Ontology Alignment

state of the art and
an application in literature search

Patrick Lambrix
Linköpings universitet

Ontologies

“Ontologies define the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary.”

(Neches, Fikes, Finin, Gruber, Senator, Swartout, 1991)

Example

GENE ONTOLOGY (GO)

- immune response
    - **i-** acute-phase response
    - **i-** anaphylaxis
    - **i-** antigen presentation
    - **i-** antigen processing
    - **i-** cellular defense response
    - **i-** cytokine metabolism
        - **i-** cytokine biosynthesis synonym cytokine production
            - …
        - **p-** regulation of cytokine biosynthesis
        - …
    - …
    - **i-** B-cell activation
        - **i-** B-cell differentiation
        - **i-** B-cell proliferation
    - **i-** cellular defense response
    - …
    - **i-** T-cell activation
        - **i-** activation of natural killer cell activity
            - …

Ontologies used …

- for communication between people and organizations
- for enabling knowledge reuse and sharing
- as basis for interoperability between systems
- as repository of information
- as query model for information sources

Key technology for the Semantic Web

Biomedical Ontologies - efforts

OBO – Open Biomedical Ontologies
http://www.obofoundry.org/
(over 50 ontologies)

” The mission of OBO is to support community members who are developing and publishing ontologies in the biomedical domain. It is our vision that a core of these ontologies will be **fully interoperable**, by virtue of a common design philosophy and implementation, thereby enabling scientists and their instruments to **communicate with minimum ambiguity**. In this way the data generated in the course of biomedical research will form a single, consistent, cumulatively expanding, and algorithmically tractable whole. This core will be known as the "OBO Foundry". .”

OBO Foundry

1. open and available
2. common shared syntax
3. unique identifier space
4. procedures for identifying distinct successive versions
5. clearly specified and clearly delineated content
6. textual definitions for all terms
7. use relations from OBO Relation Ontology
8. well documented
9. plurality of independent users
10. developed collaboratively with other OBO Foundry members

Biomedical Ontologies - efforts

National Center for Biomedical Ontology
http://bioontology.org/index.html

Funded by National Institutes of Health

”The goal of the Center is to support biomedical researchers in their knowledge-intensive work, by providing online **tools and a Web portal enabling them to access, review, and integrate disparate ontological resources** in all aspects of biomedical investigation and clinical practice. A major focus of our work involves the use of biomedical ontologies to aid in the management and analysis of data derived from complex experiments.”

Systems Biology Ontologies - efforts

- Systems Biology Ontology
- Proteomics Standard Initiative for Molecular Interaction
- BioPAX

Ontology Alignment

- Ontology alignment
- Ontology alignment strategies
- Evaluation of ontology alignment strategies
- Current issues
- Ontology-based literature search

