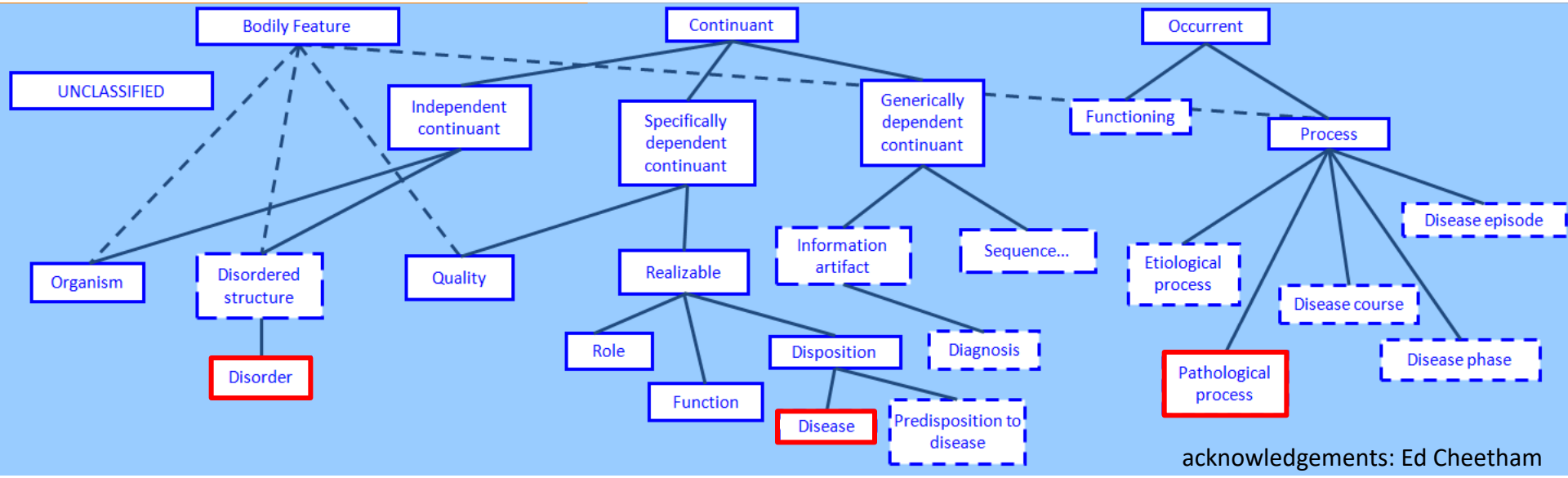
Scalable representations of diseases in biomedical ontologies and in SNOMED CT

Stefan Schulz
Institute of Medical Biometry und Medical Informatics University Medical Center Freiburg

Ontological Nature of Disease

- Hucklenbroich 2007: diseases are processes, events, or states
- Williams 2007: diseases are dispositional entities
- Scheuermann, Smith 2009: (i) diseases are dispositions, (ii) disorders are abnormal bodily components, and the (iii) manifestation of diseases are pathological processes
- SNOMED CT: Diseases under “Disorder”, “Finding”, “Event”, (rearrangement currently being discussed in the IHTSDO Event, condition, episode PG)

Diseases, disorders, pathological processes in disjoint BFO categories



90310002	Deficiency of saccadic eye movements (disorder)
194175003	Abnormal optokinetic response (finding)
370948005	Anterior capsule opacification (finding)
410568009	Anterior capsule opacification (disorder)
425558002	Azoospermia (disorder)
48188009	Azoospermia (finding)
89684003	Bends (disorder)
282977007	Does bend (finding)
399221001	Bleeding from vagina (disorder)
289530006	Bleeding from vagina (finding)
417237009	Blister of skin AND/OR mucosa (finding)
247464001	Blistering eruption (disorder)

35489007	Depressive disorder (disorder)
41006004	Depression (finding)
246815009	Excess skin of eyelid (finding)
58588007	Cutis laxa (disorder)
25702006	Alcohol intoxication (disorder)
86933000	Heavy drinker (finding)
46690002	Disorder of skin pigmentation (disorder)
3253007	Discoloration of skin (finding)
229694001	Oral dyskinesia (disorder)
9748009	Dyskinesia (finding)
242784006	Exposure to electric current, with passage of current through tissue (event)

Two Major Problems

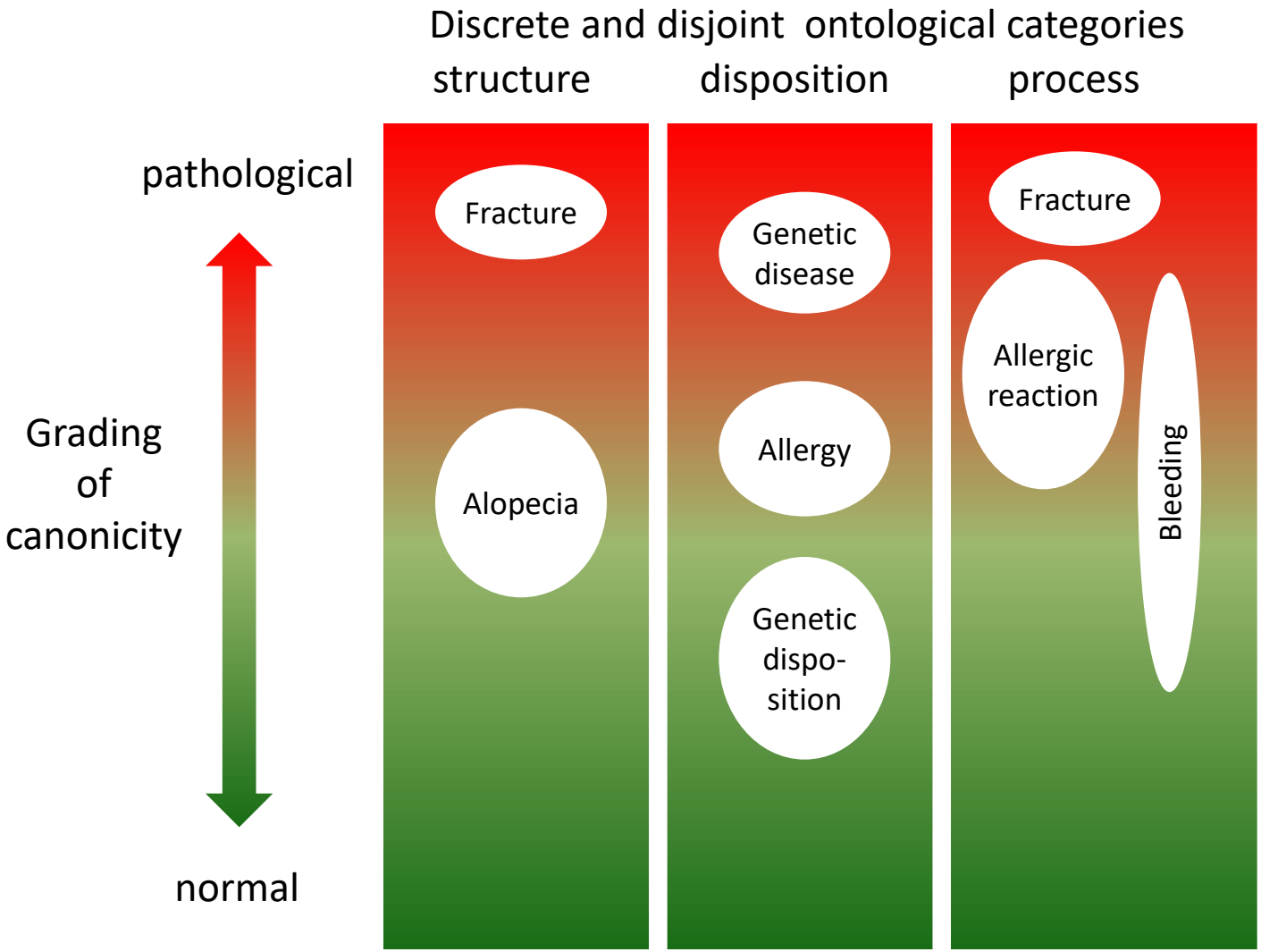
- Being pathological is rather a result of interpretation than a categorial property
  - Example: bleeding, pain, depression
- Ontologically motivated distinctions between disease, disorder, pathological process no not match the current meaning of words like “disease”, “disorder”, “abnormality” etc.

Use of “disorder” and “disease”

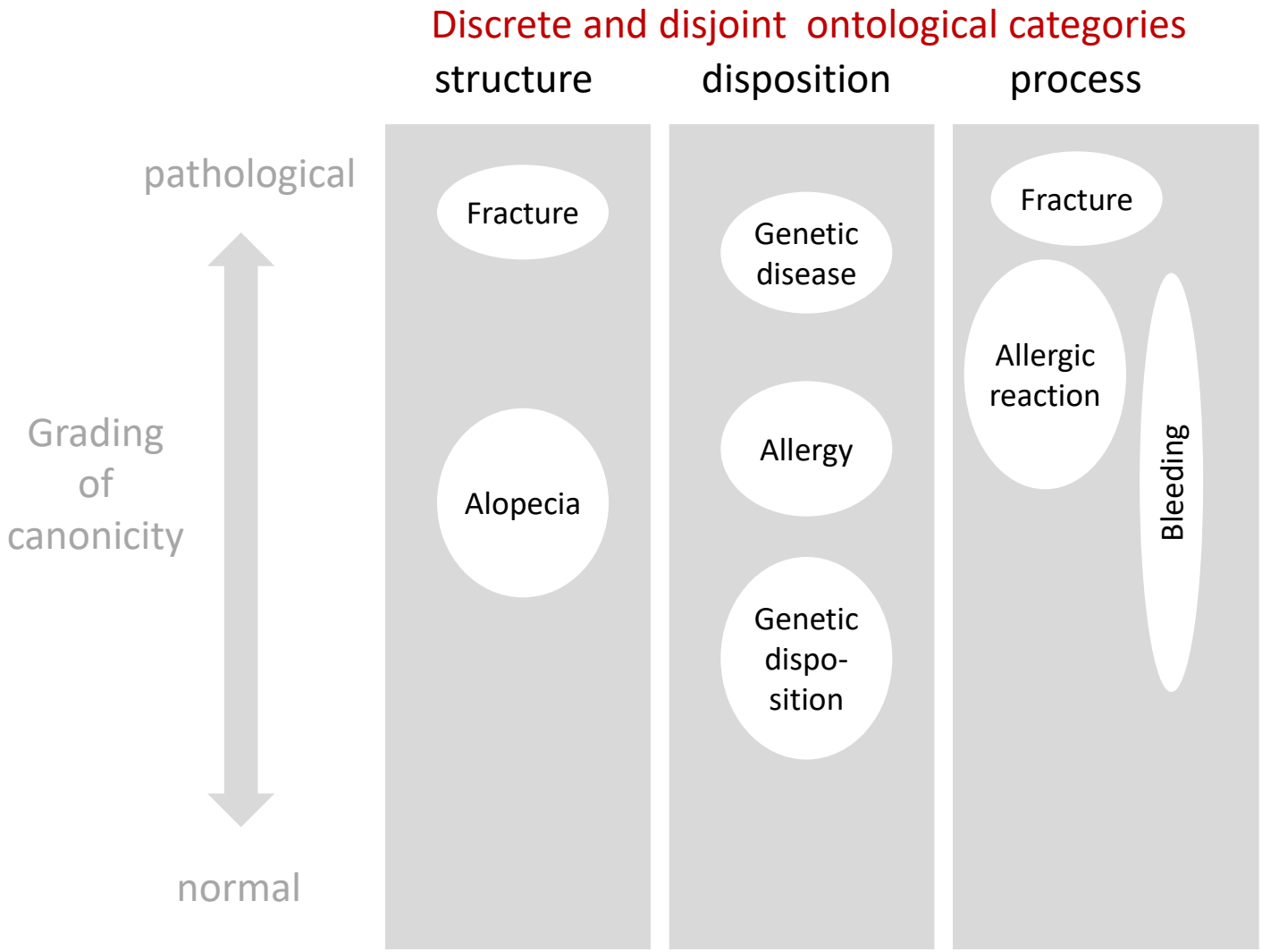
*** disease in MEDLINE**		*** disorder in MEDLINE**	
137880	heart disease	22360	bipolar disorder
77167	artery disease	20496	psychiatric disorders
66710	cardiovascular disease	14907	stress disorder
59307	liver disease	14458	depressive disorder
42607	renal disease	14115	anxiety disorders
34857	pulmonary disease	13977	mental disorders
29143	kidney disease	13935	personality disorder
27999	bowel disease	13600	panic disorder
27927	lung disease	13220	hyperactivity disorder
26376	vascular disease	11089	eating disorders

no support of the terminological suggestions by Scheuermann &Smith

Disease matrix



Disease matrix



Redefinition: avoiding ambiguous terms like disease, disorder

- ~~Disorder~~ **Pathological Structure**: a combination of bodily components of or in an organism
  1. that is not part of the life plan for an organism of the relevant type (thus aging or pregnancy are not clinically abnormal),
  2. that is causally linked to an elevated risk of pain or other feelings of illness or of death or dysfunction on the part of the organism, and
  3. that it is such that this elevated risk exceeds a certain threshold level.
- ~~Disease~~ **Pathological Disposition**: disposition
  1. to undergo pathological processes that
  2. exists in an organism because of one or more pathological structures in that organism.
- **Pathological Process**: bodily process that is a manifestation of a pathological disposition.

according to Scheuermann & Smith, 2009

Formalization of Scheuermann & Smiths definitions

PathologicalDisposition ⊑

∃ **inheresIn** .*PathologicalStructure*

PathologicalProcess ⊑

∃ **hasParticipant** .*PathologicalStructure*

PathologicalProcess ⊑

∃ **realizationOf**. *PathologicalDisposition* ?

