Terminology Systems and Protege

Ameen Abu−Hanna
Ronald Cornet
Medical Informatics
University of Amsterdam

Monica Crubezy
Samson Tu
Stanford Medical Informatics
Stanford University

Outline

- Understanding Terminology Systems
- Protege counter−parts

Terminology Systems

Many T systems

- ICD
- NHS Clinical Terms
- SNOMED
- ...

Desiderata

- Concept oriented
- Explicit Rels
- Composition Rules

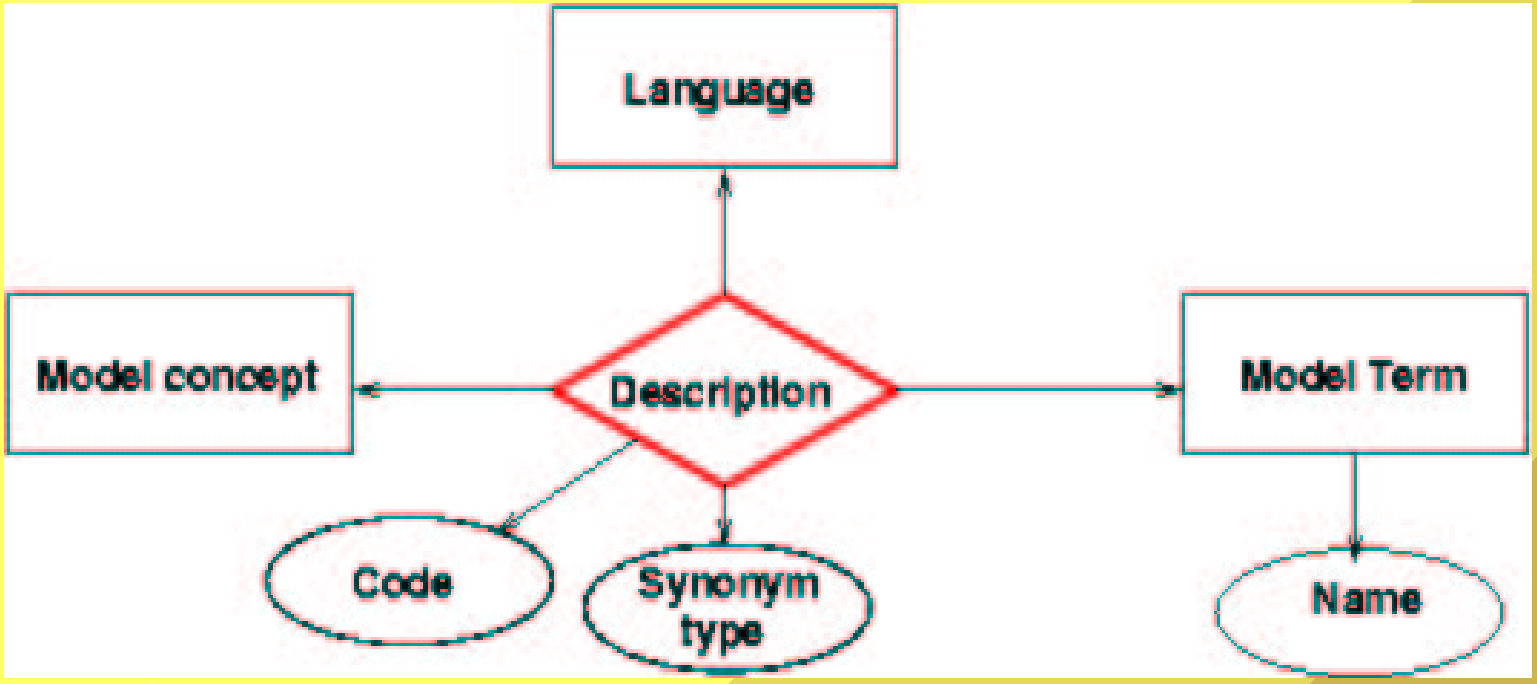
- Domain Completeness
- Multiple classification
- Use of definitions
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

Desiderata

- Concept oriented
- Explicit Rels
- Composition Rules

- Domain Completeness
- Multiple classification
- Use of definitions
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

Meta Model


Language
Model concept
Description
Model Term
Code
Synonym type
Name

FOL constraint on meta−model

For every language there is exactly one preferred term per concept

for−all c in Concept
for−all l in Language
exists−! t in Term #meaning exactly one

(and (Description(c, l, t)
Synonym−type(<c,l,t>)="preferred"))

FOL constraints on meta−model

Different concepts cannot have the same preferred term

```
for−all c1 in Concept
for−all l in Language
for−all t in Term

(==> (and Description(c1, l, t)
          Synonym−type(<c1, l, t>)="preferred")

          (not−exists c2 in Concept
          (and Description(c2, l, t)
                Synonym−type(<c2, l, t>)="preferred"
                c2 != c1)))
```

How to specify a new TS?

