An Intuitive Graphical Query Interface for Protégé Knowledge Bases

Landon Todd Detwiler,
Cornelius Rosse, Linda Shapiro

Structural Informatics Group
University of Washington

Emily

- Query interface for Protégé KBs
  - Graphical, point-and-click
  - Stand-alone Java application
  - Built on the Protégé API library
  - Class-to-class relationships
- Foundational Model of Anatomy (FMA)
- Goal: Usable by novice users with little instruction

FMA

- Foundational Model of Anatomy
- Symbolic model of the physical organization of the human body
- Goals
  - Foundation for interoperability between biological domains
  - Basis for the construction of intelligent biomedical applications

FMA Complications

- Space complexity
  - > 70,000 classes
  - > 185,000 frames
  - > 170 slots in use
  - > 1.5 million slot values
- Classes without the instances (individuals)
- Existing tools and the FMA

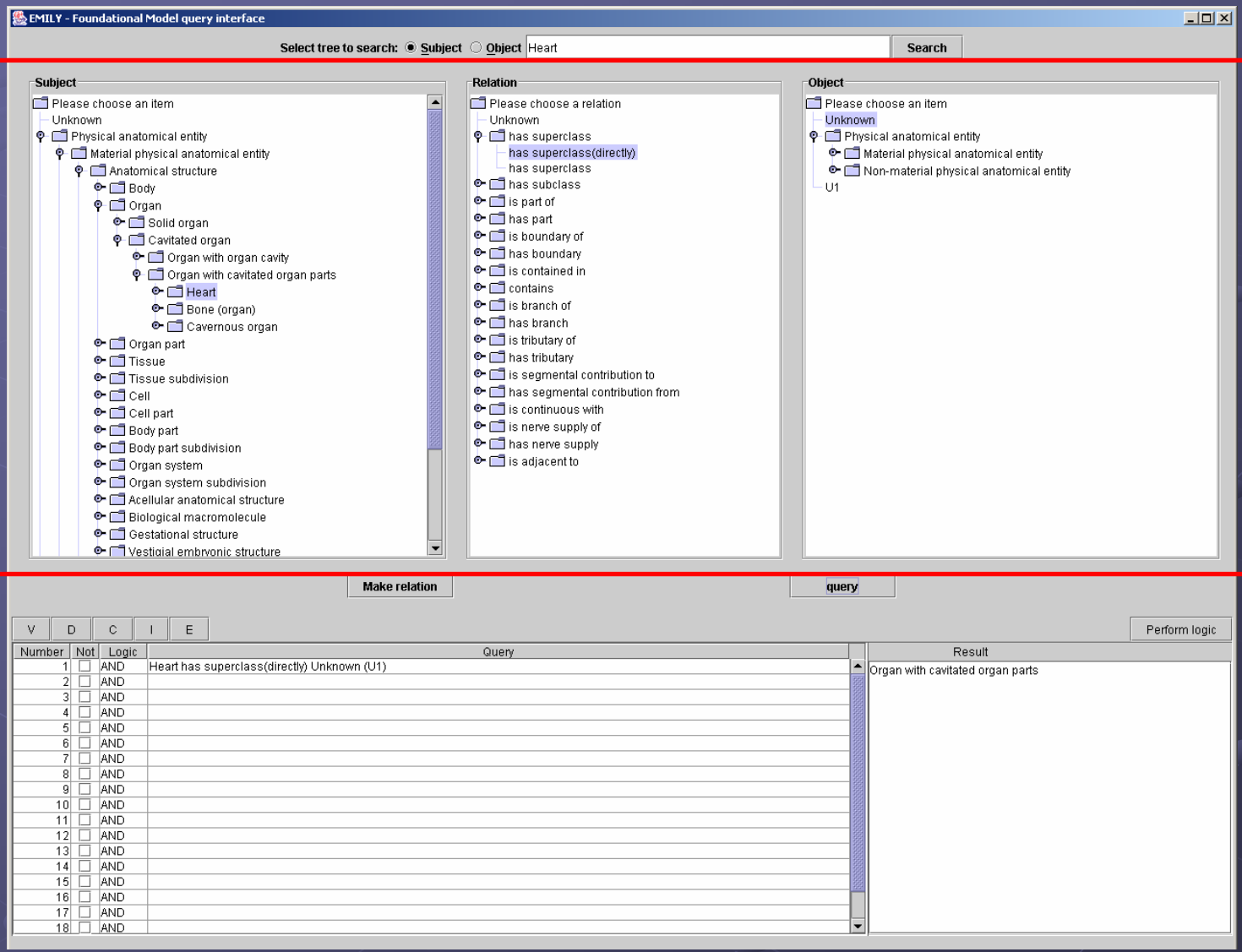
Example Questions

- What is the heart?
- What are the parts of the heart?
- Is the epicardium a part of the heart?
- Which part of the heart is continuous with the superior pulmonary vein?
- What is the relationship between the right atrium and blood?