PathologicalDisposition ⊑

∀**hasRealization**. *PathologicalProcess*

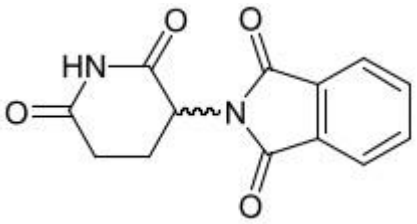
Example 1

- *Allergy* is a disposition of specific components of the immune system of an organism.
- All instances of the process type *Allergic Reaction*, are realizations of a disposition of this type, and have an allergen as their causative agent.

Image credit: http://www.topnews.in/health/files/Allergy.jpg

Example 2

- A specific binding of thalidomide to DNA forms a pathological structure on a molecular level
- This structure is the bearer of the pathological disposition realized by the misdevelopment of limbs (process) and results in a body without forearms (pathological structure)



Thalidomide

Image credit: http://www.mensch-home.com/Bilder/contergan-co-b/missbild-bild5g.jpg

Example 3

- The fracture (process) is caused by an external force, and has a fractured bone (pathological structure) as its characteristic outcome. This event is, however, not the realization of a disposition.
- A fractured bone (structure) has many pathological dispositions which can result in a variety of pathological processes (e.g. the development of a pseudarthrosis).

Image credit:
http://www.bcyr.ca/Survivor/Fracture[1].jpg

Ontological soundness vs. engineering requirements

- Ontology engineering: labor-intensive, use case-driven
- Not realistic to implement this model
  - in every well-founded ontology from the very beginning
  - for all pathological entities to be represented
- Challenge: let a coarse-grained, pragmatic representation (which ignores the structure / disposition / process distinction) gracefully evolve towards a more fully-fledged ontology?
- Can this be done in a intuitive, user-friendly, ontologically sound, computable, and scalable way?

Disjunctive top level category

- *PathologicalEntity* ≡

  PathologicalStructure ⊔

  PathologicalDisposition ⊔

  PathologicalProcess

- Top node of disease / disorder hierarchy

  (regardles of whether a distinction is made between processes, structures, dispositions)

Relation to organism parts / locations

… crucial for defining pathological entities

Different relations (e.g. OBO RO, BioTop)

- *Pathological Structures*: **part-of** / **located-in**
- *Pathological Dispositions*: **inheres- in**
- *Pathological Processes*: **has-participant**

  located-in

Redesign of relation hierarchy

… allows connection to organism parts or locations, without commitment to structure, disposition, or process

part-of $\sqsubseteq$ **has-locus**

has-location $\sqsubseteq$ **has-locus**

inheres-in $\sqsubseteq$ **has-locus**

has-participant $\sqsubseteq$ **locus-of**

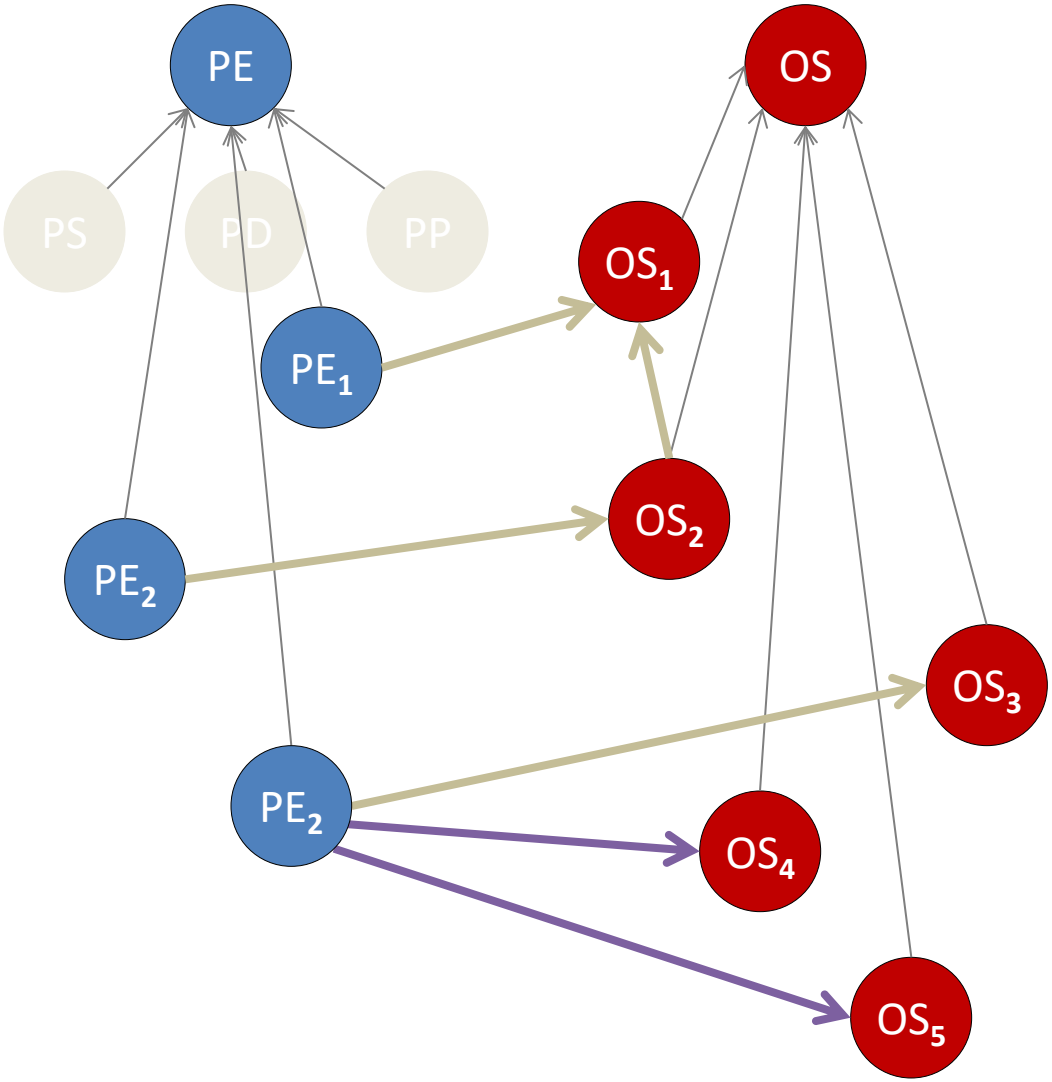
locus-of $\equiv$ **has-locus**$^{-1}$ **:** reflexive and transitive …

Corollaries of relation abstraction

- a disposition of a part is also borne by the whole 
- a pathological structure located in a part is also located in the whole 
- a process located in a part is also located in the whole 
- all participants of a process are located where the process is located 

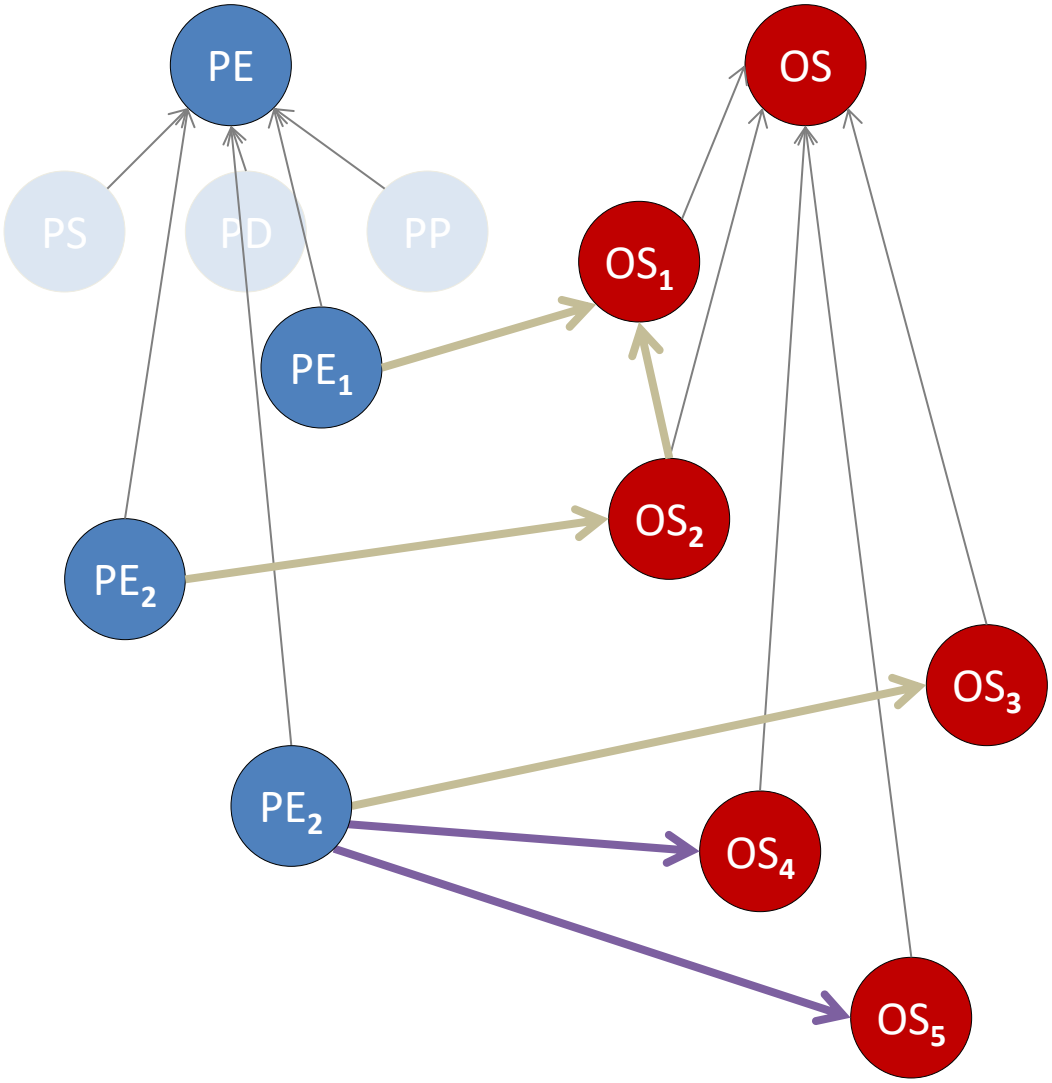
Construction of basic disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**



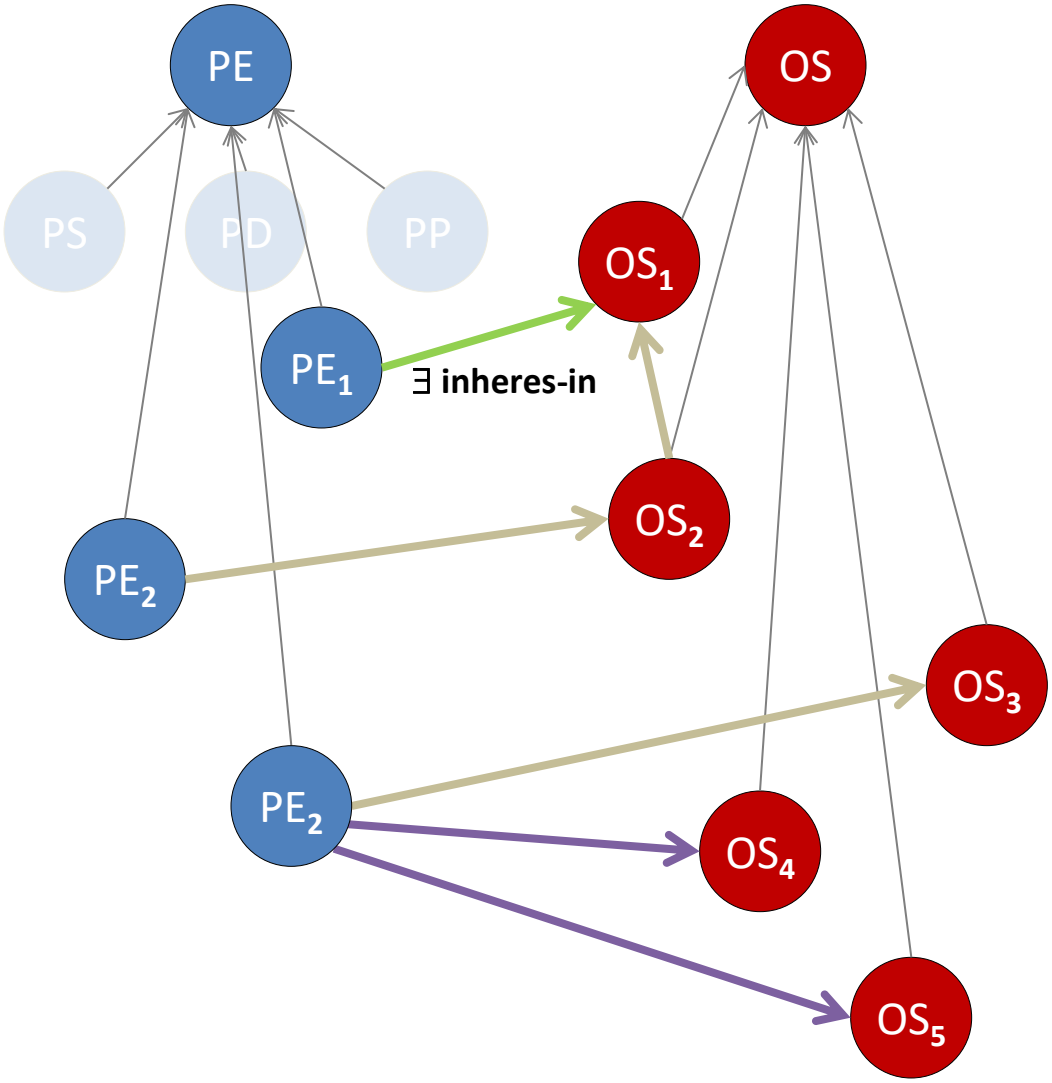
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PS *PathologicalStructure*
  - PD *PathologicalDisposition*
  - PP *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