Intensive Care

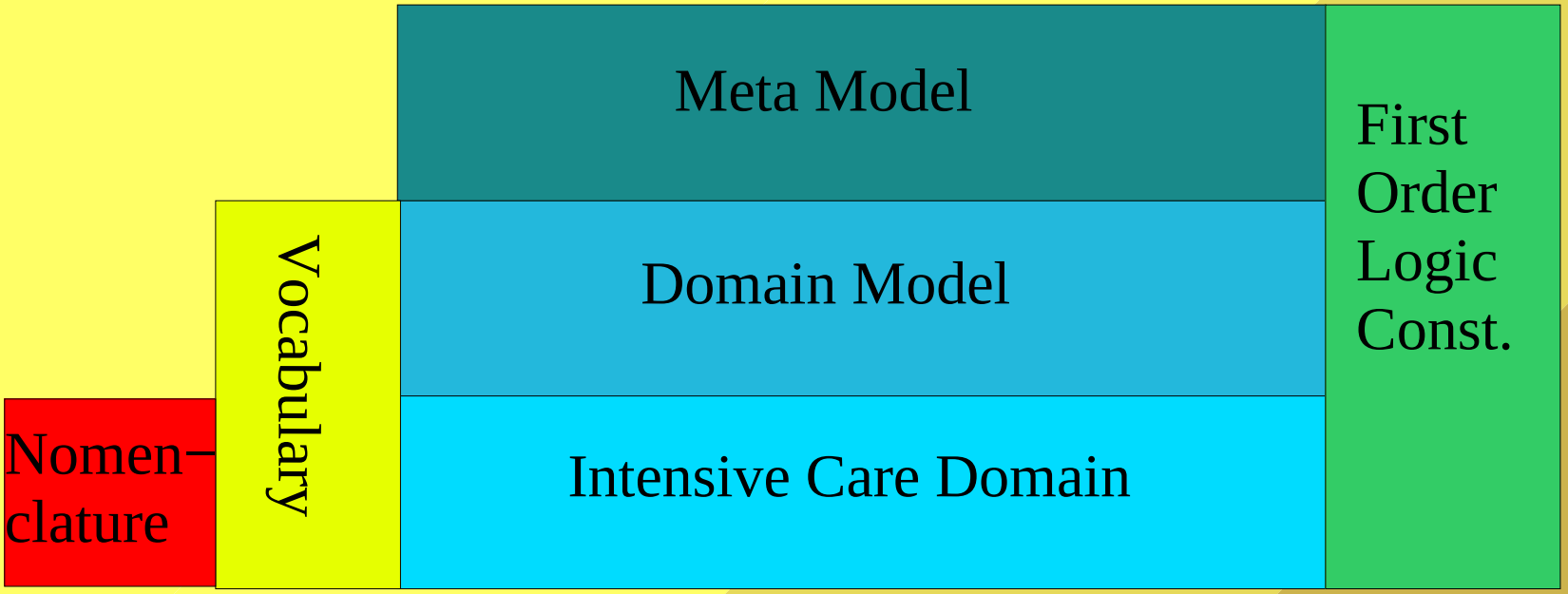
Intensive Care Knowledge Base

± 2500 concepts: Diseases

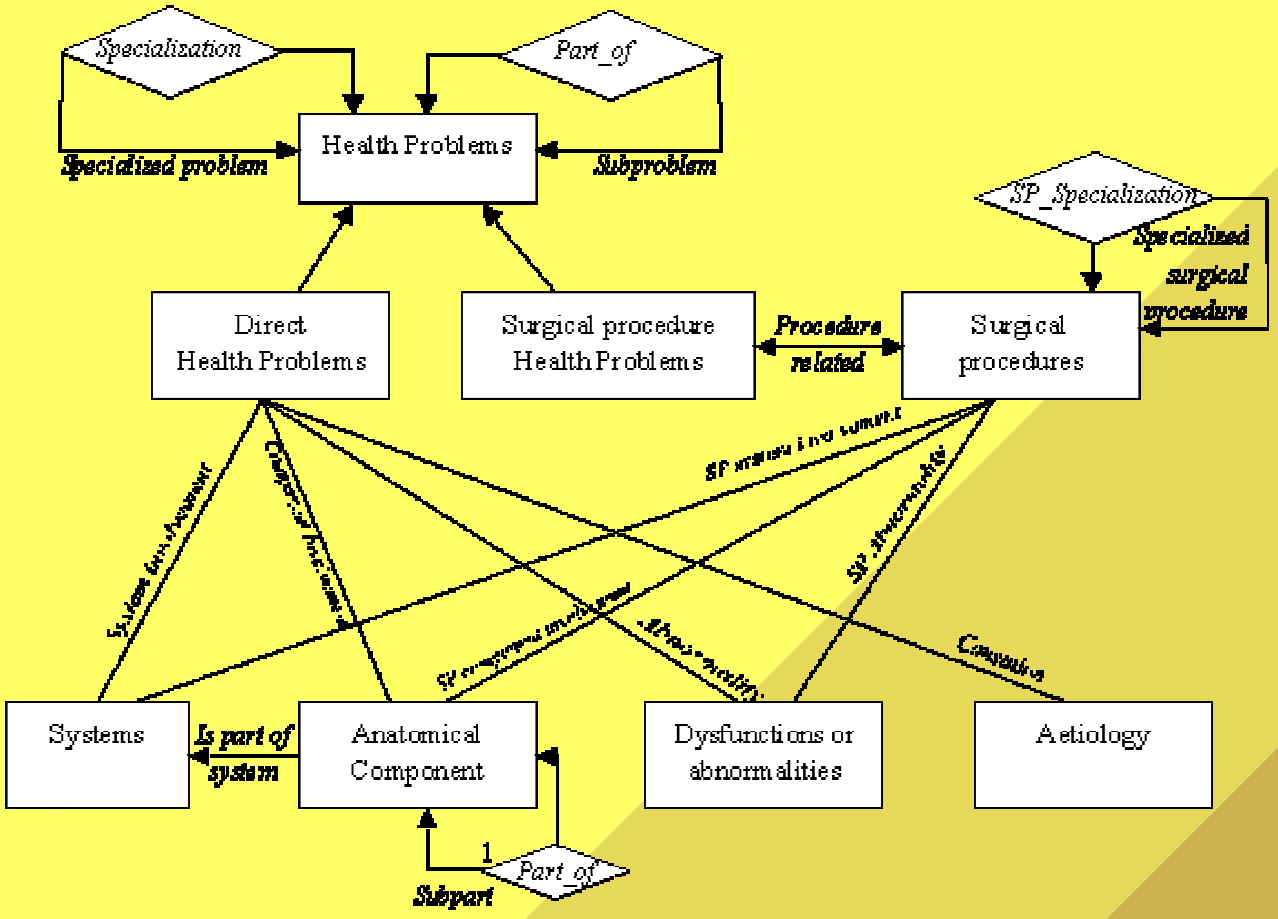
Anatomy, Systems, Abnormalities, Etiologies

- **Hepatitis**: ***isa* Disease**,
  location **Liver**,
  abnormality **Inflammation**
- **Infective Hepatitis**: *isa* **Hepatitis**,
  cause **Micro−organism**