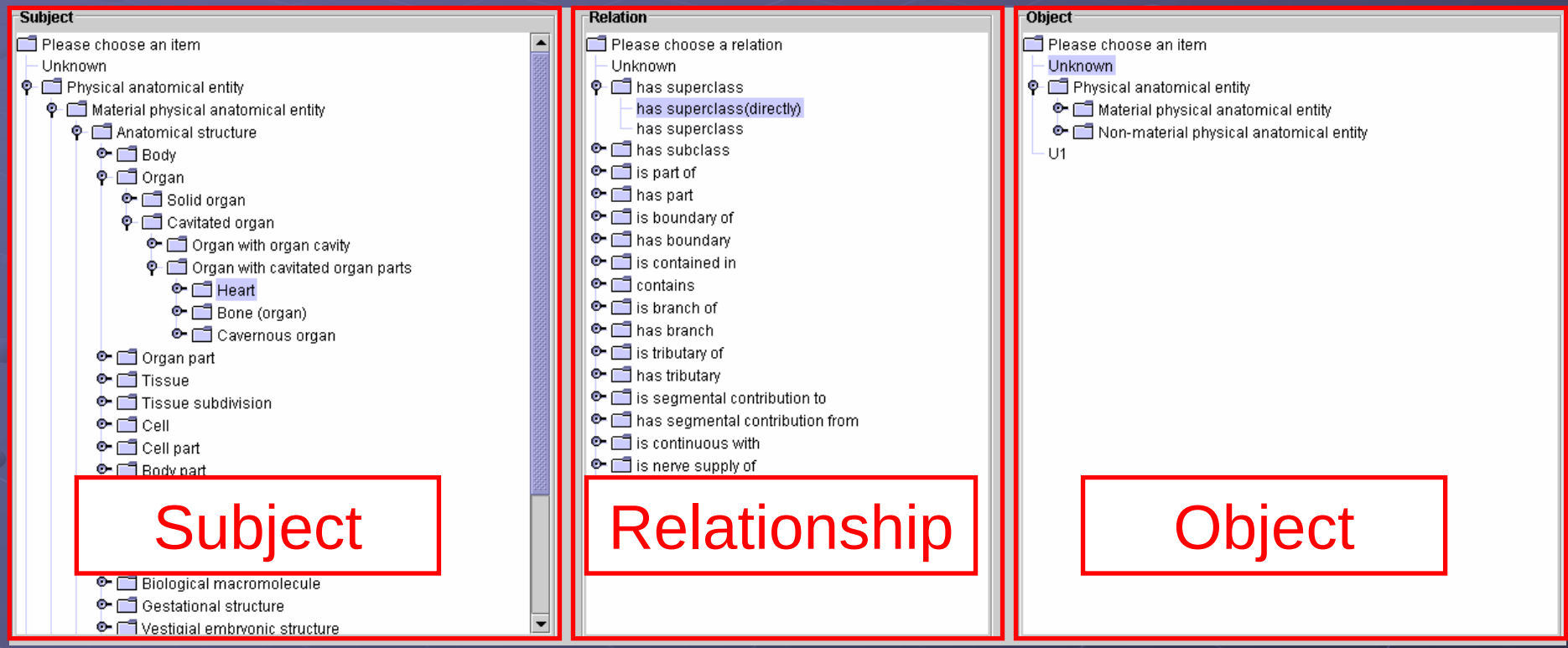
Example Questions

- What is the heart?
- What are the parts of the heart?
- Is the epicardium a part of the heart?
- Which part of the heart is continuous with the superior pulmonary vein?
- What is the relationship between the right atrium and blood?

