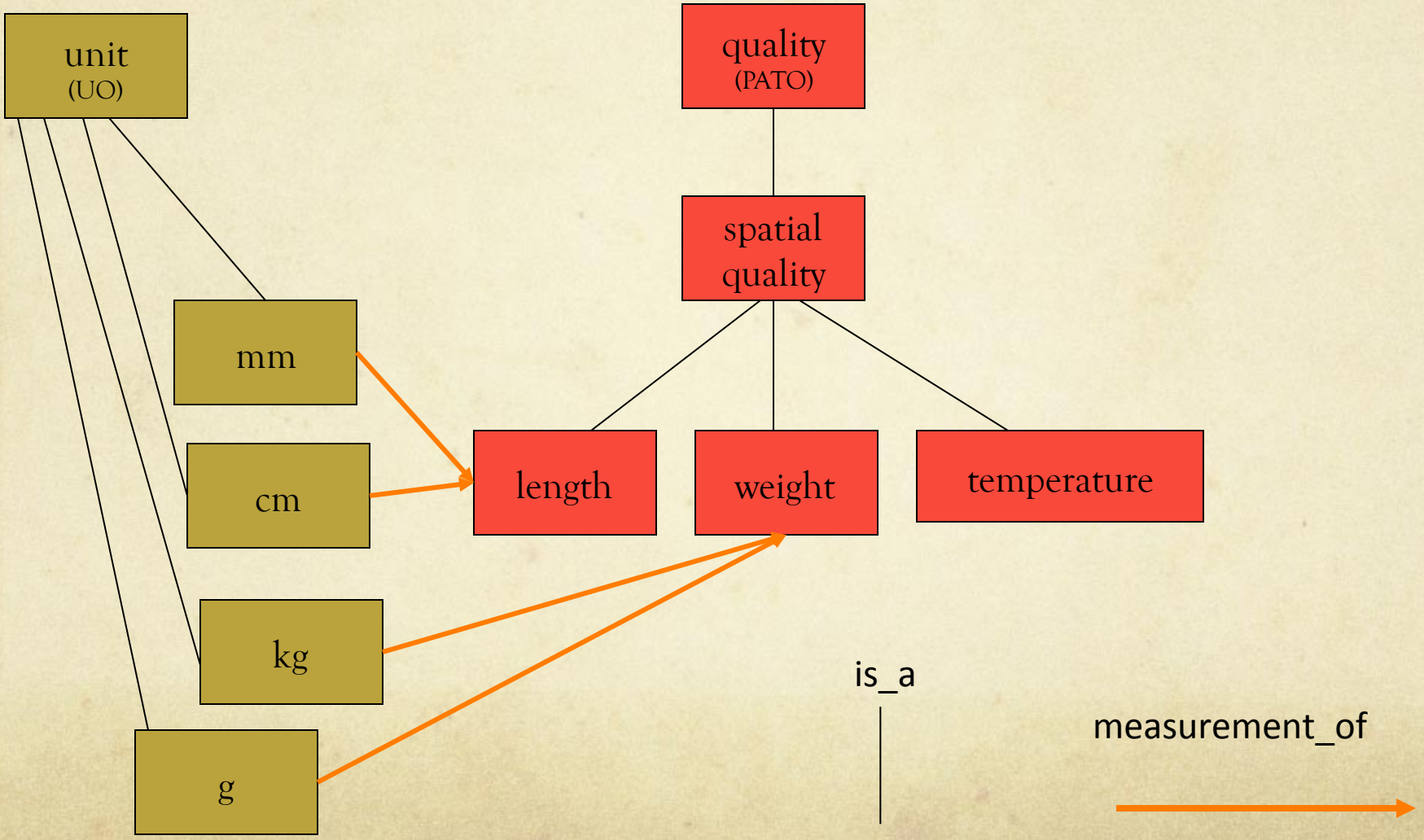
The tail of my mouse is 2.1 cm

- A scalar quality can be partitioned on a linear scale
- Scalar qualities can be measured
- Measurements involve units

PATO & scalar qualities



Representation of measurements



Mapping PATO to the UO