Knowledge Components in DICE

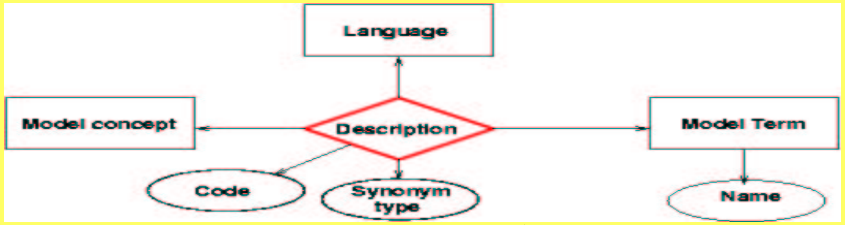


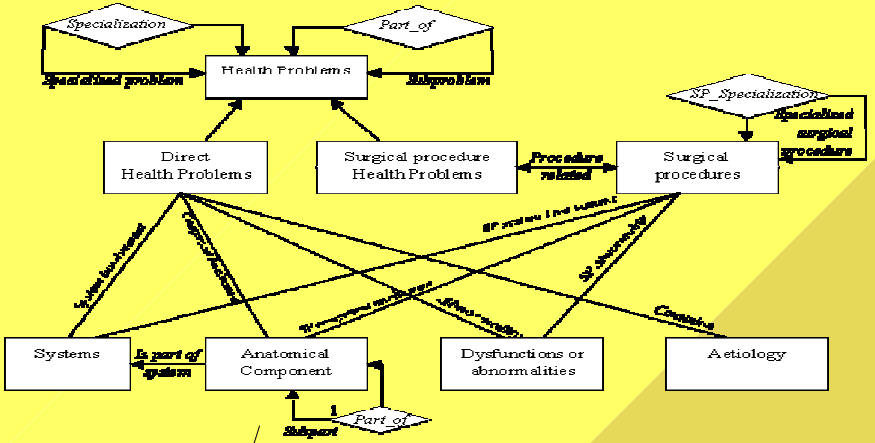
Domain Model in DICE



Nomenclature/Defs in DICE

Disease	Anatomical Component	Abnormality	Aetiology
Hepatitis	Def: Liver	Def: Infection	OR: (HepV, E−BV, CytomegaloV)





Meta Model

Domain Model

Intensive Care Domain

First Order Logic Const.

Unique−codes

Vocabulary

Nomen−clature

Hepatitis Def(..)

Protege?

Knowledge model in Protege

- **Concept oriented**
- **Explicit Rels**
- Composition Rules

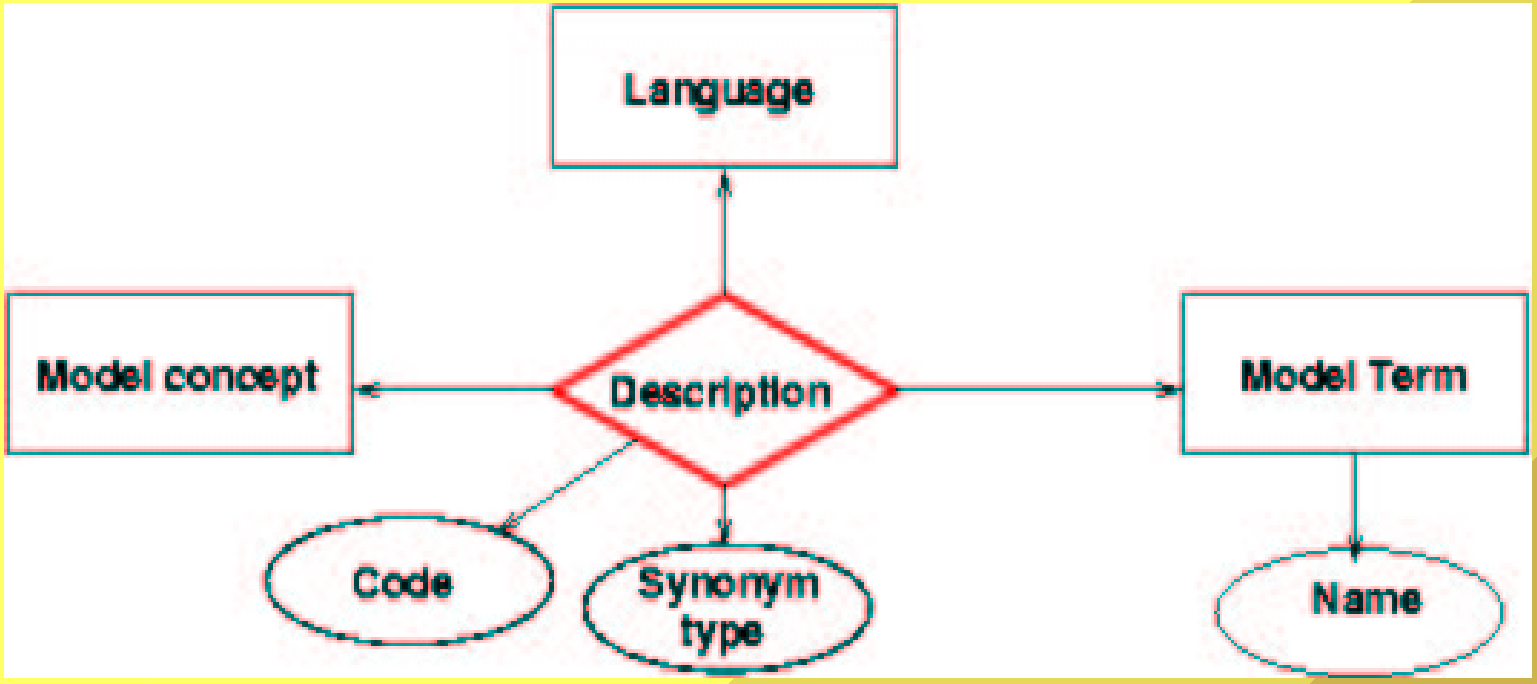
- **Domain Completeness**
- **Multiple classification**
- **Use of definitions**
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

Knowledge model in Protege

- **Concept oriented**
- **Explicit Rels**
- Composition Rules

- **Domain Completeness**
- **Multiple classification**
- **Use of definitions**
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