Composing a Simple Query



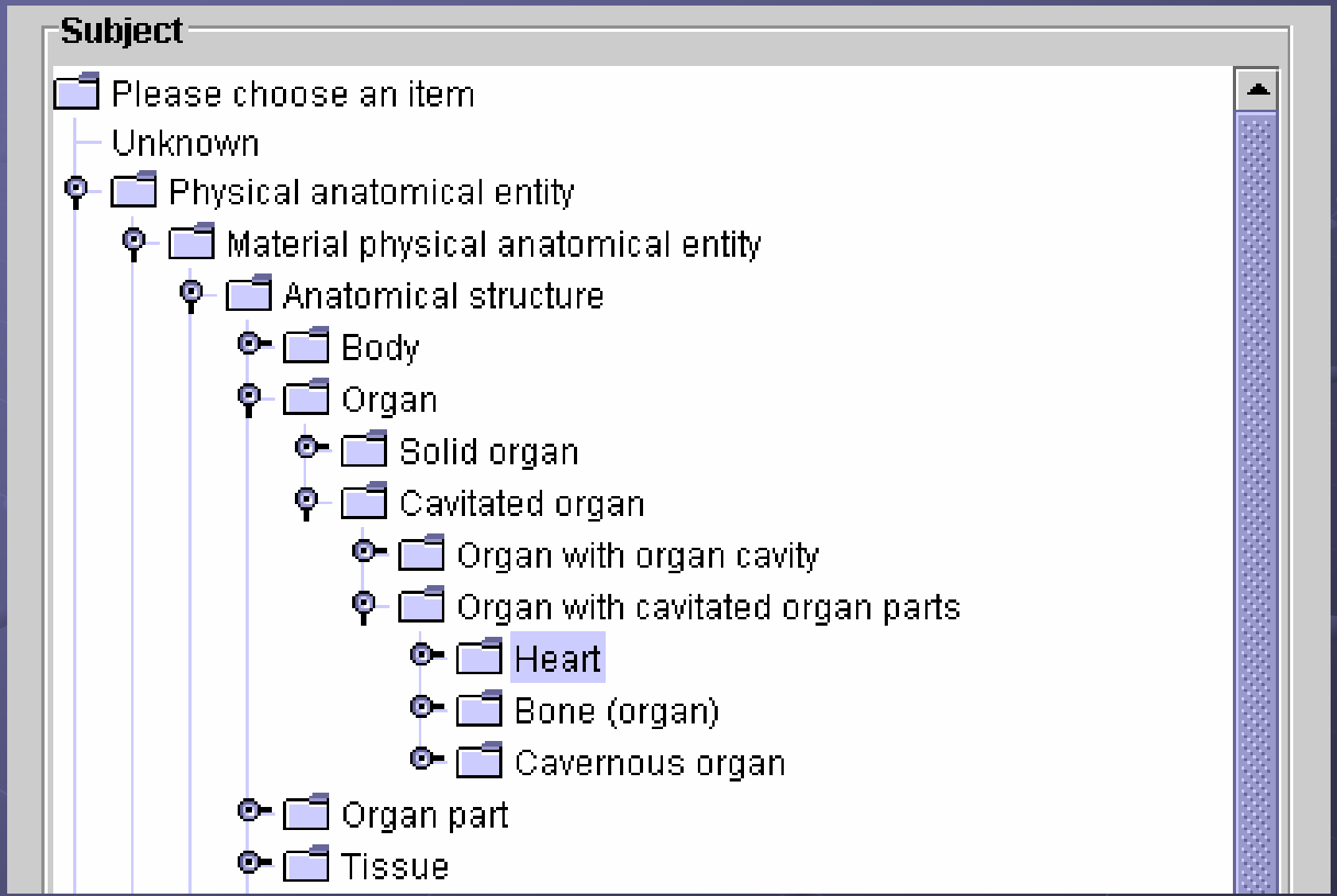
Composing a Simple Query



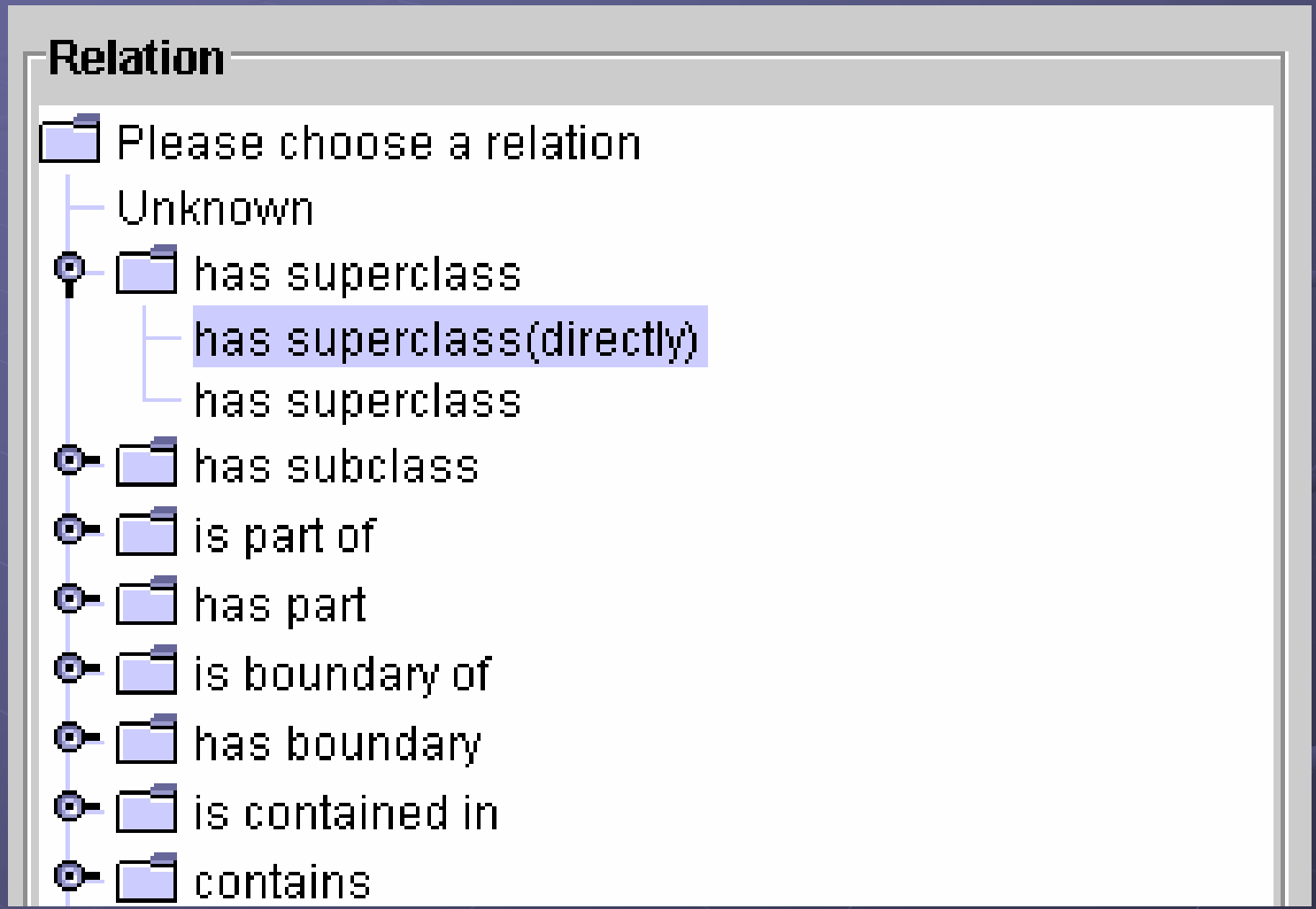
What is the heart?