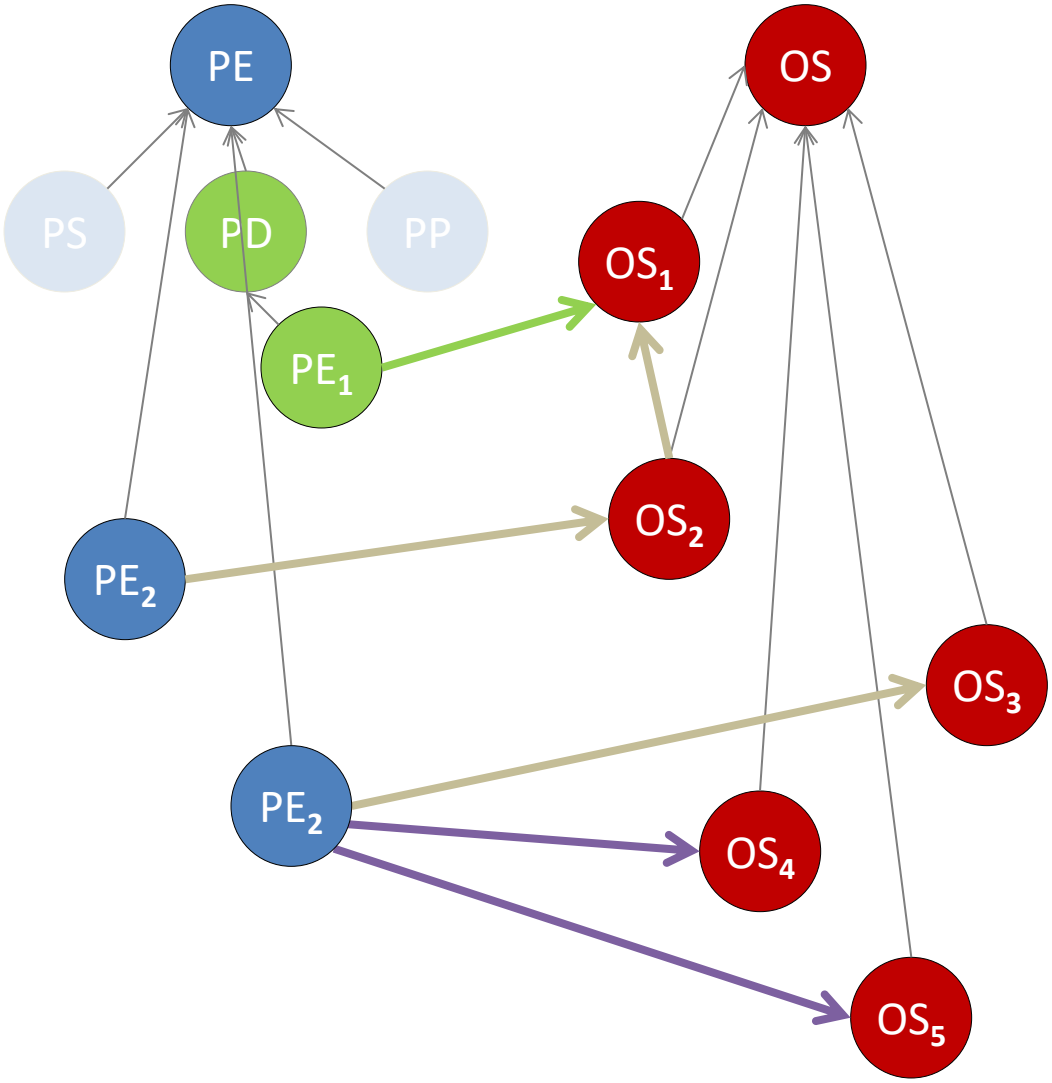
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PS *PathologicalStructure*
  - PD *PathologicalDisposition*
  - PP *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



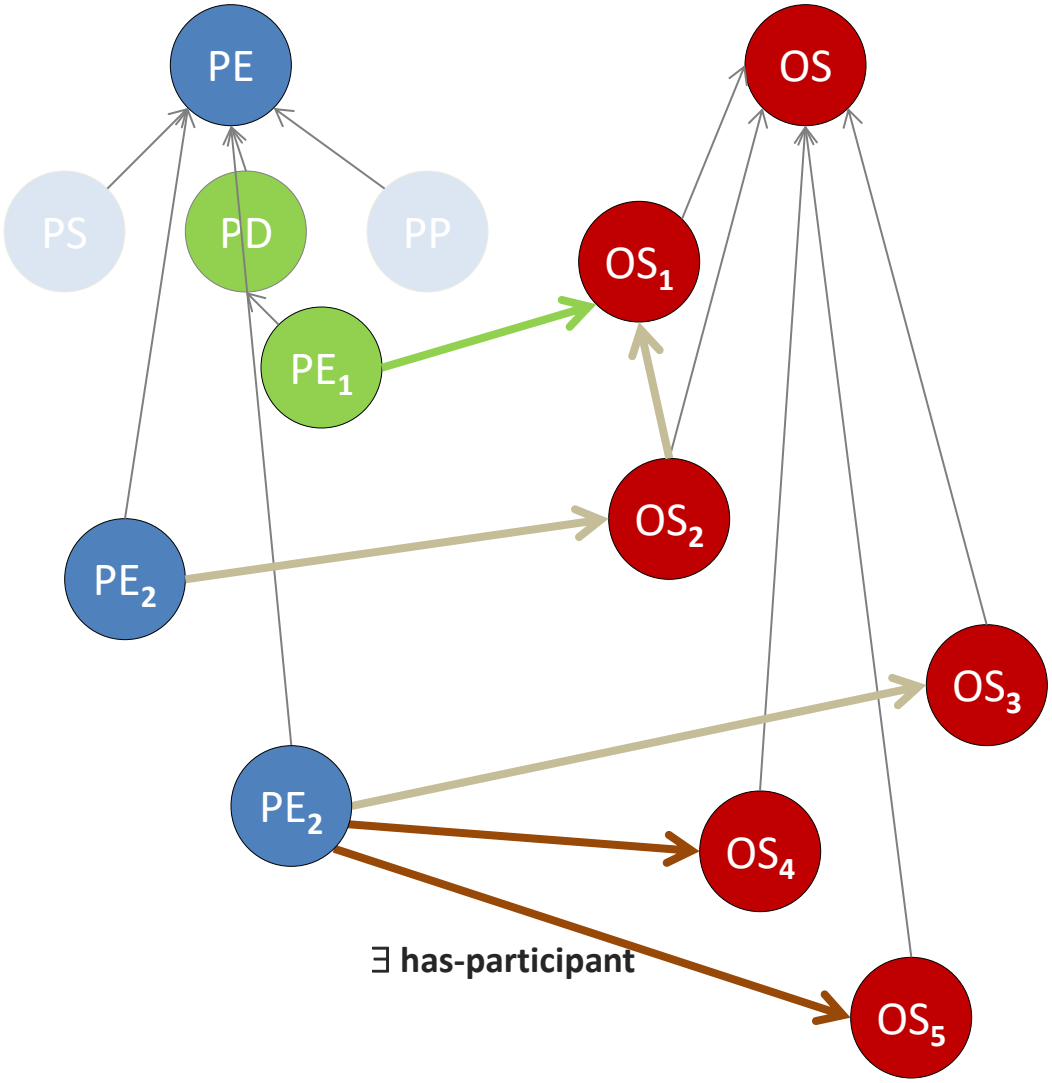
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



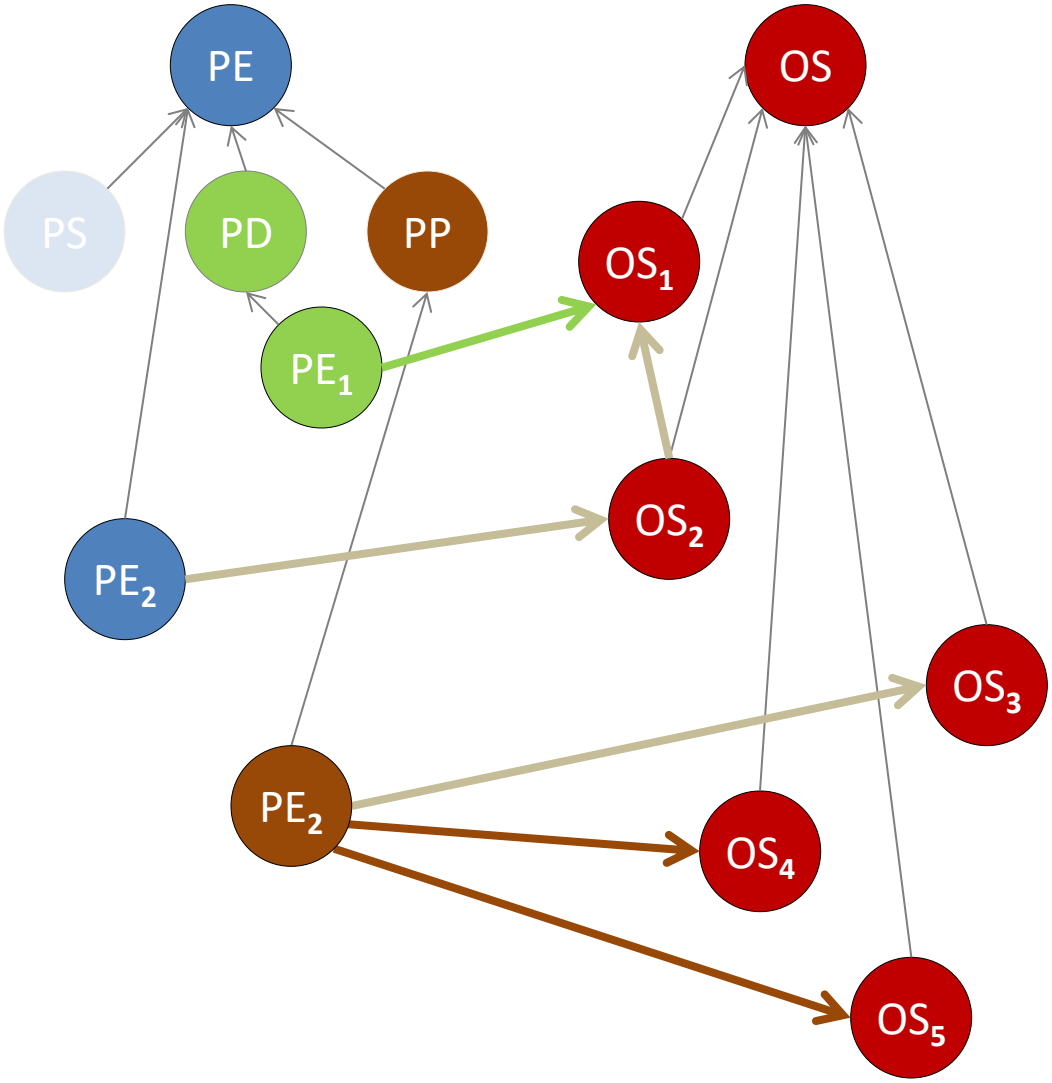
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