Meta Model

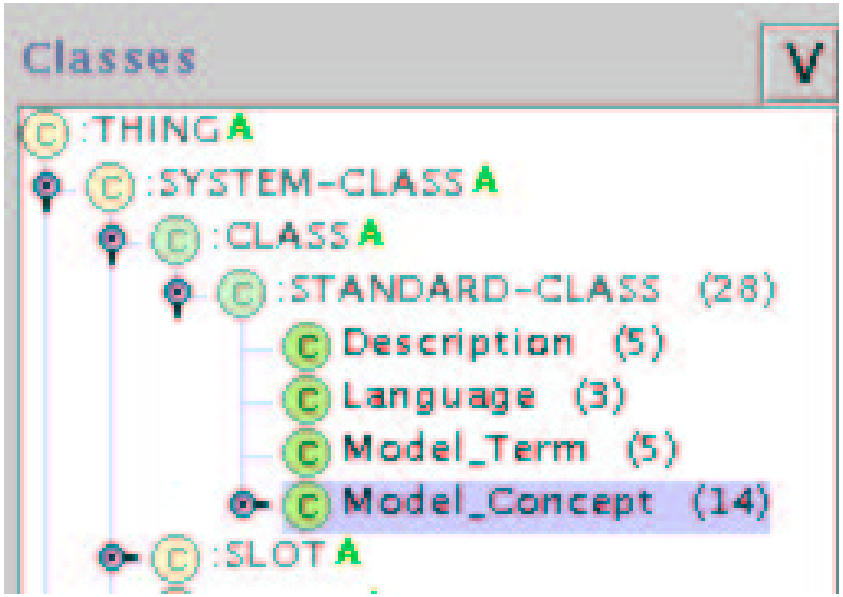

Language
Model concept
Description
Model Term
Code
Synonym type
Name

Meta Model in Protege: Using **Meta Classes**

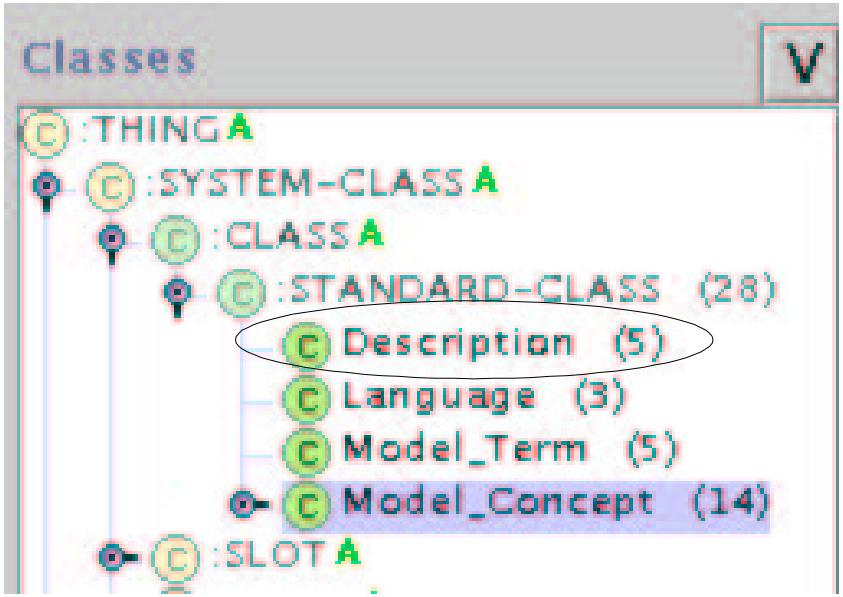
Standard class

- Description
- Language
- Model concept
- Model term

Meta Classes



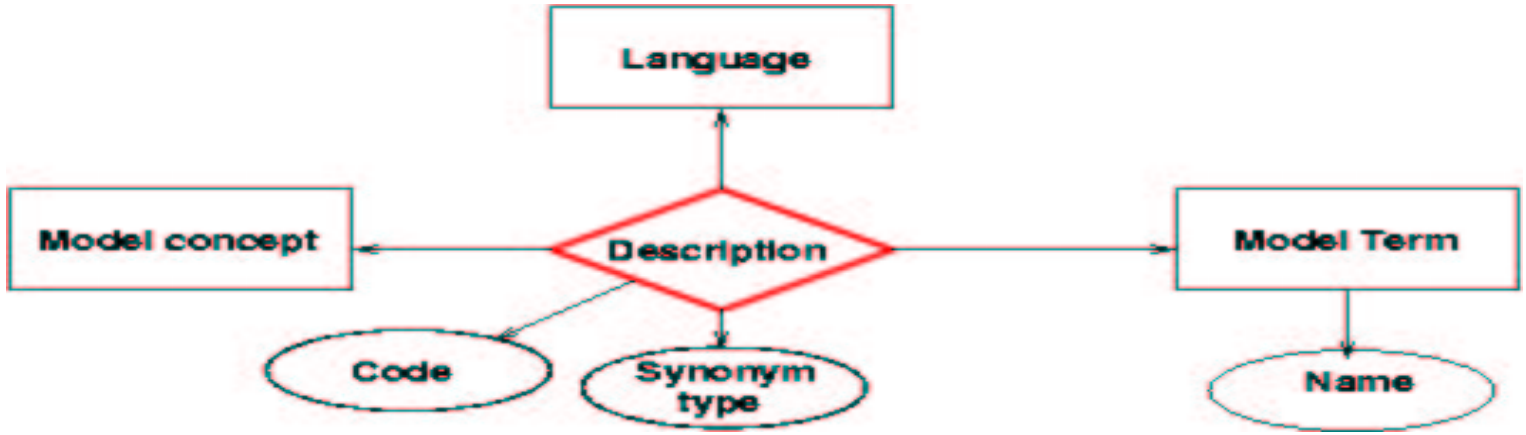
Meta Classes



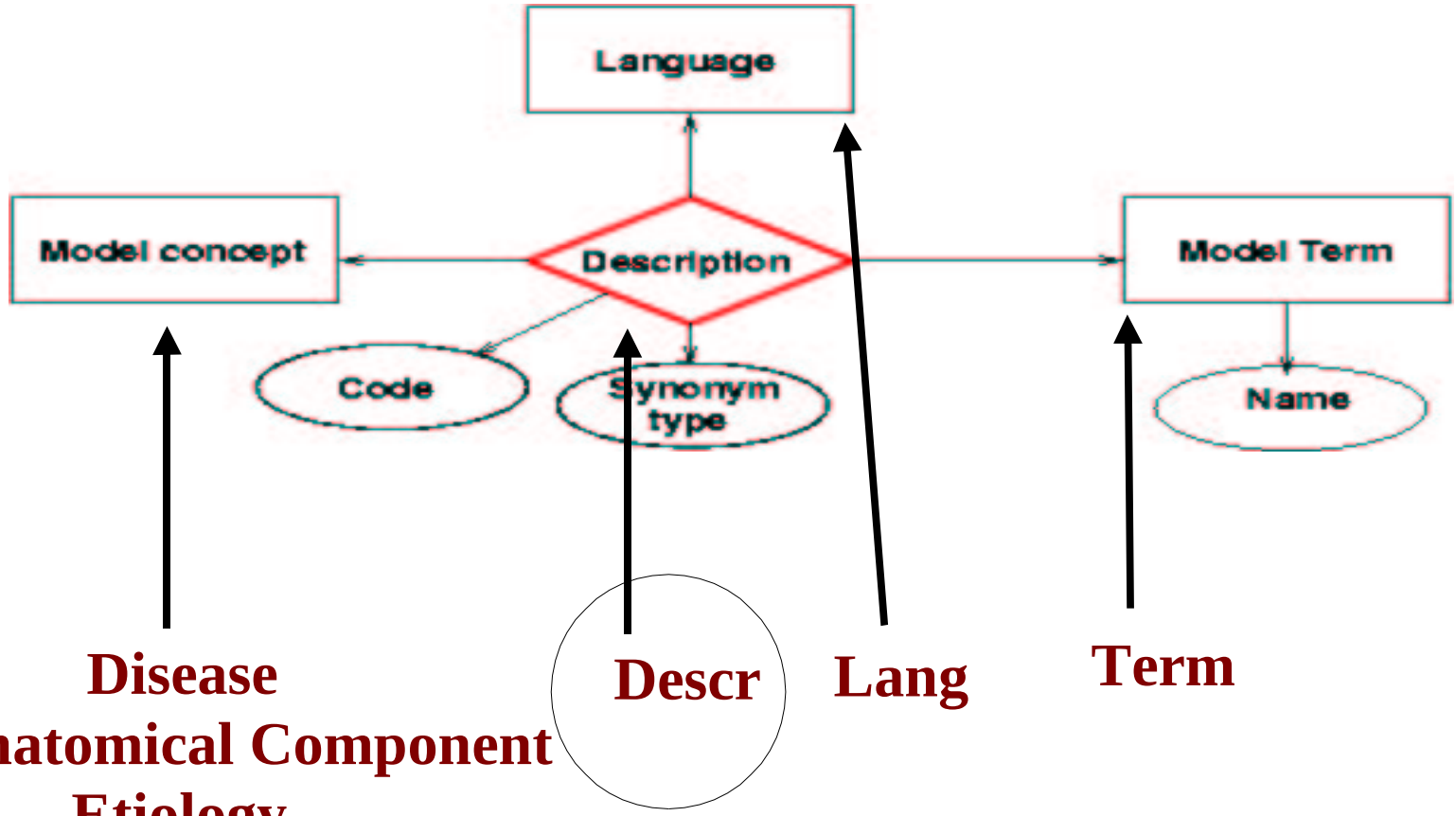
Description