Subject	Relationship	Object
Heart	has superclass (directly)	Unknown

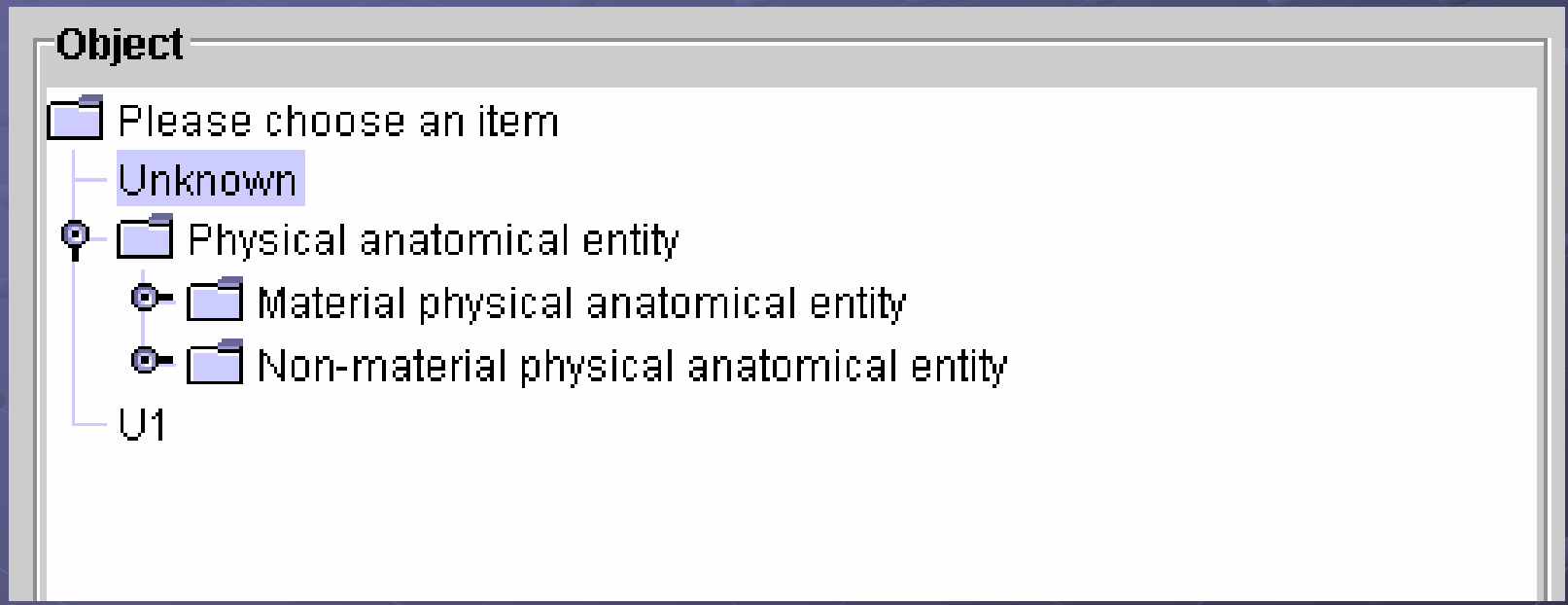
Subject Tree



Relationship Tree



Object Tree


Object
Please choose an item
Unknown
Physical anatomical entity
Material physical anatomical entity
Non-material physical anatomical entity
U1

What is the heart?

Subject	Relationship	Object
Heart	has superclass (directly)	Unknown

Results
Organ with cavitated organ parts

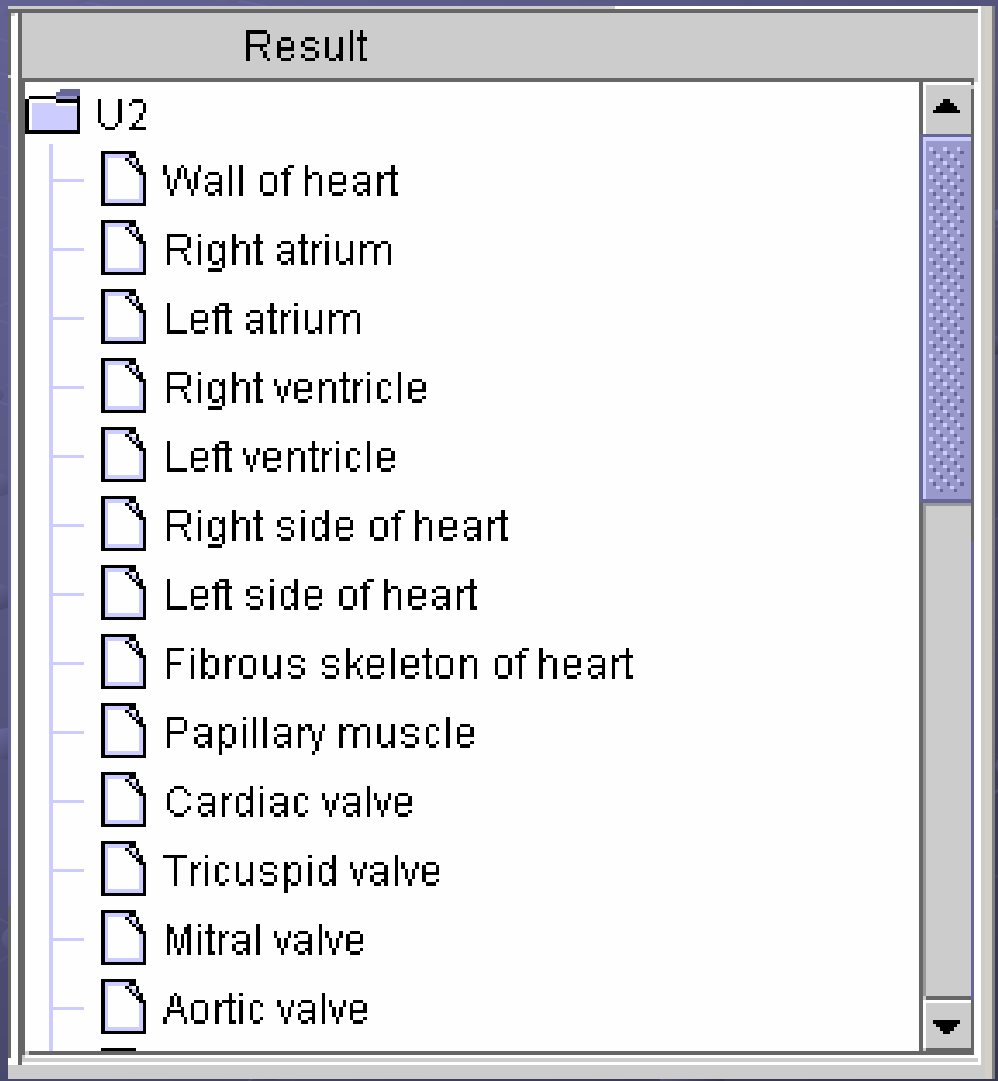
Example Questions

- What is the heart?
- What are the parts of the heart?
- Is the epicardium a part of the heart?
- Which part of the heart is continuous with the superior pulmonary vein?
- What is the relationship between the right atrium and blood?

What are the parts of the heart?

Subject	Relationship	Object
Heart	has part (directly)	Unknown

What are the parts of the heart?