Construction of advanced disease ontology

- Basic components:
  - Top nodes

PE *Pathological Entity*

OS *Organism Structure*

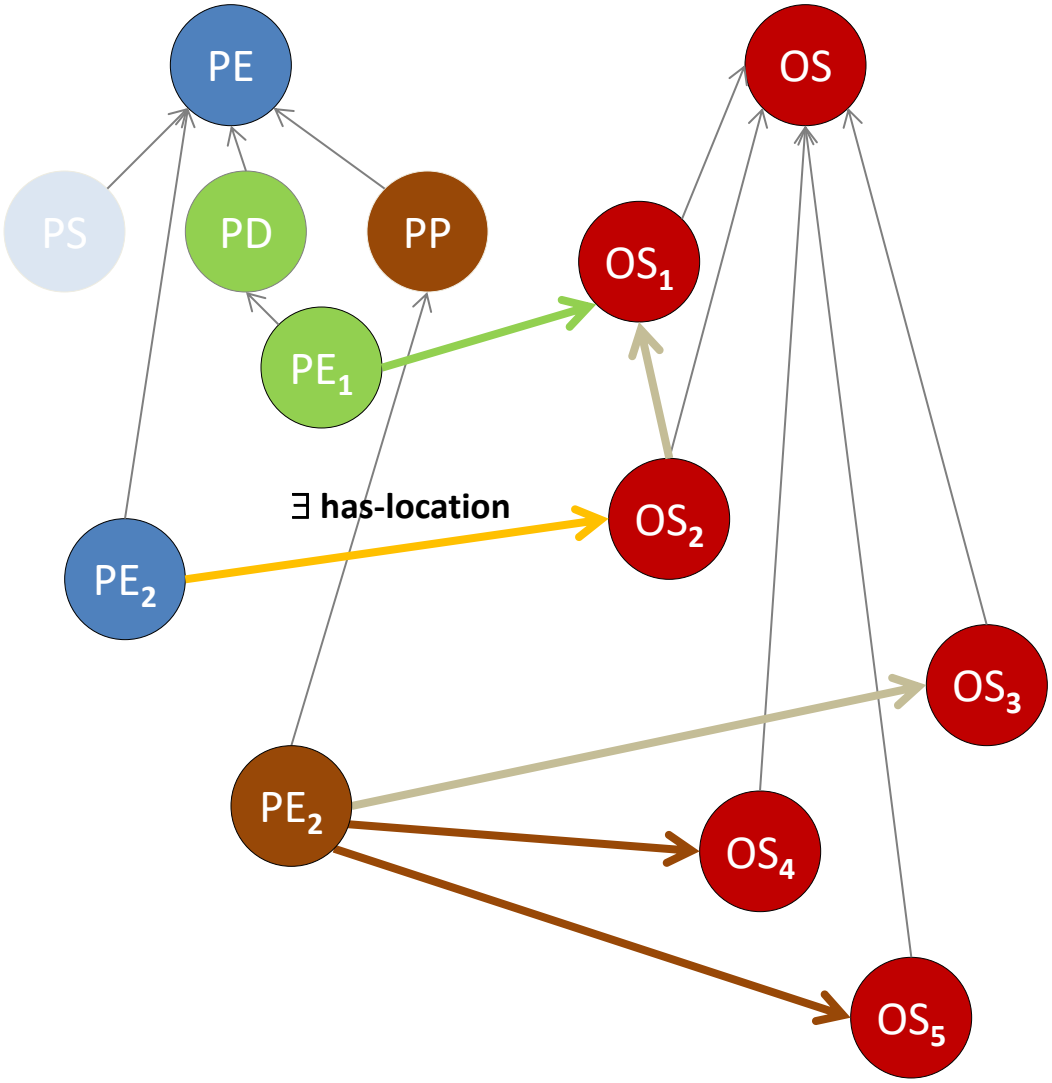
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components

PE *PathologicalStructure*

OS *PathologicalDisposition*

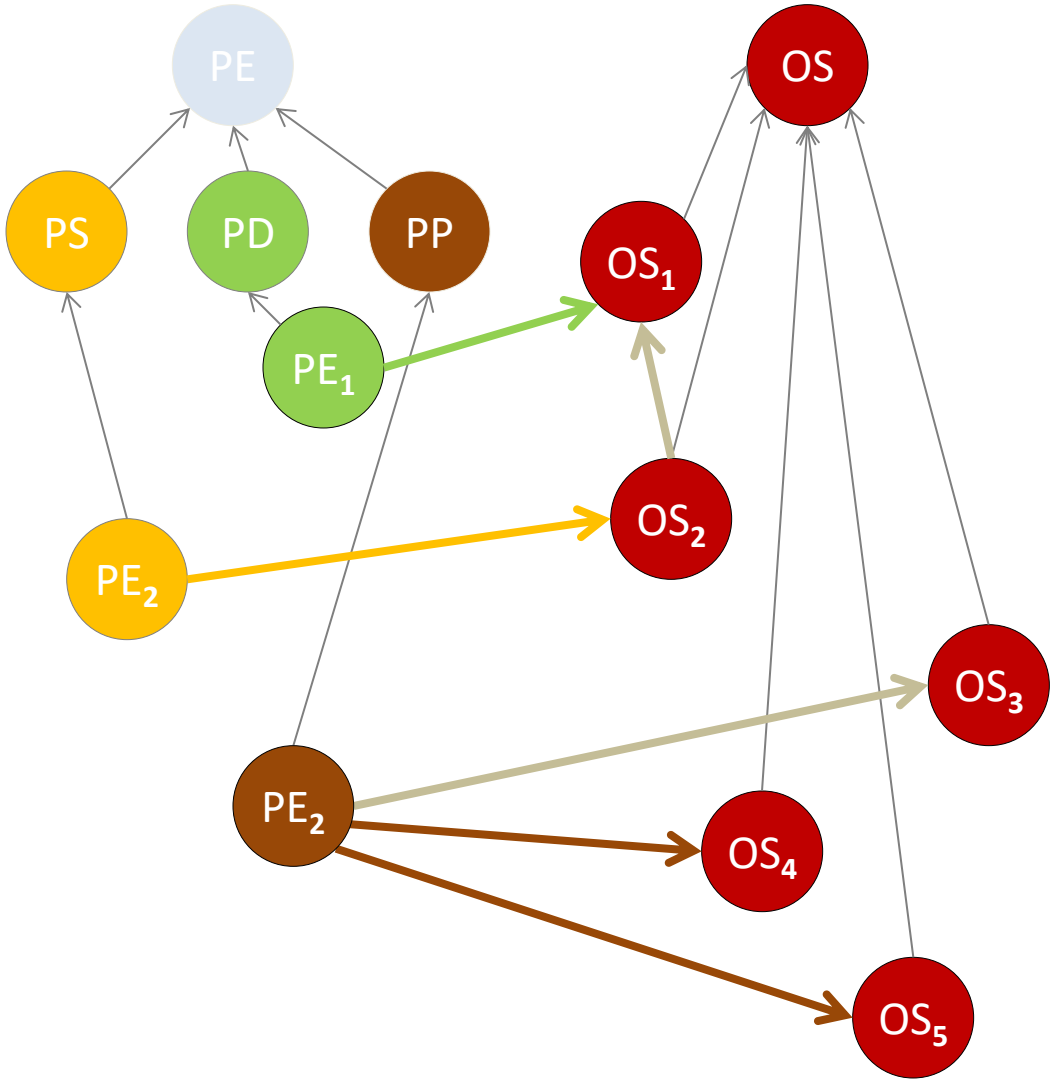
OS *Pathological Process*

  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



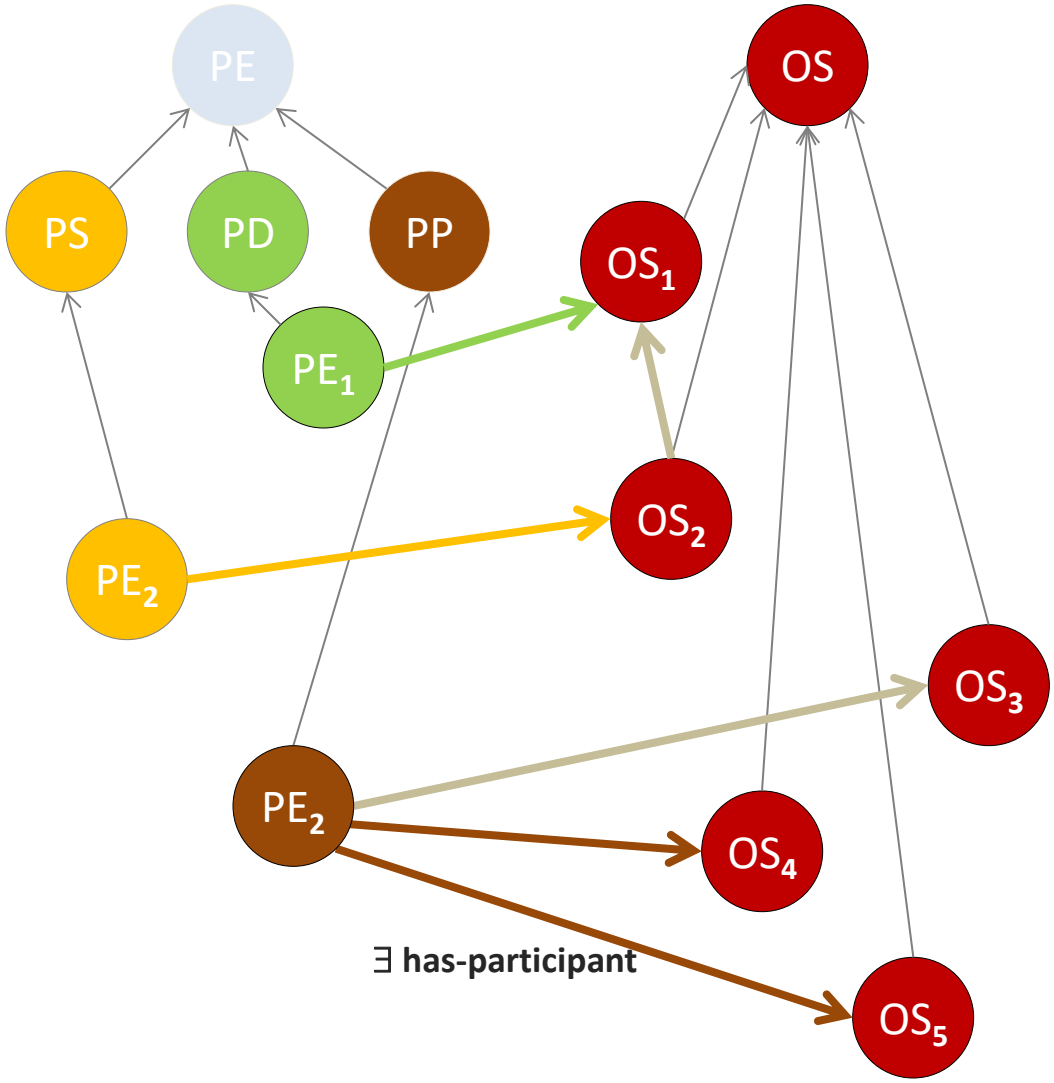
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



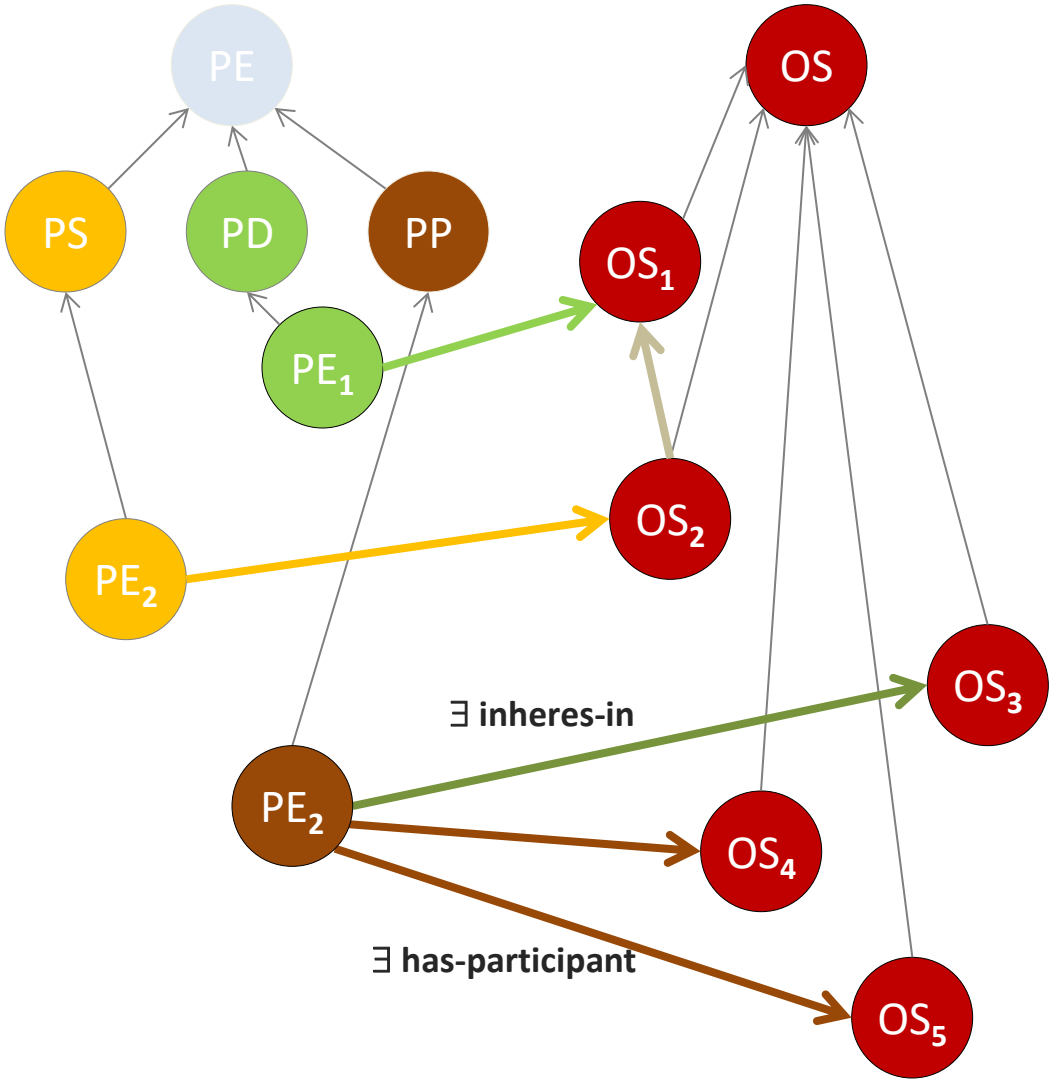
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