Template Slots

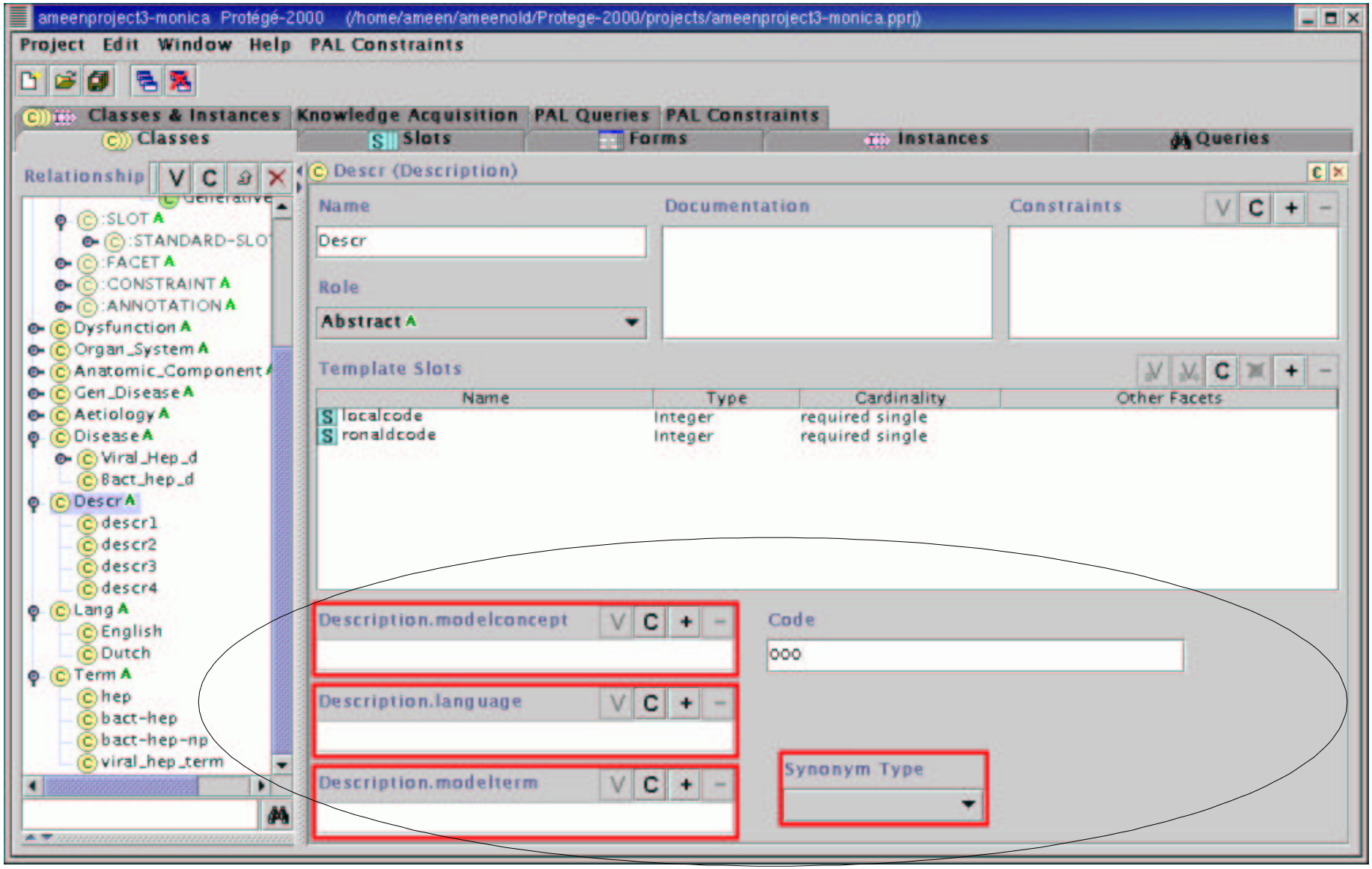
Name	Type	Cardinality	Other Facets
code	String	single	
synonym_type	Symbol	required single	allowed-values={preferred,synonym} ...
description.modelterm I	Instance	required single	classes={Model_Term}
description.language I	Instance	required single	classes={Language}
description.modelconcept I	Instance	required single	classes={Model_Concept}