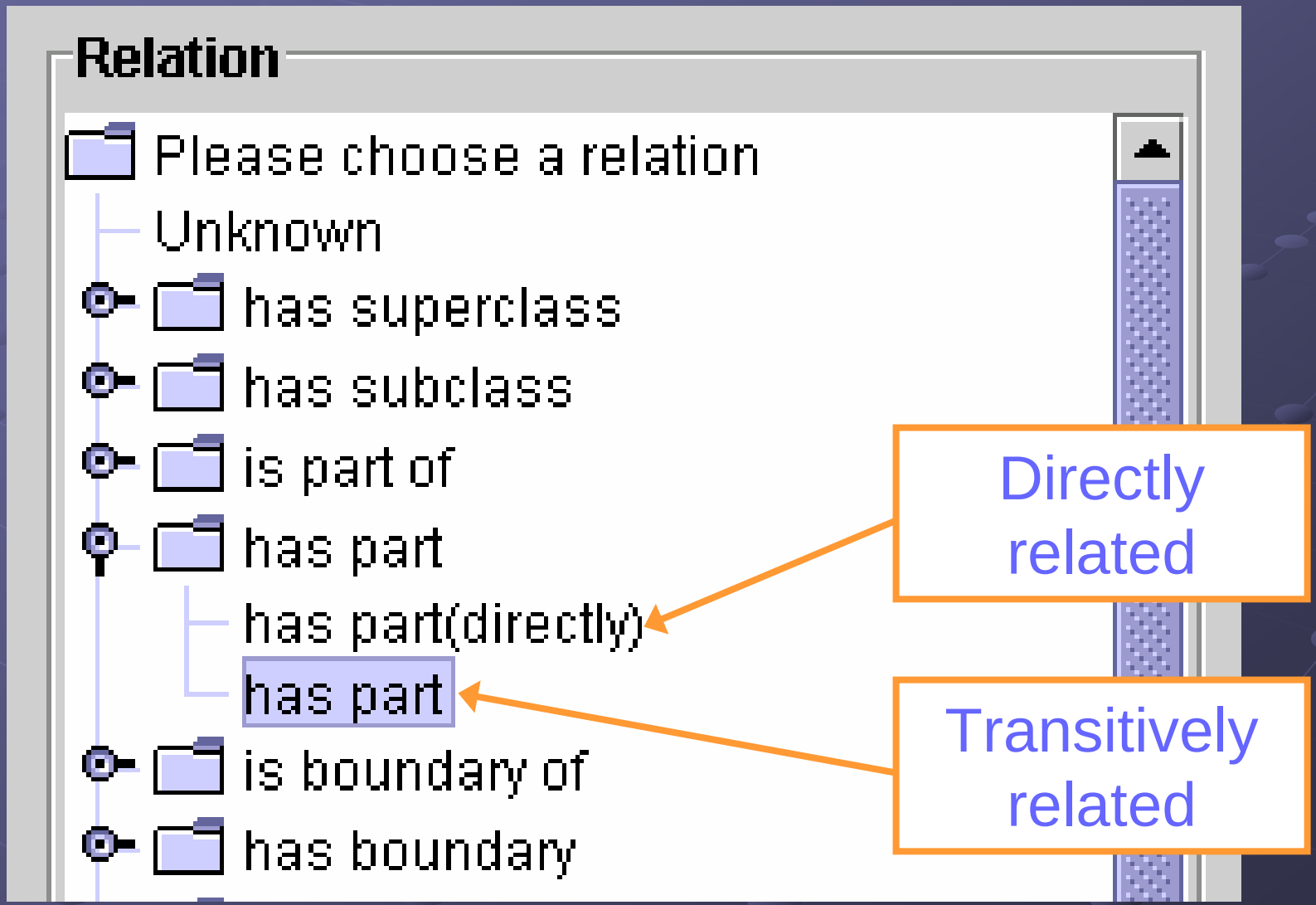
Example Questions

- What is the heart?
- What are the parts of the heart?
- Is the epicardium a part of the heart?
- Which part of the heart is continuous with the superior pulmonary vein?
- What is the relationship between the right atrium and blood?

Is the epicardium a part of the heart?

- Why was Epicardium not in the previous result set?
  - Epicardium not a direct part of heart
  - Heart has part Wall of heart which has part Epicardium
- But aren't the parts of any direct part of the Heart still a part of the Heart?
  - Yes, because part is transitive

Relationship Tree



Is the epicardium a part of the heart?

Subject	Relationship	Object
Heart	has part	Epicardium

Results
Yes

Example Questions

- What is the heart?
- What are the parts of the heart?
- Is the epicardium a part of the heart?
- Which part of the heart is continuous with the superior pulmonary vein?
- What is the relationship between the right atrium and blood?

Which part of the heart is continuous with the superior pulmonary vein?

- What is the Subject?
- What is the Relationship?
- What is the Object?

Compound queries

- Two methods for creating compound queries
  - Query chaining
  - Set operations

Which part of the heart is continuous with the superior pulmonary vein?

- Subject: part of the heart
- Relationship: is continuous with
- Object: Superior pulmonary vein

Which part of the heart is continuous with the superior pulmonary vein?