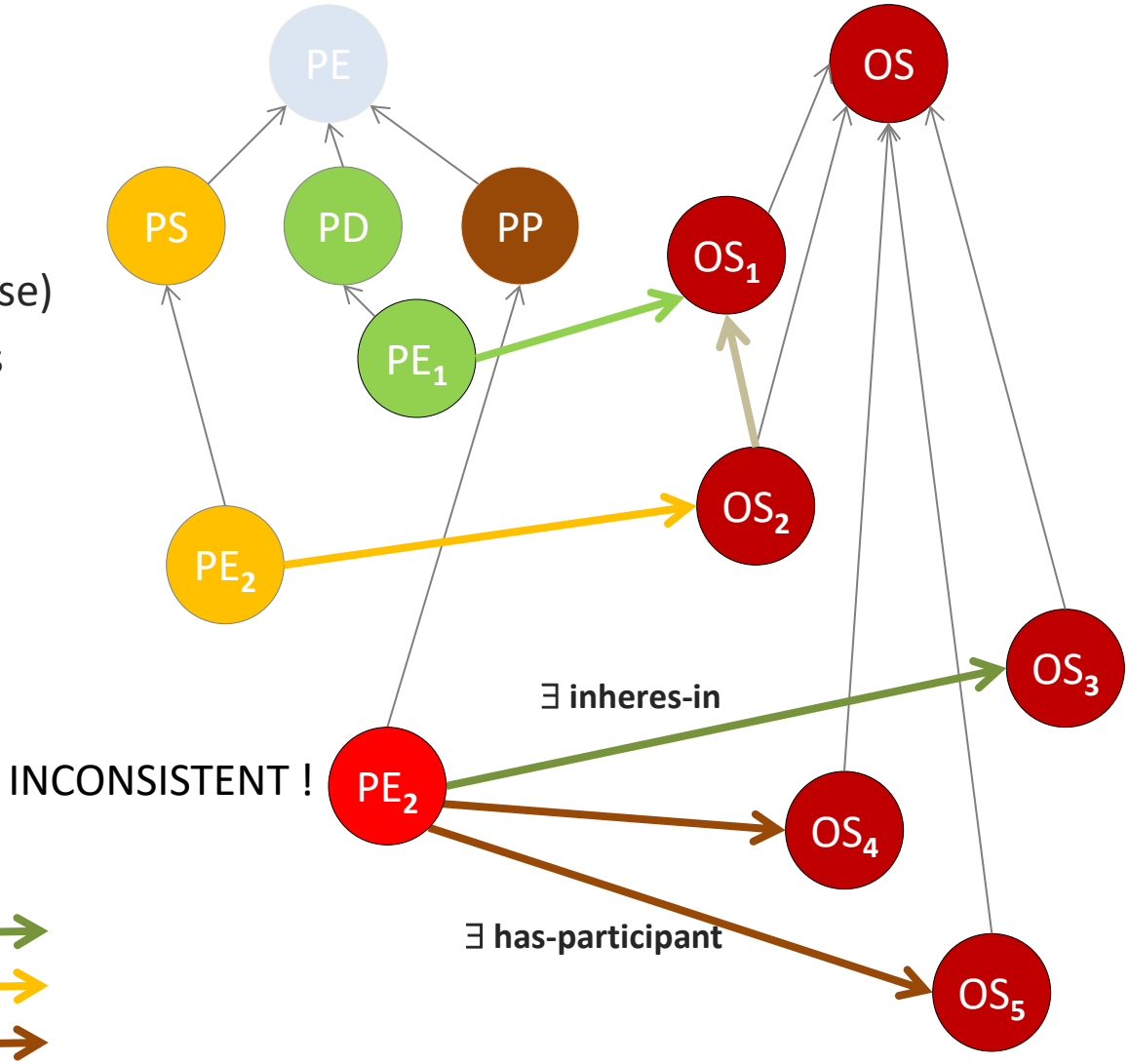
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



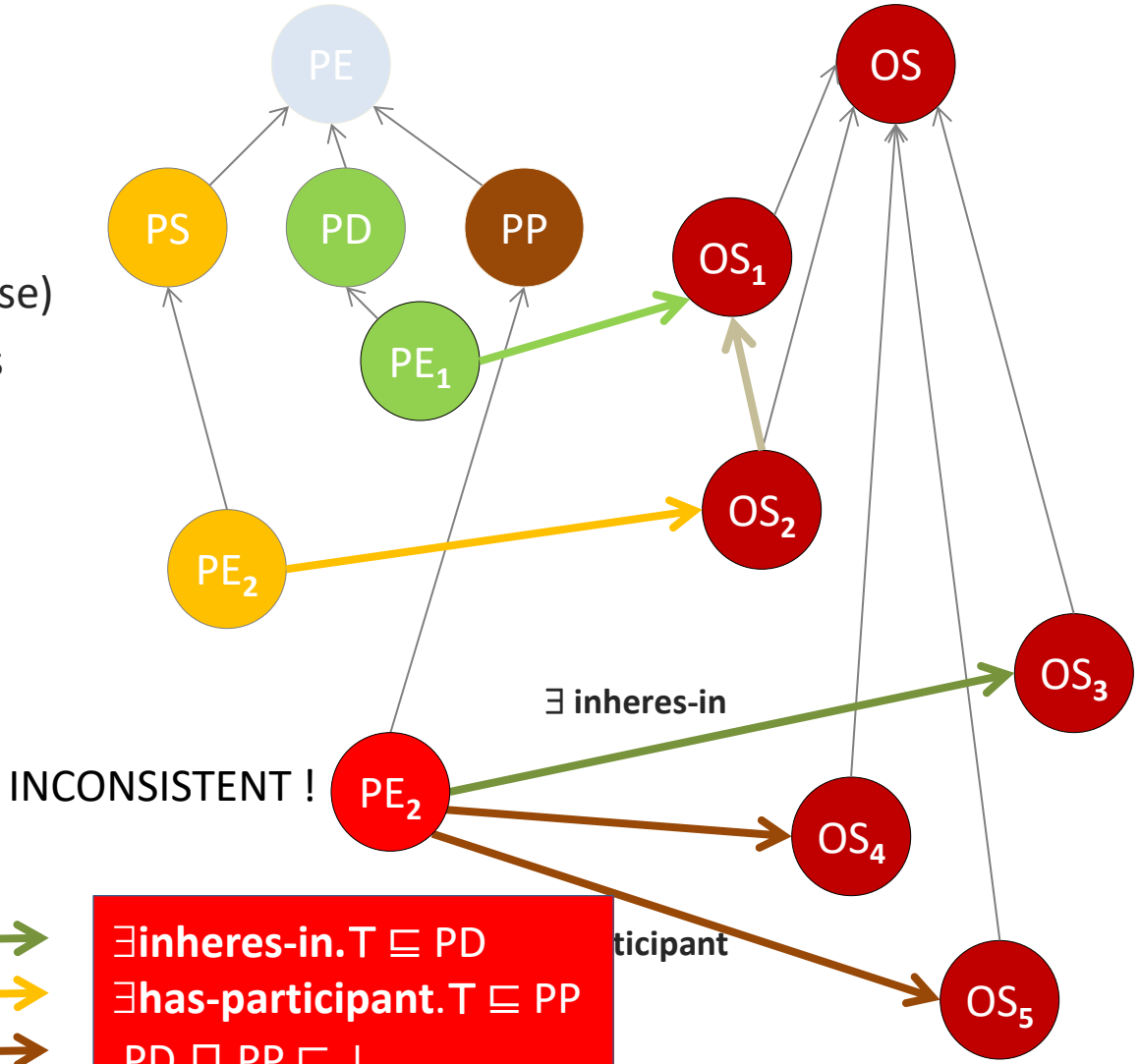
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



Construction of advanced disease ontology

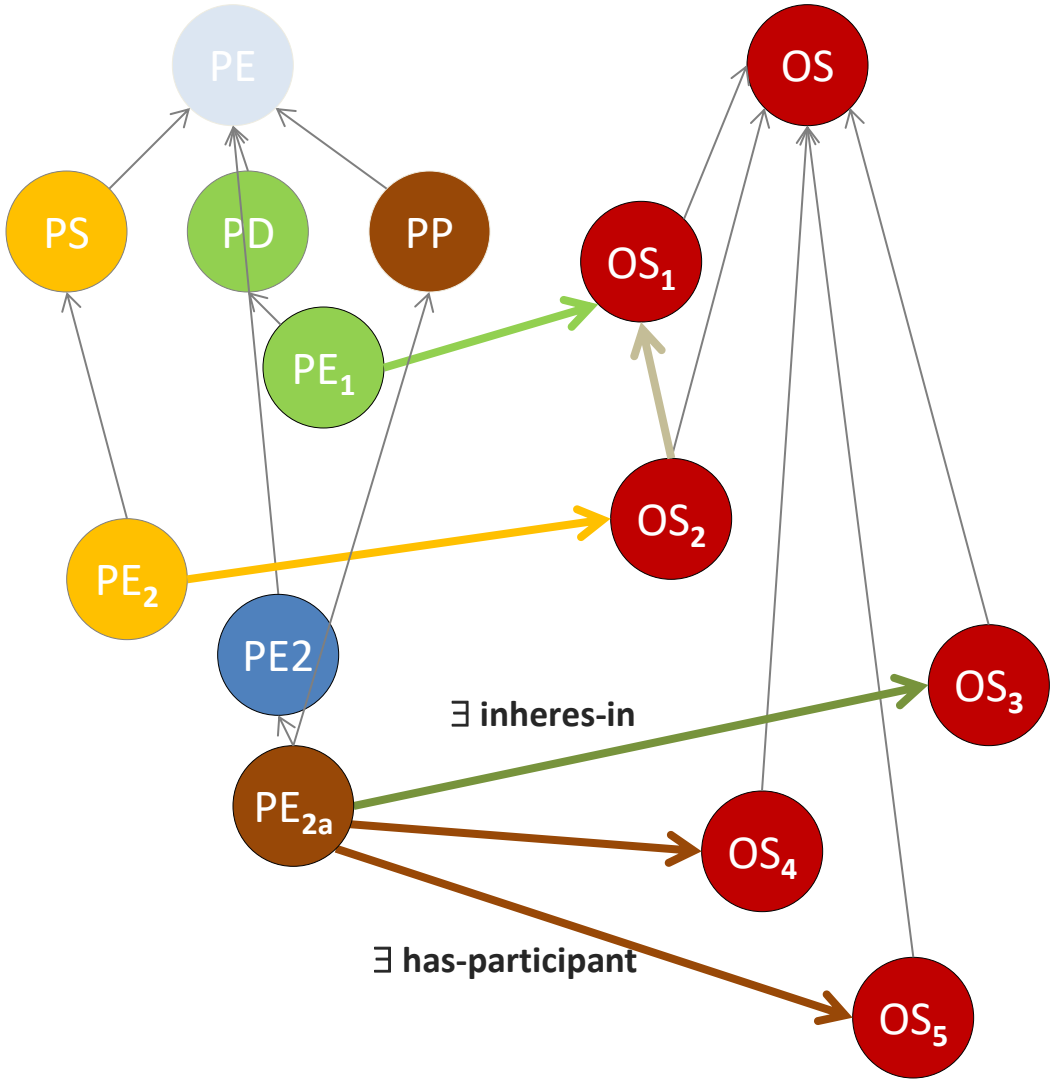
- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



∃**inheres-in**.T ⊑ PD
∃**has-participant**.T ⊑ PP
PD ⊓ PP ⊑ ⊥

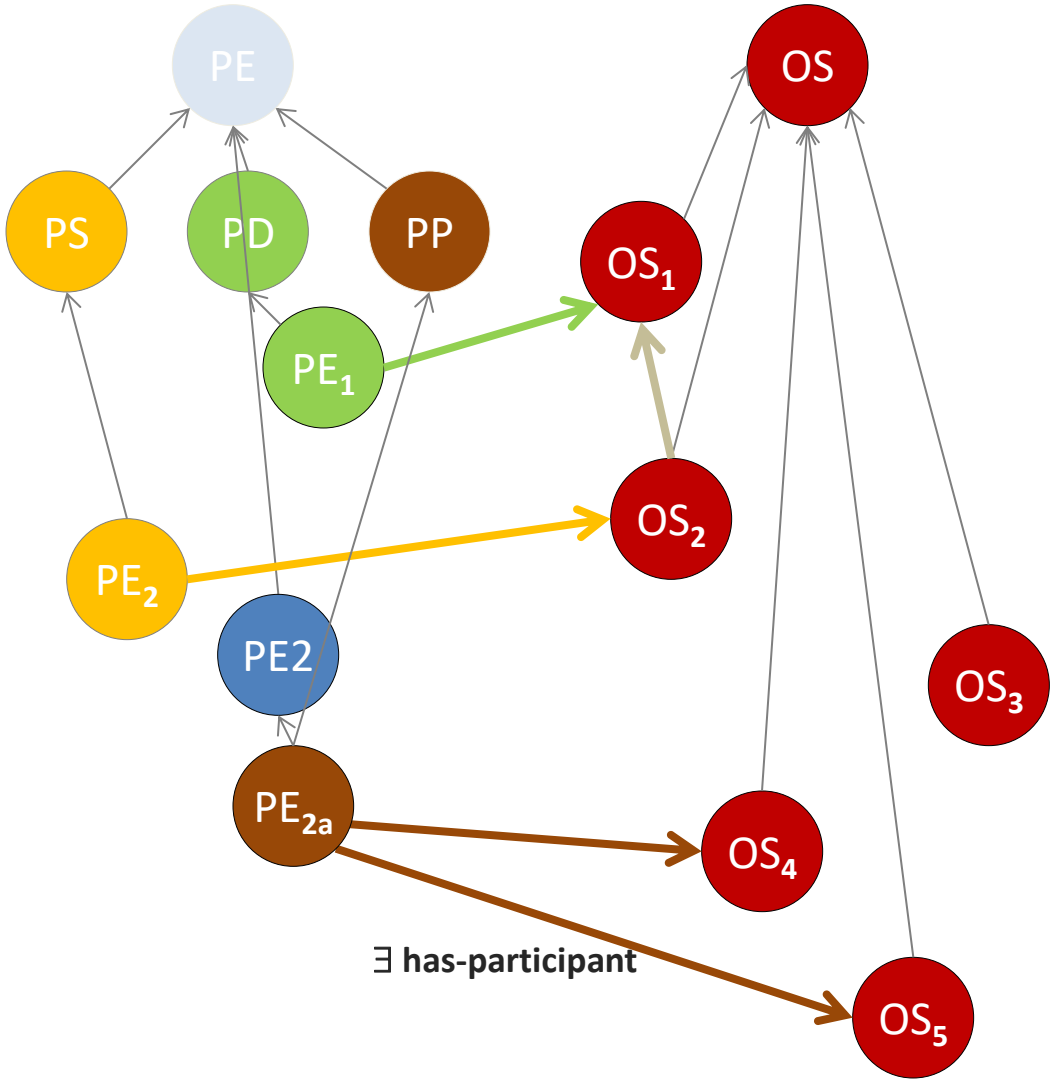
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