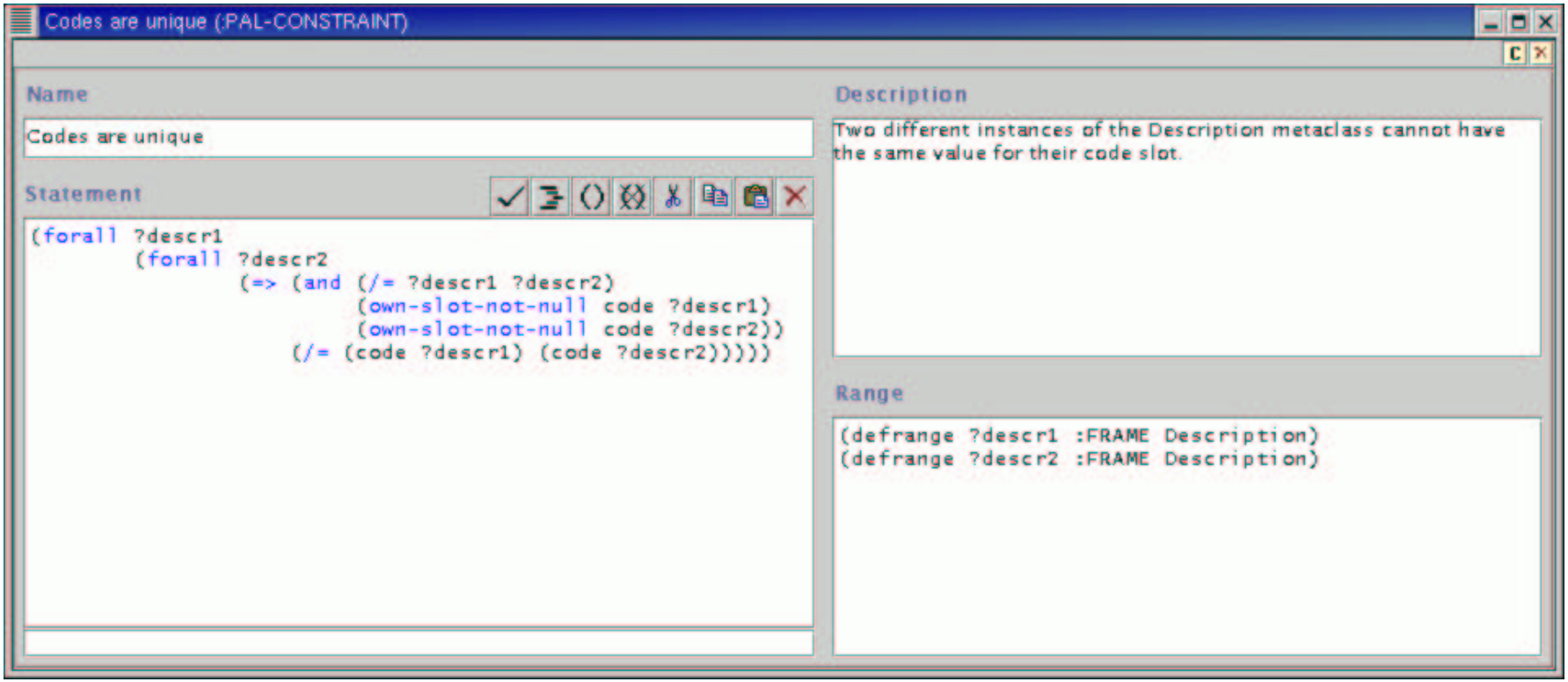
Instance Classes



Instances of Meta Classes



PAL: Codes are unique



PAL: 1 preferred term per concept

```
(forall ?language
    (forall ?description1
        (forall ?description2
            (=> (and (language.description ?language ?description1)
                (language.description ?language ?description2)
                (/= ?description1 ?description2)
                (= (description.modelconcept ?description1)
                  (description.modelconcept ?description2))
                (synonym_type ?description1 (coerce−to−symbol "preferred")))
              (and (not (synonym_type ?description2 (coerce−to−symbol "preferred")))
                (exists ?term1
                    (and (description.modelterm ?description1 ?term1)
                        (modelterm.description ?term1 ?description1)
                        (not (exists ?term2
                            (and (/= ?term2 ?term1)
                                (description.modelterm ?description1 ?term2)
                                (modelterm.description ?term2 ?description1)))))))))))
```

PAL: 1 concept per preferred term

```
(forall ?desc1
      (forall ?desc2
            (=> (and
                  (/= ?desc1 ?desc2)
                  (synonym_type ?desc1 (coerce−to−symbol "preferred"))
                  (synonym_type ?desc2 (coerce−to−symbol "preferred"))
                  (= (description.language ?desc1)
                     (description.language ?desc2))
                  (/= (description.modelconcept ?desc1)
                      (description.modelconcept ?desc2)))

                (/= (description.modelterm ?desc1)
                     (description.modelterm ?desc2)))))
```

Knowledge model in Protege

- Concept oriented
- Explicit Rels
- Composition Rules

- Domain Completeness
- Multiple classification
- Use of definitions
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

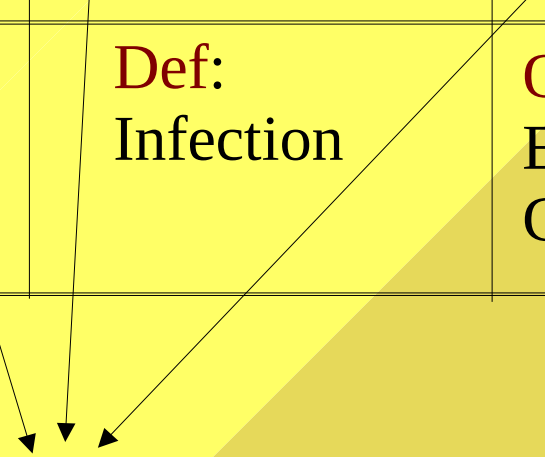
Knowledge model in Protege

- Concept oriented
- Explicit Rels
- Composition Rules

- Domain Completeness
- Multiple classification
- Use of definitions
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