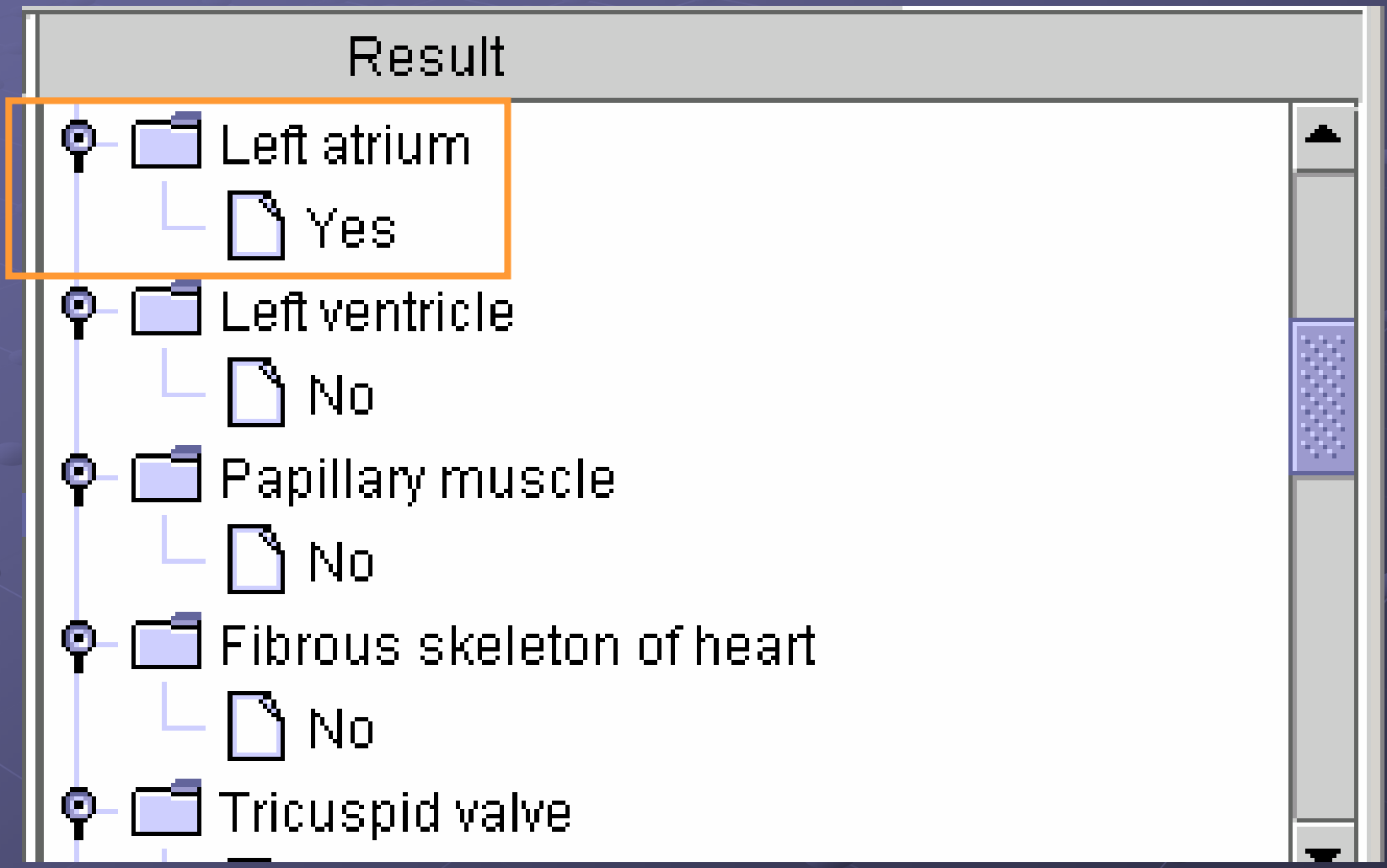
- Subject: part of the heart
  - Subject: Heart
  - Relationship: has part
  - Object: Unknown

  } U1
- Relationship: is continuous with
- Object: Superior pulmonary vein

Which part of the heart is continuous with the superior pulmonary vein?

- Subject: U1
- Relationship: is continuous with
- Object: Superior pulmonary vein

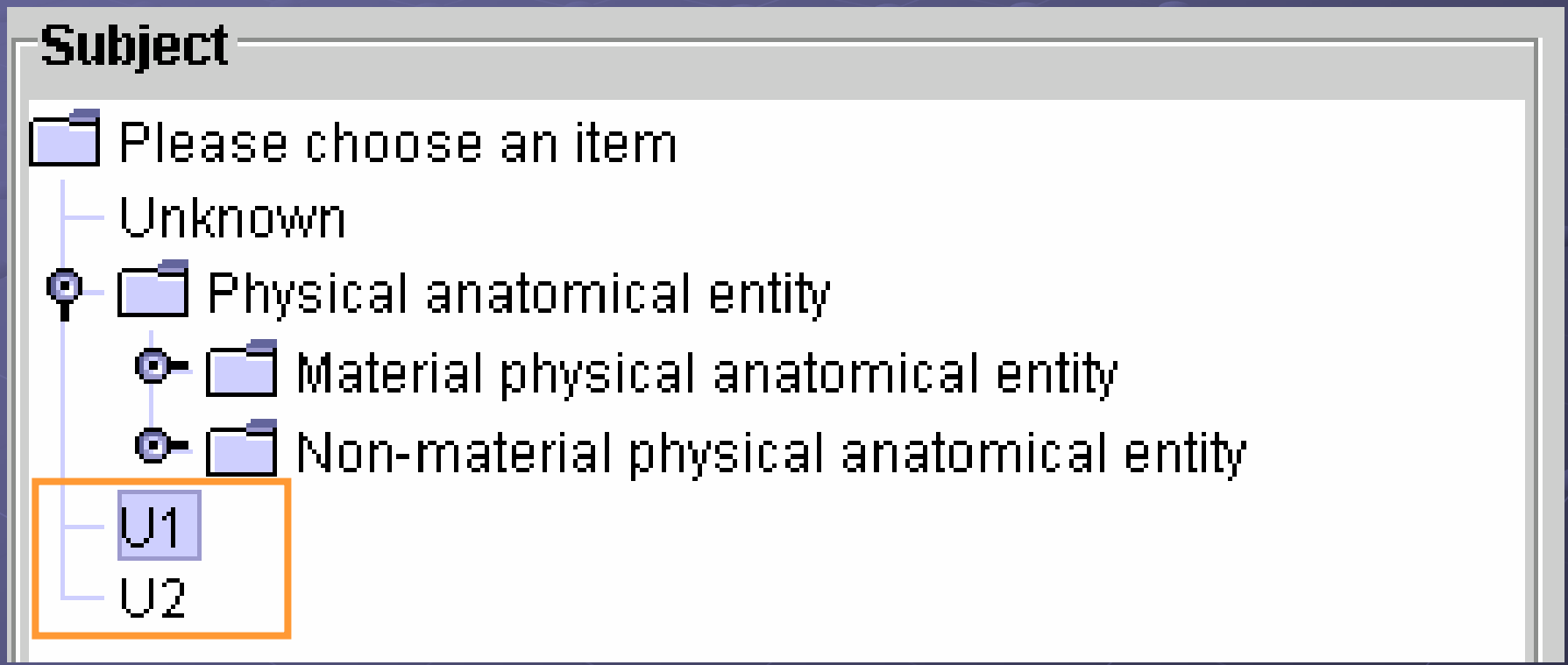
Which part of the heart is continuous with the superior pulmonary vein?