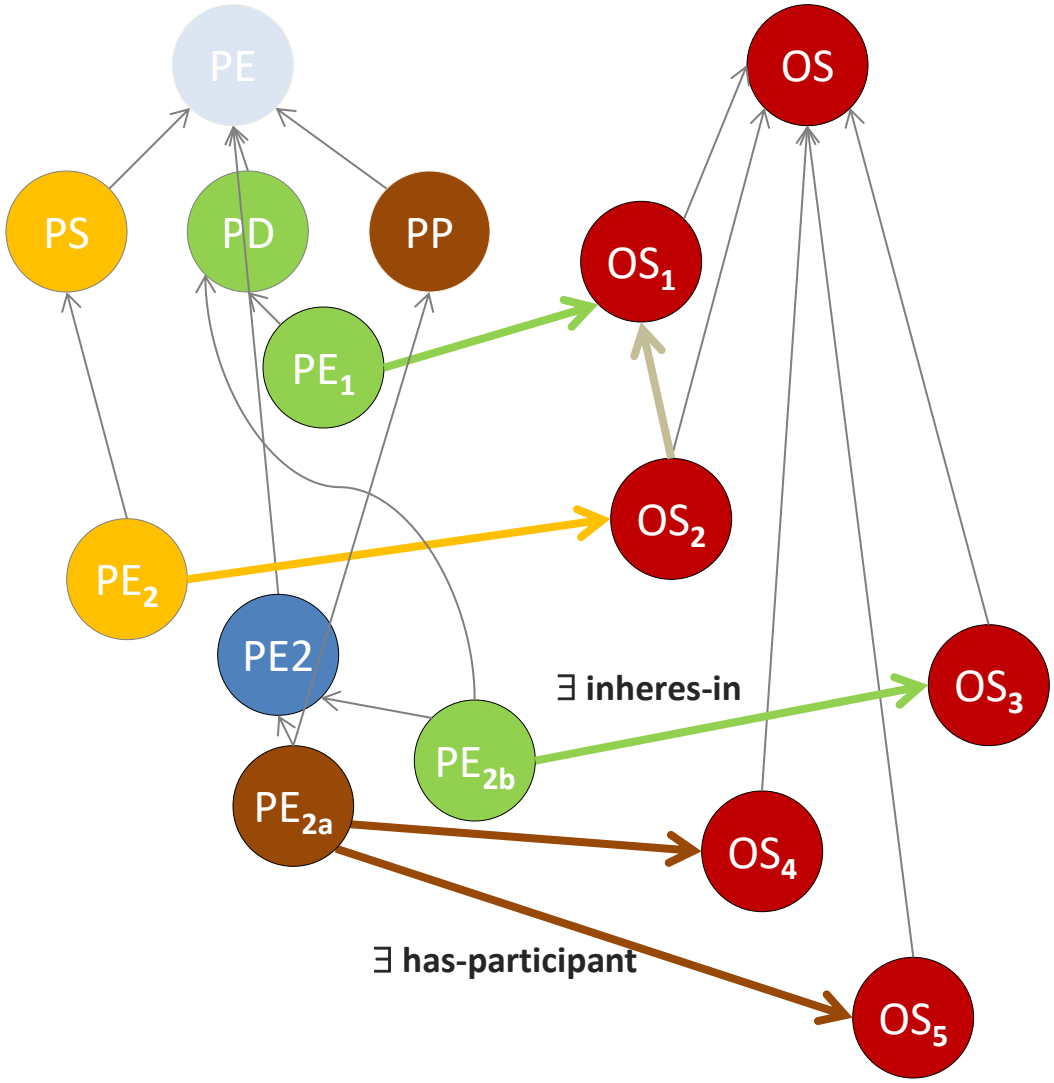
Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



Construction of advanced disease ontology

- Basic components:
  - Top nodes
    - PE *Pathological Entity*
    - OS *Organism Structure*
  - Disease classes (broad sense)
  - Organism structure classes
  - transitive relations
    - ∃ **has-locus**
    - ∃ **locus-of**
- Advanced components
  - PE *PathologicalStructure*
  - OS *PathologicalDisposition*
  - OS *Pathological Process*
  - Relations
    - ∃ **inheres-in**
    - ∃ **has-location**
    - ∃ **has-participant**



Diseases in SNOMED CT

- Can the tripartition Process / Structure / Disposition be implemented in SNOMED CT?
- Current state: Root classes:
  - Disorder: “classical” diseases
  - Findings: signs and symptoms
  - Morphology: abnormal structures
  - Events, e.g. causes of injury
- Fuzzy boundaries

Relations involved

Root	Relation to anatomical concept
Disorder	Finding Site (to anatomical site) AssociatedMorphology (to morphology)
Finding	Finding Site
Morphology	no
Event	no

Fitting of SNOMED relations with proposed model

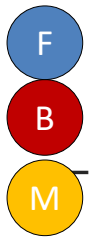
- biotop:**has-locus** corresponds to sct:**FindingSite**
- biotop:**has-participant** (with range restricted by pathological structures) corresponds to sct:**AssociatedMorphology**
- sct:**FindingSite** can be added to subconcepts of *Event* (where necessary)

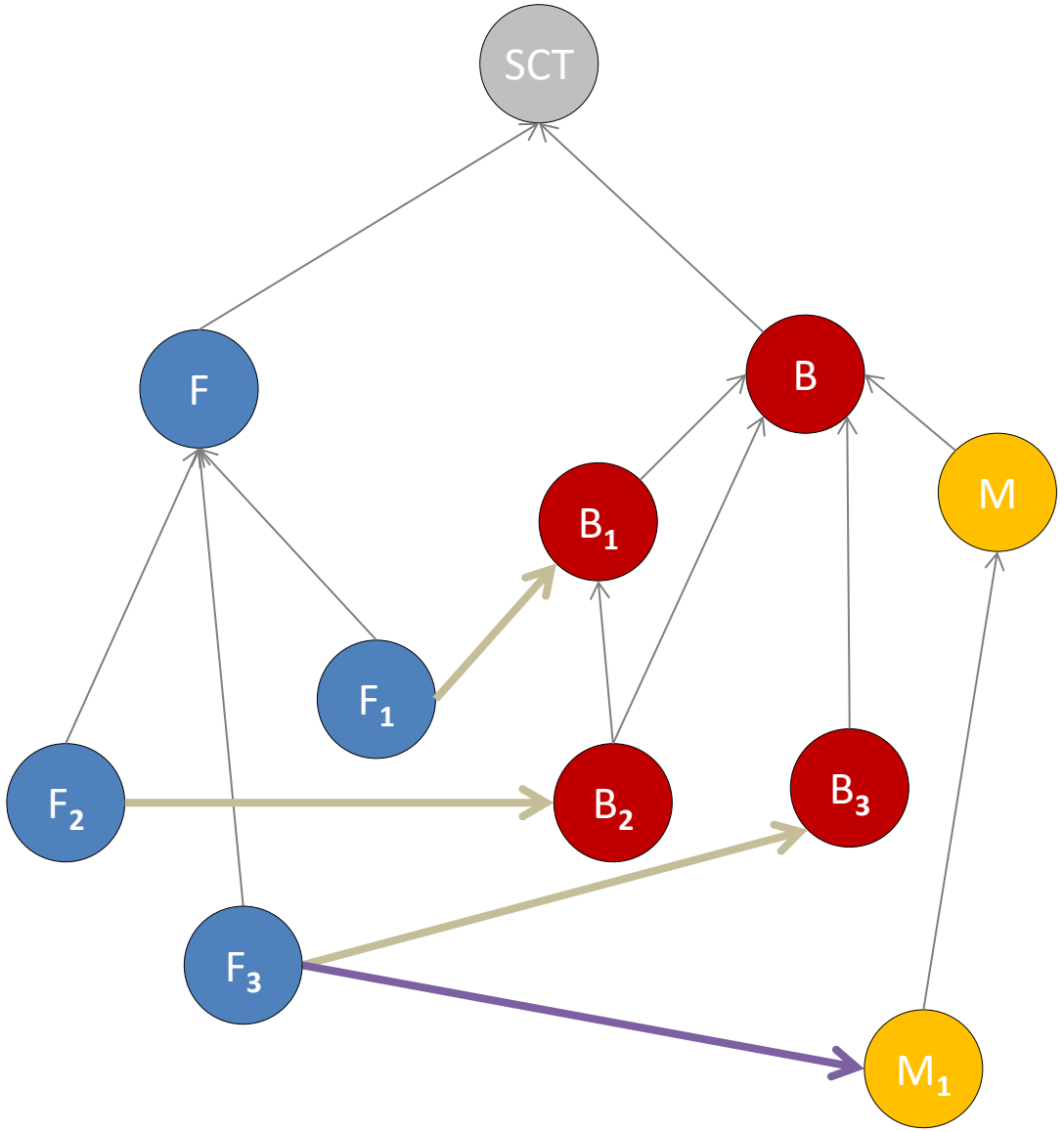
Alternative root concepts

- *Pathological Structure* can mostly be equated with *Morphology*
- New: *Pathological process* and *Pathological Disposition*
- Some *Events* are *Pathological processes*
- Default: all *Finding* and *Disease* concepts are considered disjunctions of dispositions and processes.
  - Where it is obvious that a concept (and all of its subconcepts) are to interpreted as dispositions only, the parent *Pathological Disposition* is added. Then their relations both to anatomical entities and to morphologies (*FindingSite*, *AssocMorphology*) can be understand as inherence (although this relation is not introduced)
  - Where it is obvious that a concept (and all of its subconcepts) are to interpreted as processes only, the parent *Pathological Process* is added. Then their relations both to anatomical entities and to morphologies (*FindingSite*, *AssocMorphology*) can be interpreted as location and participation (although this relation is not introduced)

Current state in SNOMED

- Basic components:
  - Top nodes
    - *Finding*
    - *BodyStructure (S-Nodes)*
    - *MorphologicalAbnormality (Pathological Structure)*



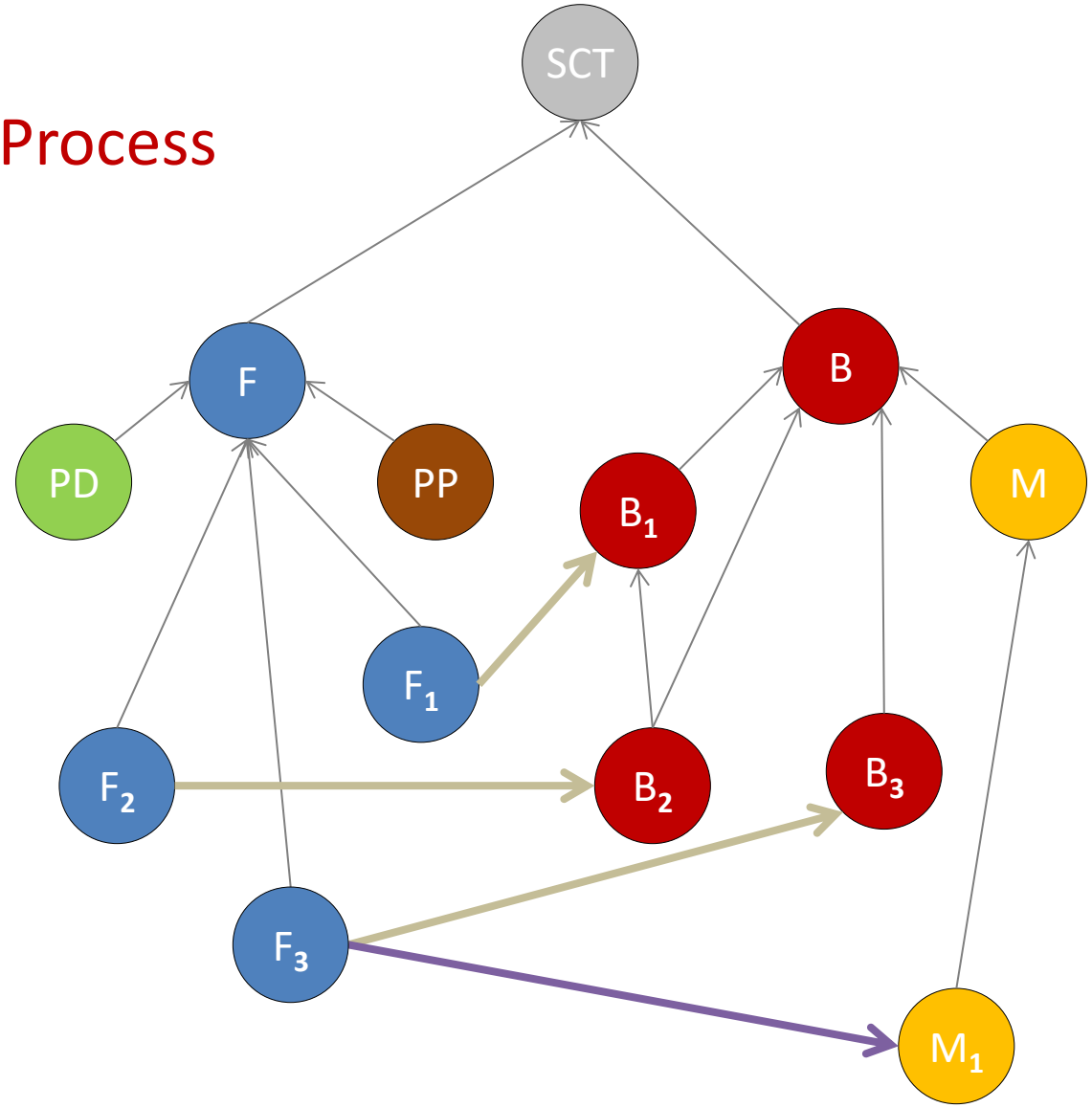


  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

Step 1:

Addition of Pathological Process and Disposition

- Basic components:
  - Top nodes
    - F *Finding*
    - B *BodyStructure (S-Nodes)*
    - M *MorphologicalAbnormality (Pathological Structure)*
    - PD *PathologicalDisposition*
    - PP *PathologicalProcess*

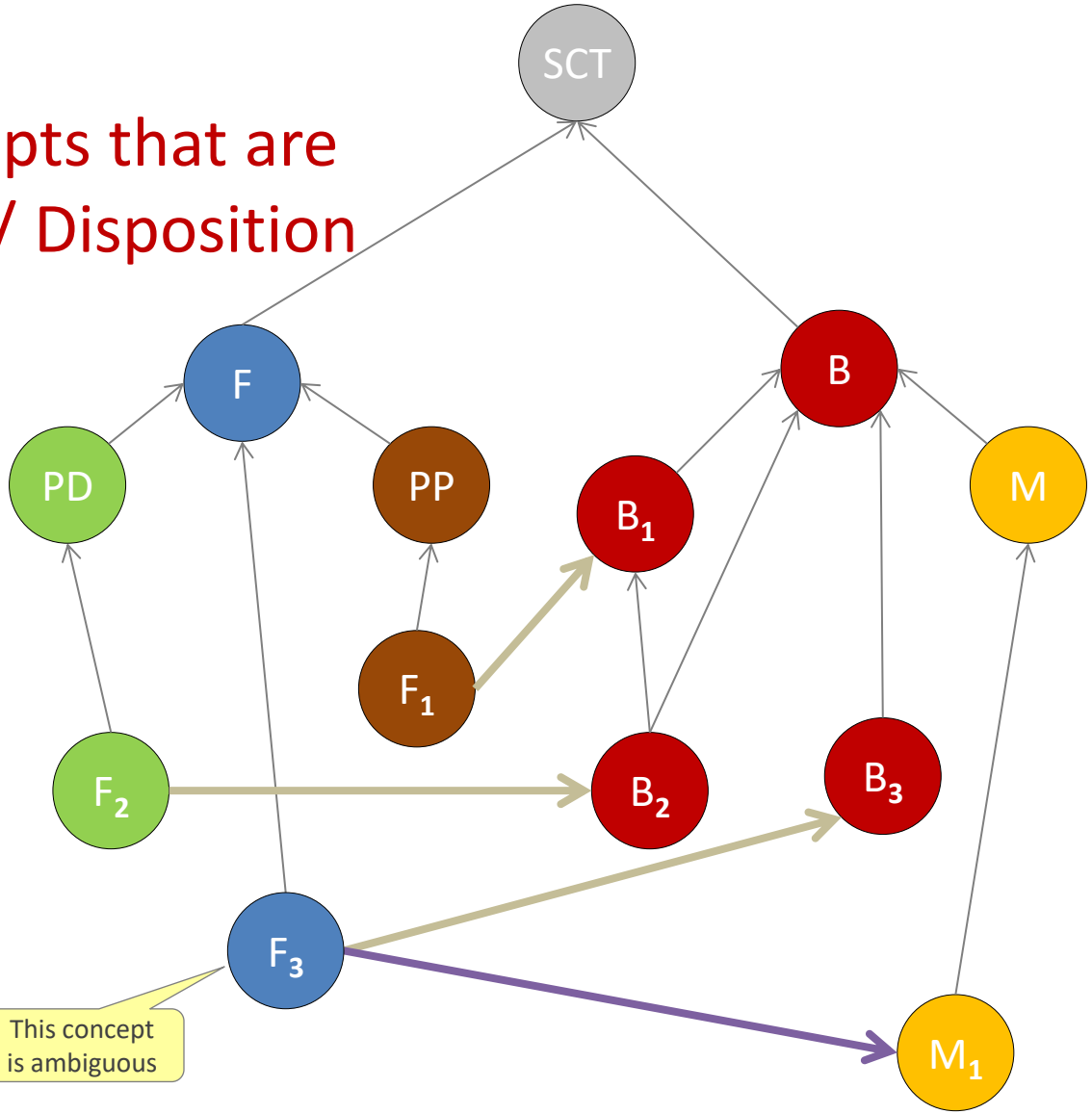


  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

Step 2:

Rearrangement of Concepts that are unambiguous re Process/ Disposition

- Basic components:
  - Top nodes
    - F *Finding*
    - B *BodyStructure (S-Nodes)*
    - M *MorphologicalAbnormality (Pathological Structure)*
    - PD *PathologicalDisposition*
    - PP *PathologicalProcess*

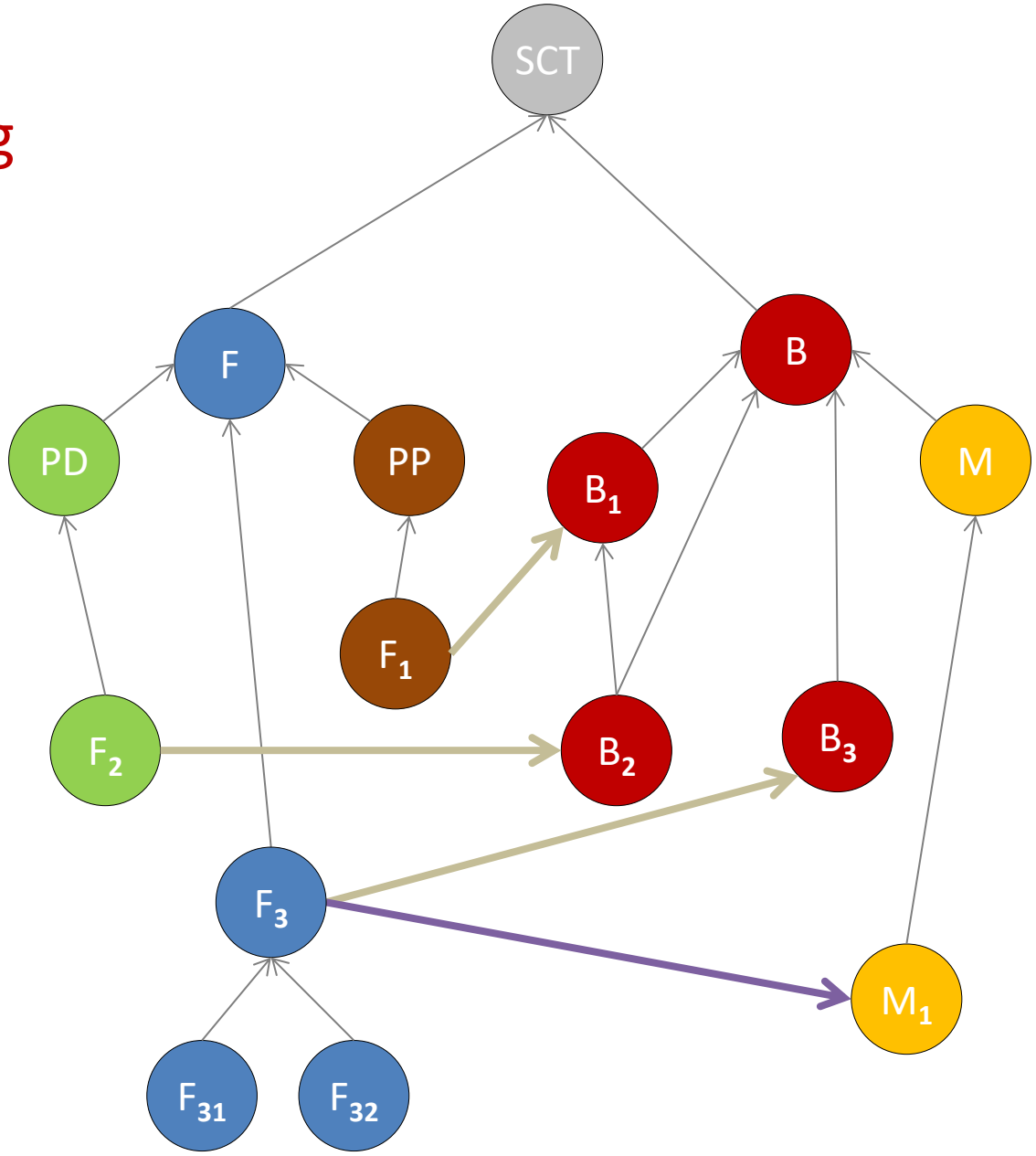


  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

Step 3:

Disambiguate by creating subconcepts

- Basic components:
  - Top nodes
    - F *Finding*
    - B *BodyStructure (S-Nodes)*
    - M *MorphologicalAbnormality (Pathological Structure)*
    - PD *PathologicalDisposition*
    - PP *PathologicalProcess*

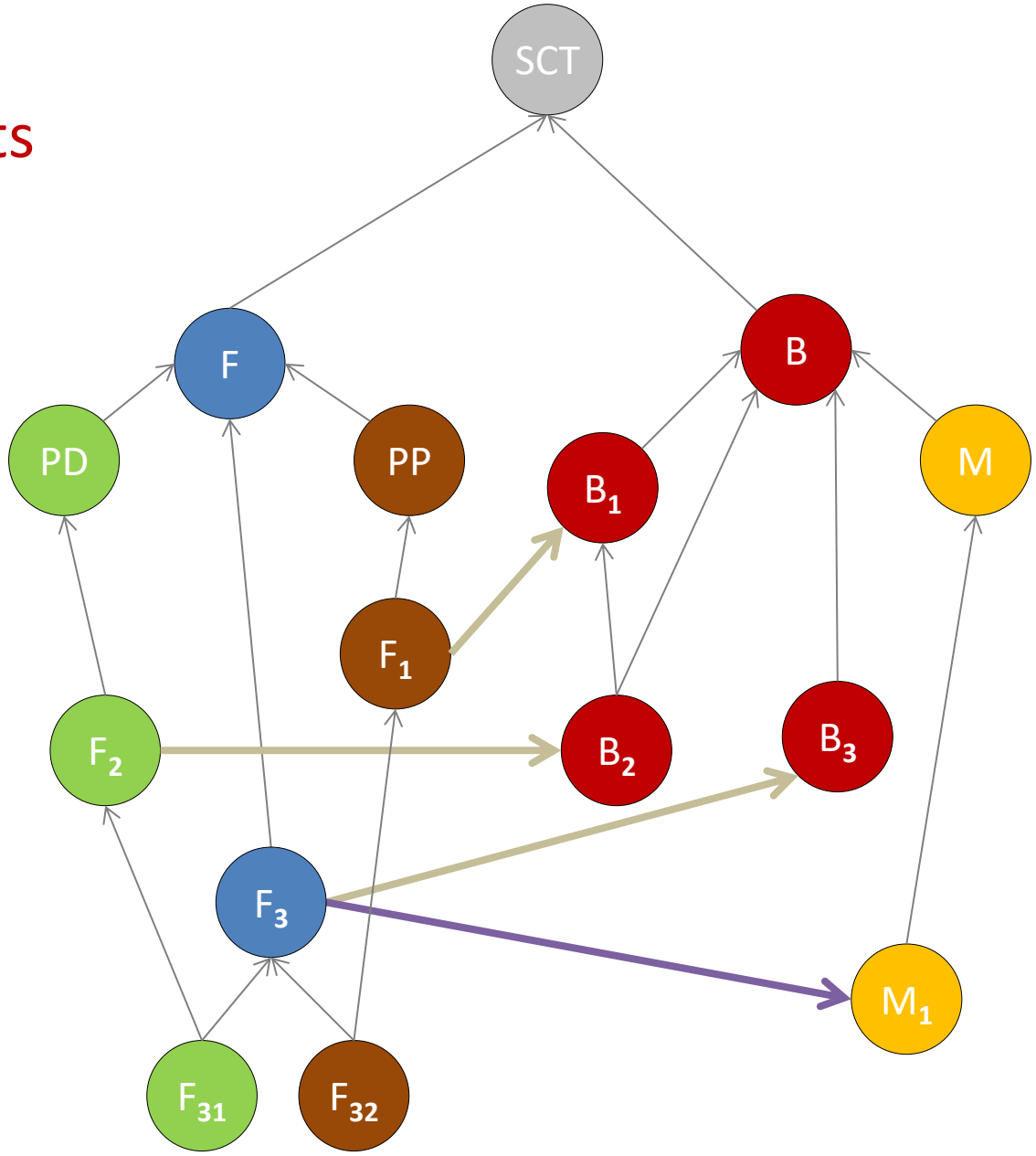


  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

Step 4:

fully defining subconcepts by conjunction

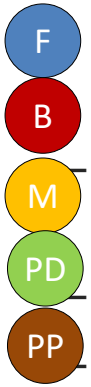
- Basic components:
  - Top nodes
    - F *Finding*
    - B *BodyStructure (S-Nodes)*
    - M *MorphologicalAbnormality (Pathological Structure)*
    - PD *PathologicalDisposition*
    - PP *PathologicalProcess*