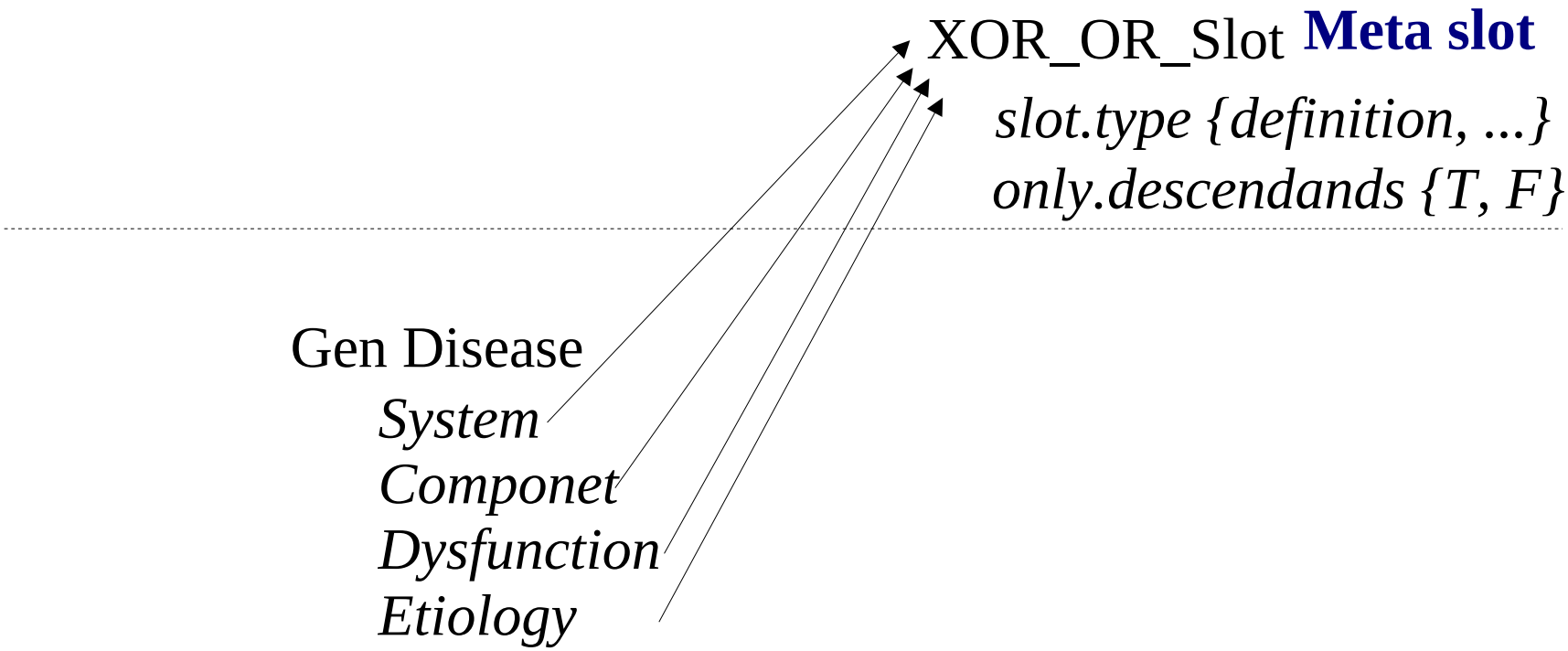
Nomenclature/Defs in DICE

Disease	Anatomical Component	Abnormality	Aetiology
Hepatitis	Def: Liver	Def: Infection	OR: (HepV, E−BV, CytomegaloV)