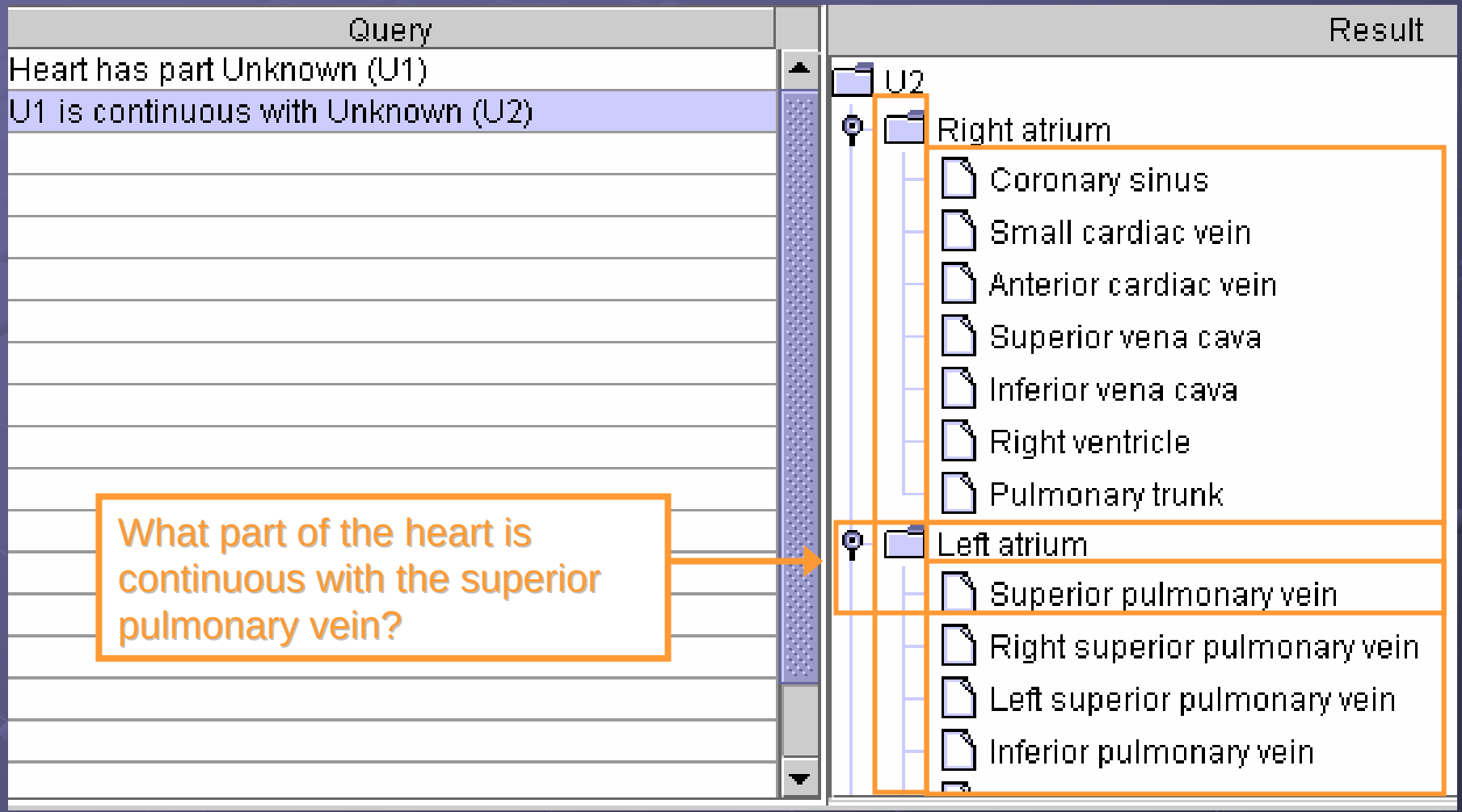
Which part of the heart is continuous with the superior pulmonary vein?

Subject	Relationship	Object
Heart	has part	Unknown (U1)
U1	is continuous with	Unknown (U2)

Result Sets


Subject
Please choose an item
Unknown
Physical anatomical entity
Material physical anatomical entity
Non-material physical anatomical entity
U1
U2

Query History Panel



Compound queries

- Two methods for creating compound queries
  - Query chaining
  - Set operations

Set Operations

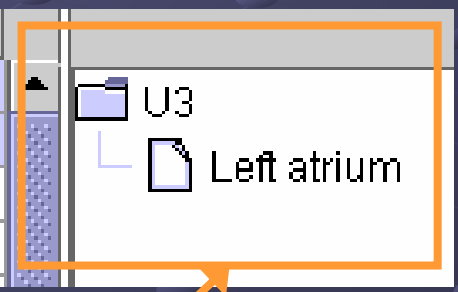
- Results of some queries are sets
- Available operations
  - AND
  - OR
  - NOT

Which part of the heart is continuous with the superior pulmonary vein?

Subject	Relationship	Object
Heart	has part	Unknown (U1)
Unknown (U2)	is continuous with	Superior pulmonary vein
U1	AND	U2

Set Operations

Not	Logic	Query
☐	AND	Heart has part Unknown (U1)
☐	AND	Unknown (U2) is continuous with(directly) Superior pulmon...
☐	AND	U3 = [(U1)] AND [(U2)]
☐	AND	



What part of the heart is continuous with the superior pulmonary vein?

Example Questions

- What is the heart?
- What are the parts of the heart?
- Is the epicardium a part of the heart?
- Which part of the heart is continuous with the superior pulmonary vein?
- What is the relationship between the right atrium and blood?

Unknown Relationship

- What if the user chooses the ‘Unknown’ node for the query relationship?
  - FMA is highly connected
    - Many paths between classes
    - Not practical (nor useful) to find all paths
  - First path
    - Assumption: The simplest path is likely to be the most desirable

Unknown Relationship

- Algorithm
  - Direct connections
  - Transitive closure connections
  - Predetermined connection types
  - Depth-limited breadth-first search

What is the relationship between the right atrium and blood

Subject	Relationship	Object
Right atrium	Unknown	Blood

Results
Right atrium part Cavity of right atrium contains Blood in right atrium :DIRECT-SUPERCLASSES Blood

Issues

- Translating into S-R-O triple
- Translating into FMA terms
- Attributed relationships (e.g. right lateral adjacency)

Conclusion

- Simple, point-and-click query construction
- Query types
  - Simple
  - Compound
- “Real world” limits
  - Triple formulation
  - Knowledge base content