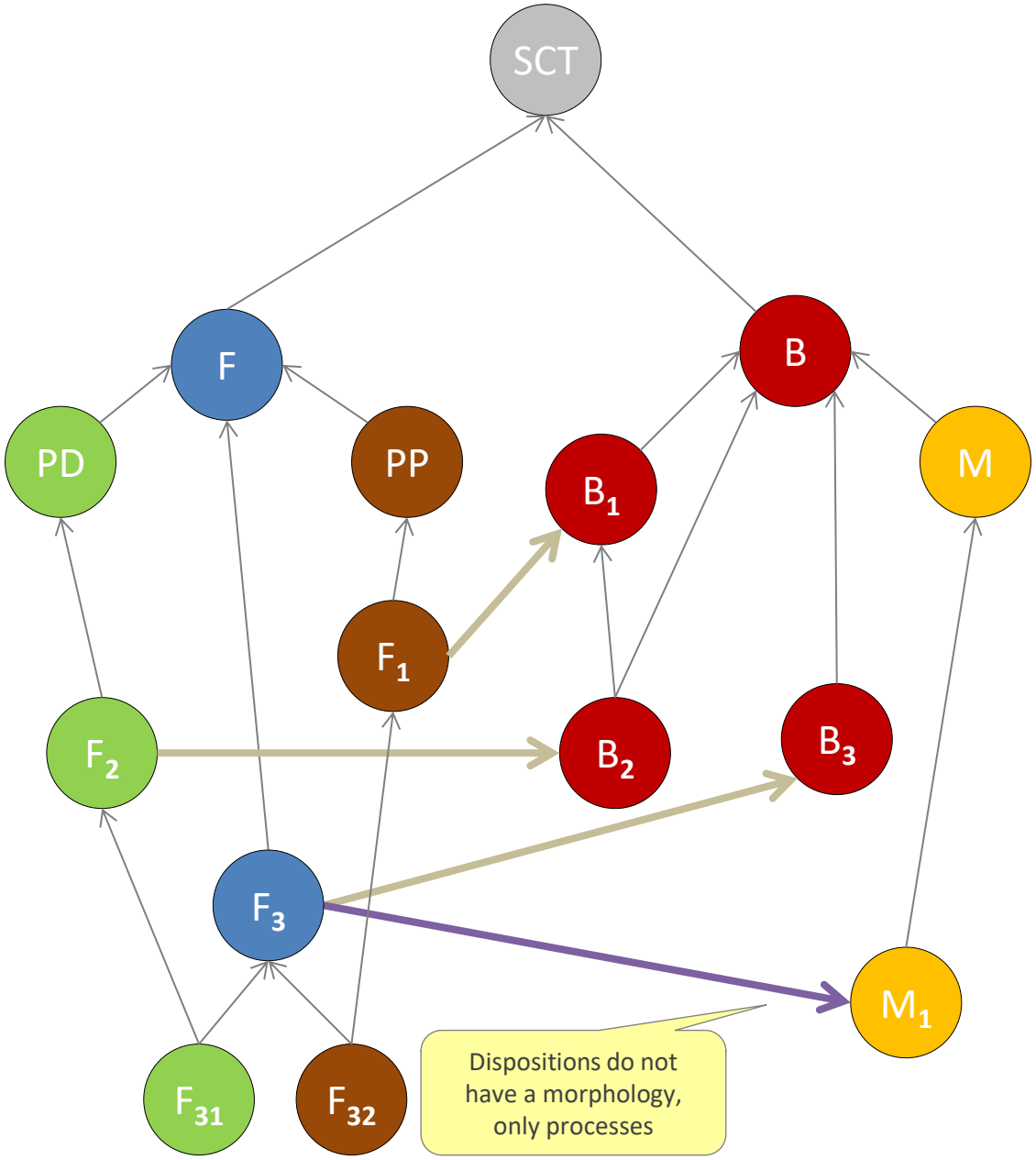


  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

Step 5:
rearrangement of relations

- Basic components:
  - Top nodes
    - *Finding*
    - *BodyStructure (S-Nodes)*
    - *MorphologicalAbnormality (Pathological Structure)*
    - *PathologicalDisposition*
    - *PathologicalProcess*

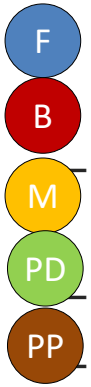


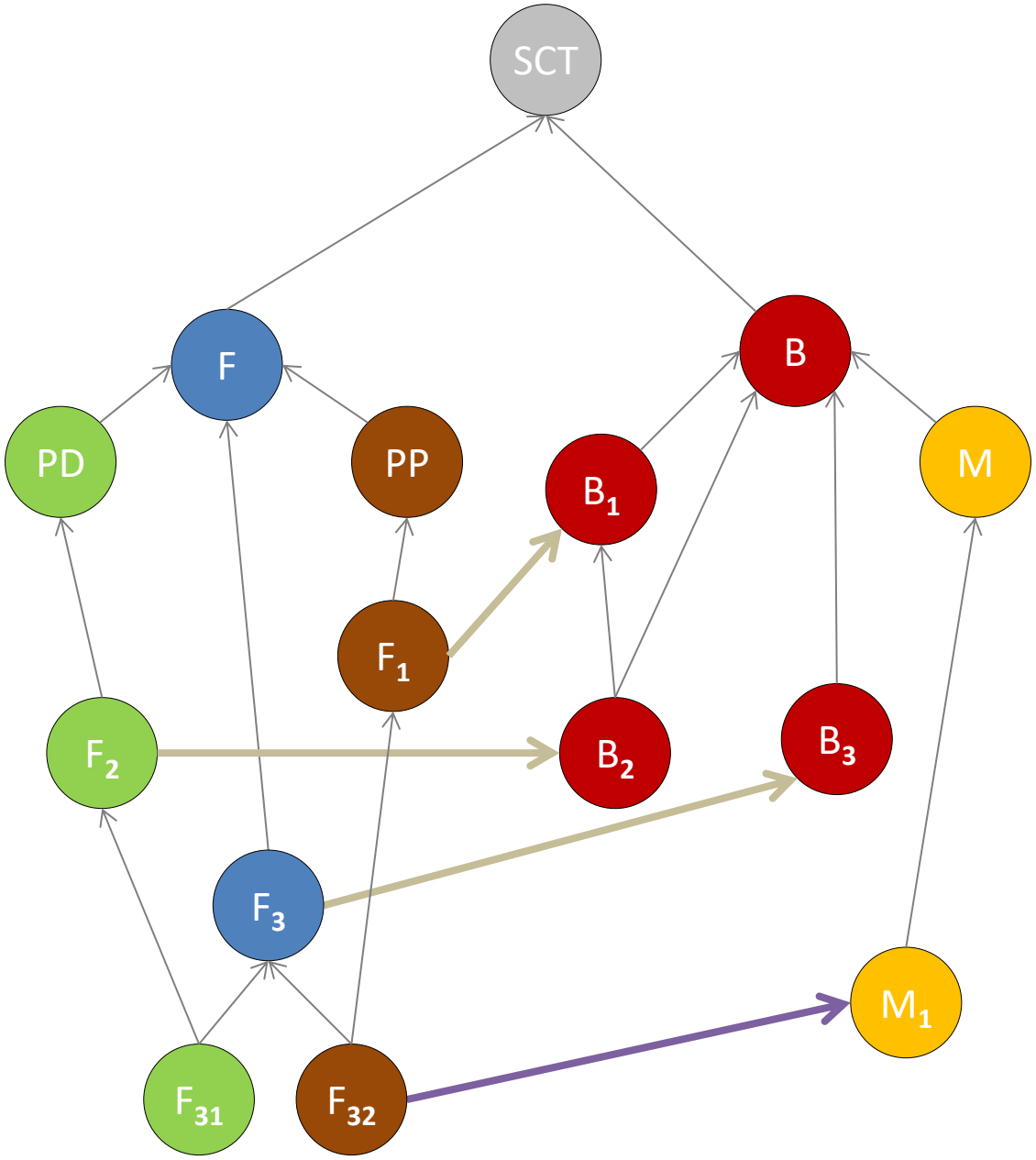


  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

Step 5:
rearrangement of relations

- Basic components:
  - Top nodes
    - *Finding*
    - *BodyStructure (S-Nodes)*
    - *MorphologicalAbnormality (Pathological Structure)*
    - *PathologicalDisposition*
    - *PathologicalProcess*





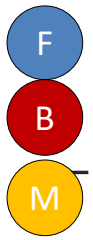
  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**

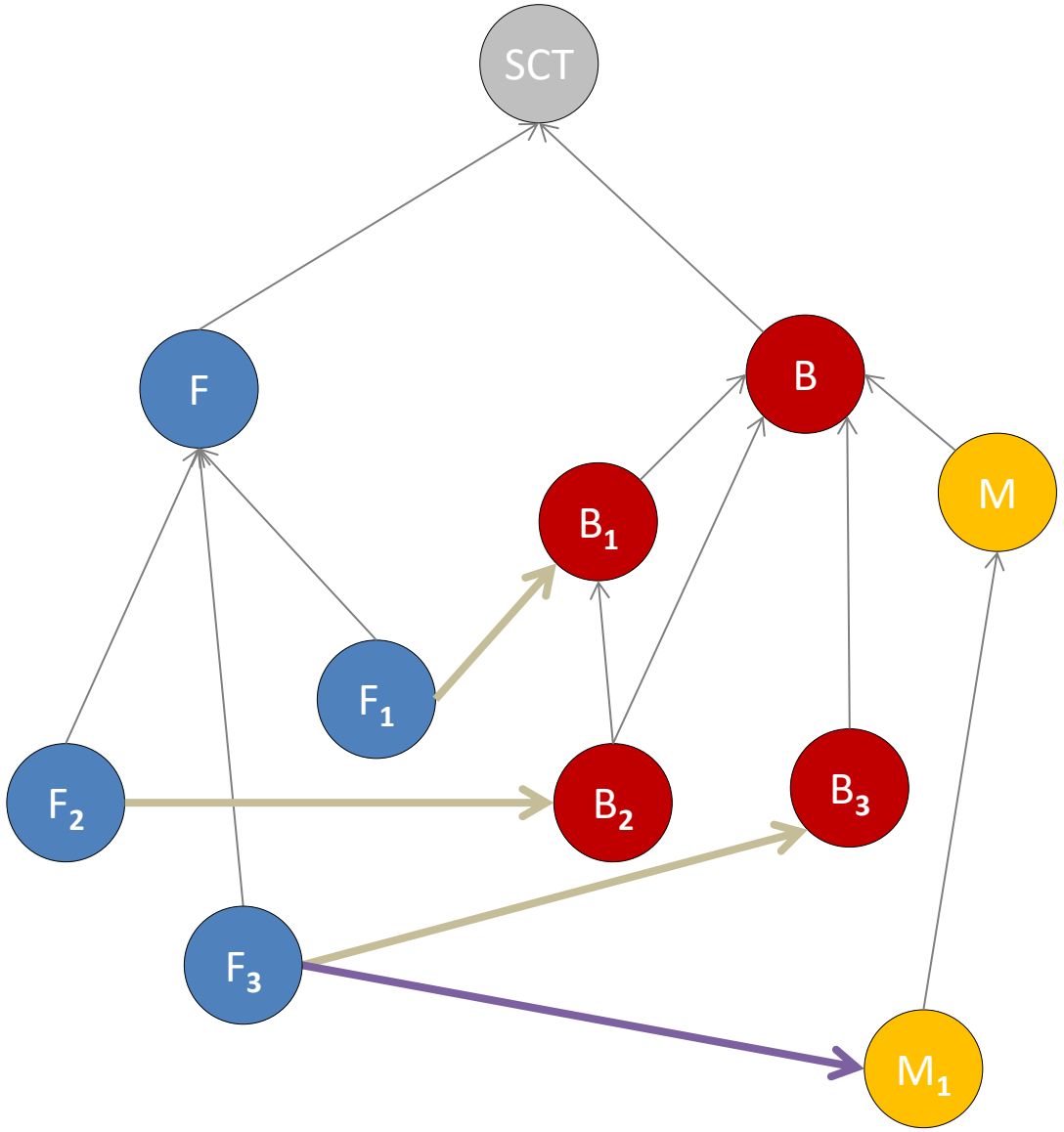
Conclusions

- "Disease": ontologically polymorphic category
- Refinement of disease classes into pathological structures, dispositions, and processes often not necessary
- Introduction of umbrella category *Pathological entity*, together with the high-level relation **has-locus**:
  - construction of simple model which already supports important inferences
  - permits graceful evolution towards more sophisticated models in which the above distinctions are introduced where necessary
- Can be implemented in SNOMED with maintaining the **findingSite** relation for **has-locus**
- Open question: what to do with the pathological structure concepts that are not in the *Morphology* branch?

Current state in SNOMED

- Basic components:
  - Top nodes
    - *Finding*
    - *BodyStructure (S-Nodes)*
    - *MorphologicalAbnormality (Pathological Structure)*





  - relations
    - ∃ **FindingSite**
    - ∃ **AssocMorphology**