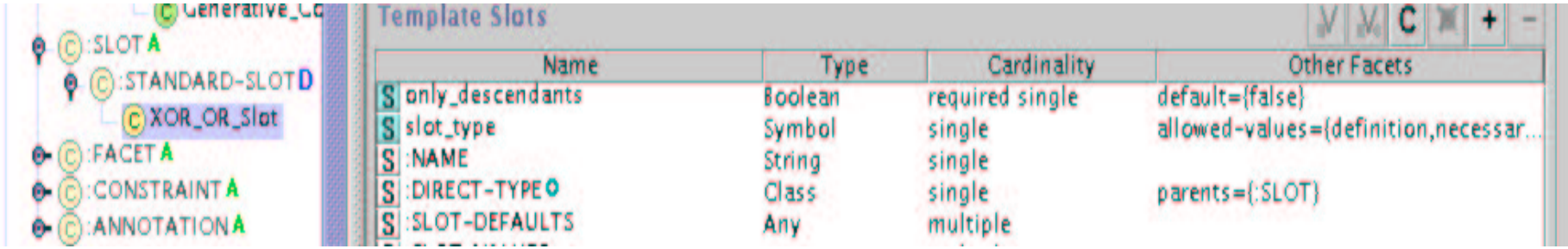


Meta Slots

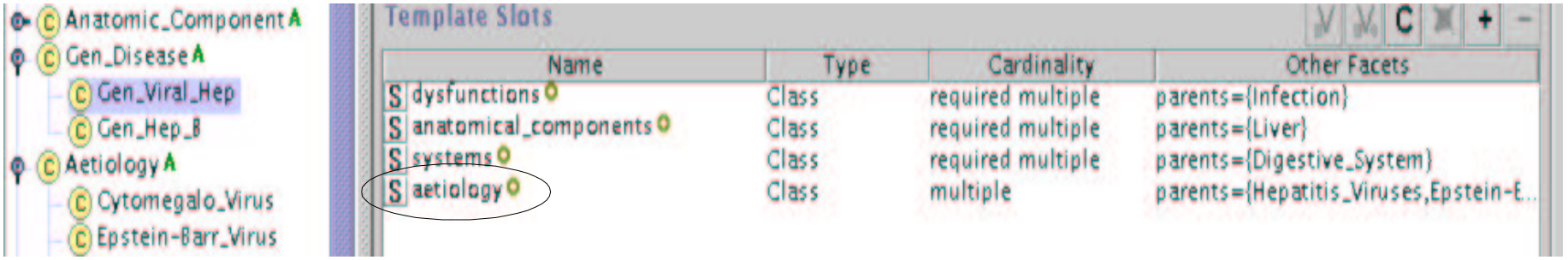
Nomenclature



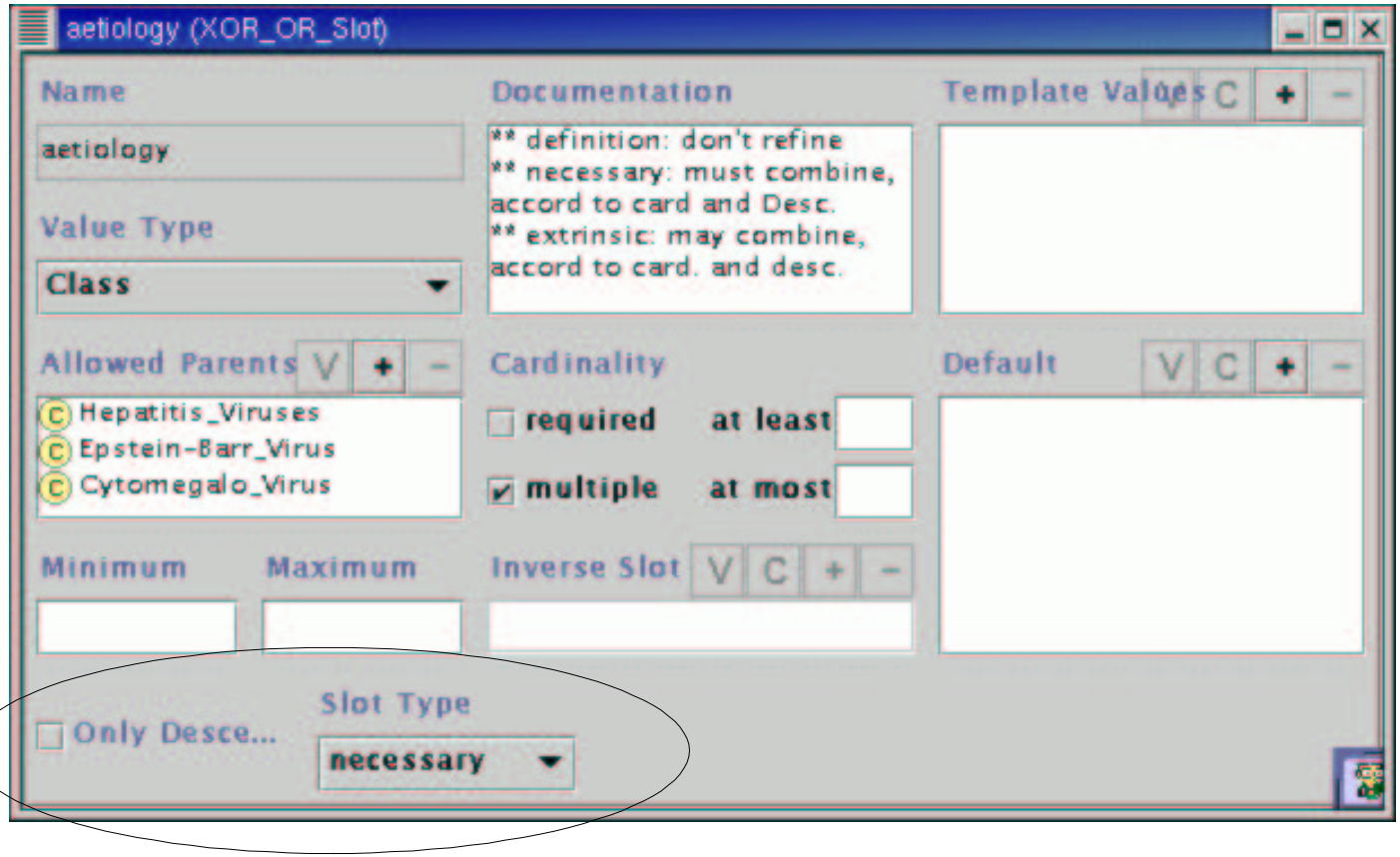
Meta slot: XOR_OR_Slot



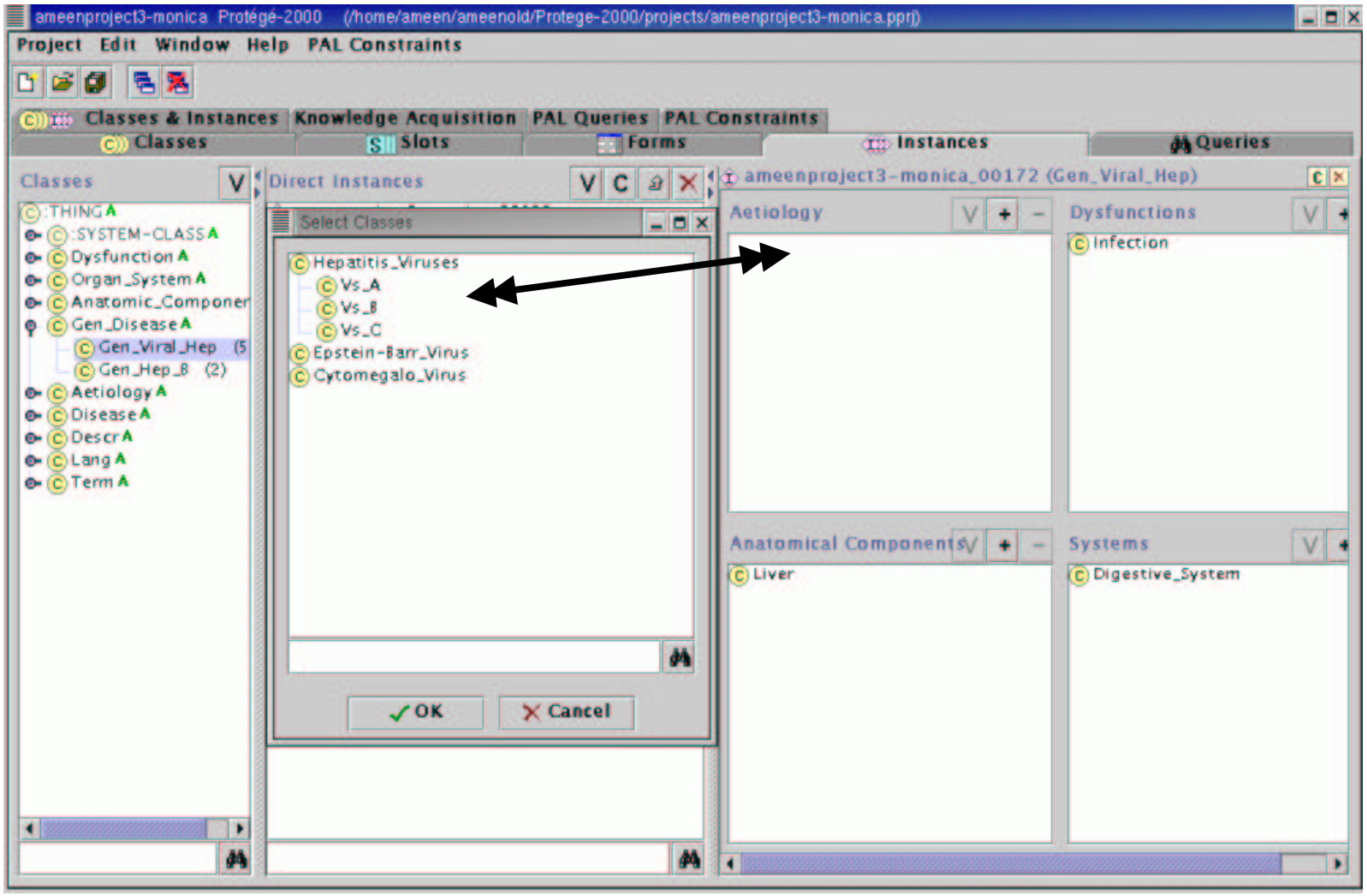
Slot instances used to specify legal combinations



Specification of etiology within combination rule of Viral Hepatitis



Refinement at instantiation of Gen Disease



Knowledge model in Protege

- Concept oriented
- Explicit Rels
- Composition Rules

- Domain Completeness
- Multiple classification
- Use of definitions
- Synonyms
- Multilingual
- Non redundant
- Unique codes
- Non ambiguous
- Non vague
- Context−free codes

Conclusions

- **Knowledge model** of Protege forms a good basis for terminology systems.
- Meta model of Concept−Language−Term can be naturally specified by **Meta Classes** and enforced by **PAL constraints**.
- Nomenclature can be specified by **Meta Slots** but is not enforced *within* Protege.