Future Work

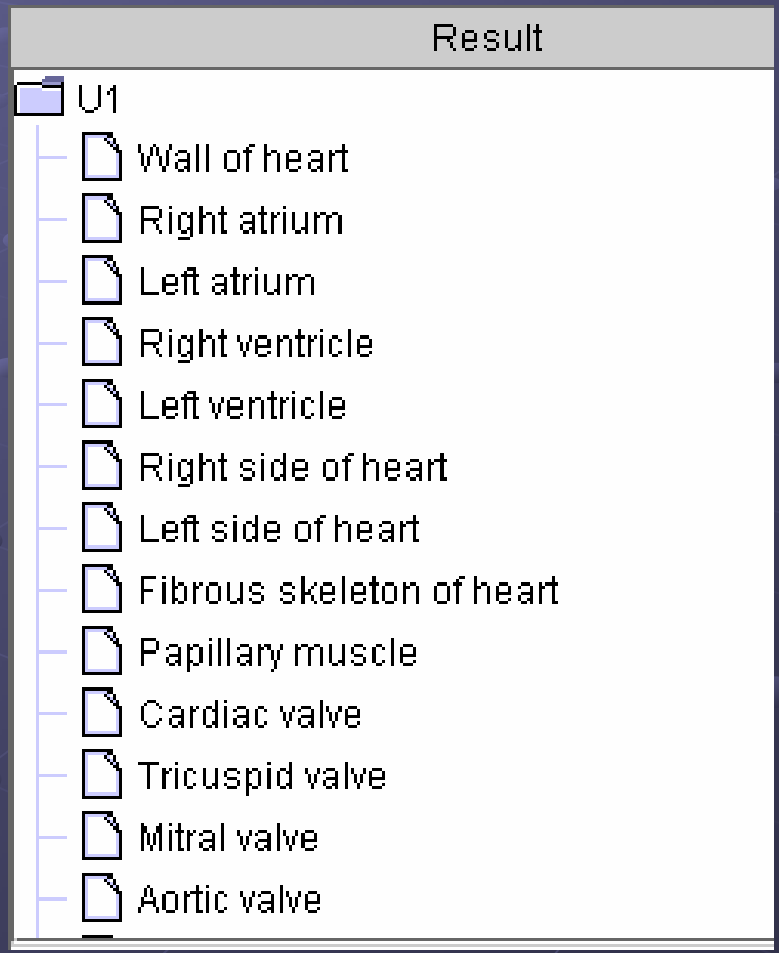
- Web accessibility
- Improved transitive closure performance
- Attributed relationship handling
- Hierarchical transitive closures
- Relationship synonyms

Simple Queries

- Subject – Relationship – Object triplets
  - Basic English sentence structure
  - Choose an icon from each field
  - At most 1 may be “Unknown”

Query chaining

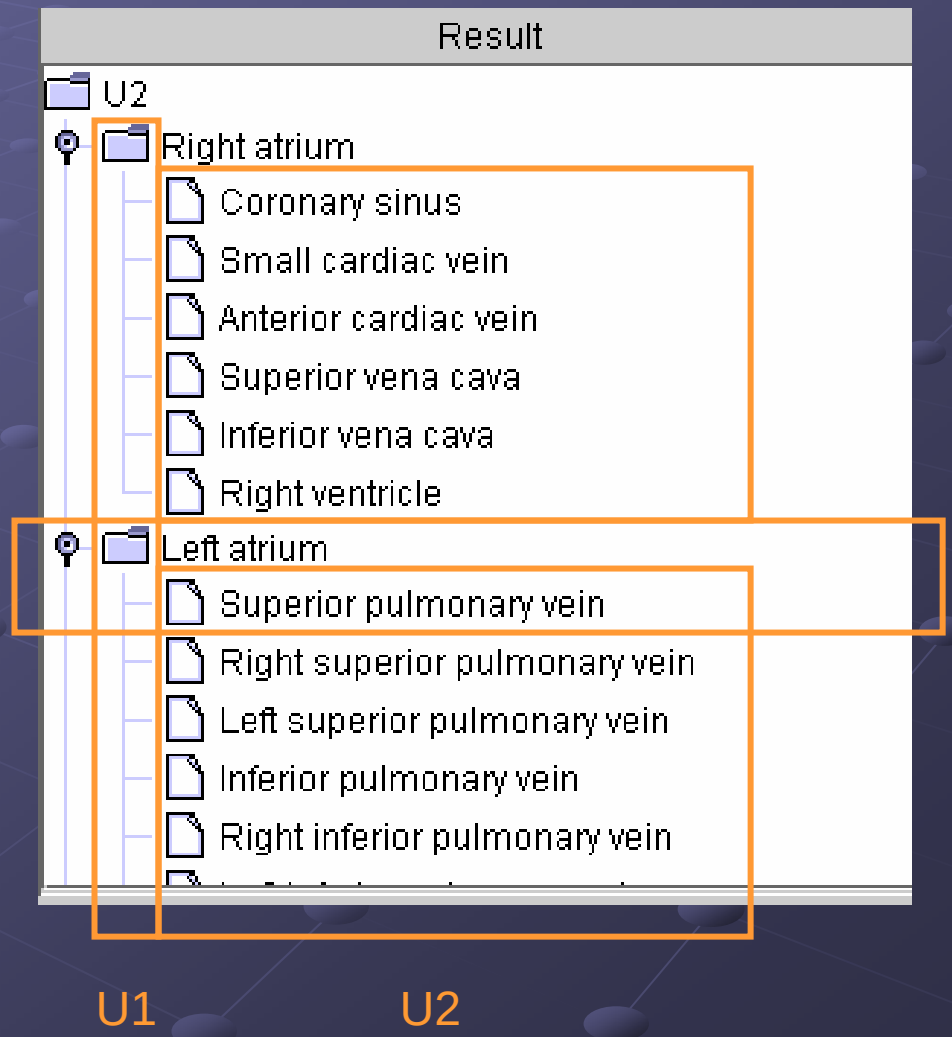
Heart has part Unknown (U1)



Query chaining

U1 is continuous with Unknown (U2)

What part of the heart is continuous with the superior pulmonary vein?



Which part of the heart is continuous with the superior pulmonary vein?

- Subject: U1
- Relationship: is continuous with
- Object: Unknown