Ontologies in biomedical research

- many biomedical ontologies

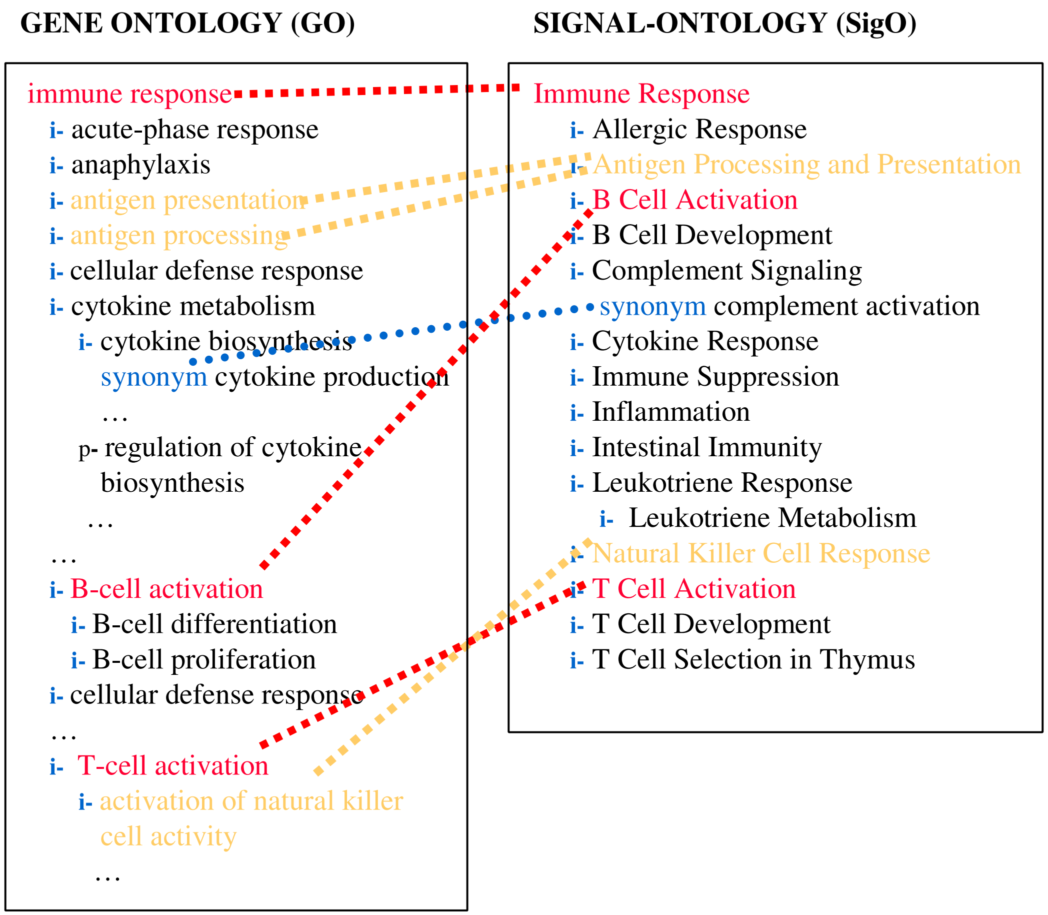
- practical use of biomedical ontologies

  e.g. databases annotated with GO

GENE ONTOLOGY (GO)

```
immune response
  i- acute-phase response
  i- anaphylaxis
  i- antigen presentation
  i- antigen processing
  i- cellular defense response
  i- cytokine metabolism
    i- cytokine biosynthesis
      synonym cytokine production
      …
    p- regulation of cytokine
      biosynthesis
    …
  …
  i- B-cell activation
    i- B-cell differentiation
    i- B-cell proliferation
  i- cellular defense response
  …
  i-  T-cell activation
    i- activation of natural killer
      cell activity
     …
```

Ontologies with overlapping information



GENE ONTOLOGY (GO)

immune response
- i- acute-phase response
- i- anaphylaxis
- i- antigen presentation
- i- antigen processing
- i- cellular defense response
- i- cytokine metabolism
  - i- cytokine biosynthesis
    synonym cytokine production
  - …
  - p- regulation of cytokine biosynthesis
  - …
- …
- i- B-cell activation
  - i- B-cell differentiation
  - i- B-cell proliferation
- i- cellular defense response
- …
- i- T-cell activation
  - i- activation of natural killer cell activity
  - …

SIGNAL-ONTOLOGY (SigO)

Immune Response
- i- Allergic Response
- i- Antigen Processing and Presentation
- i- B Cell Activation
- i- B Cell Development
- i- Complement Signaling
  synonym complement activation
- i- Cytokine Response
- i- Immune Suppression
- i- Inflammation
- i- Intestinal Immunity
- i- Leukotriene Response
  - i- Leukotriene Metabolism
- i- Natural Killer Cell Response
- i- T Cell Activation
- i- T Cell Development
- i- T Cell Selection in Thymus

Ontologies with overlapping information

- Use of multiple ontologies

  e.g. custom-specific ontology + standard ontology

  different views on same domain

  connecting related areas

- Bottom-up creation of ontologies

  experts can focus on their domain of expertise

important to know the inter-ontology relationships

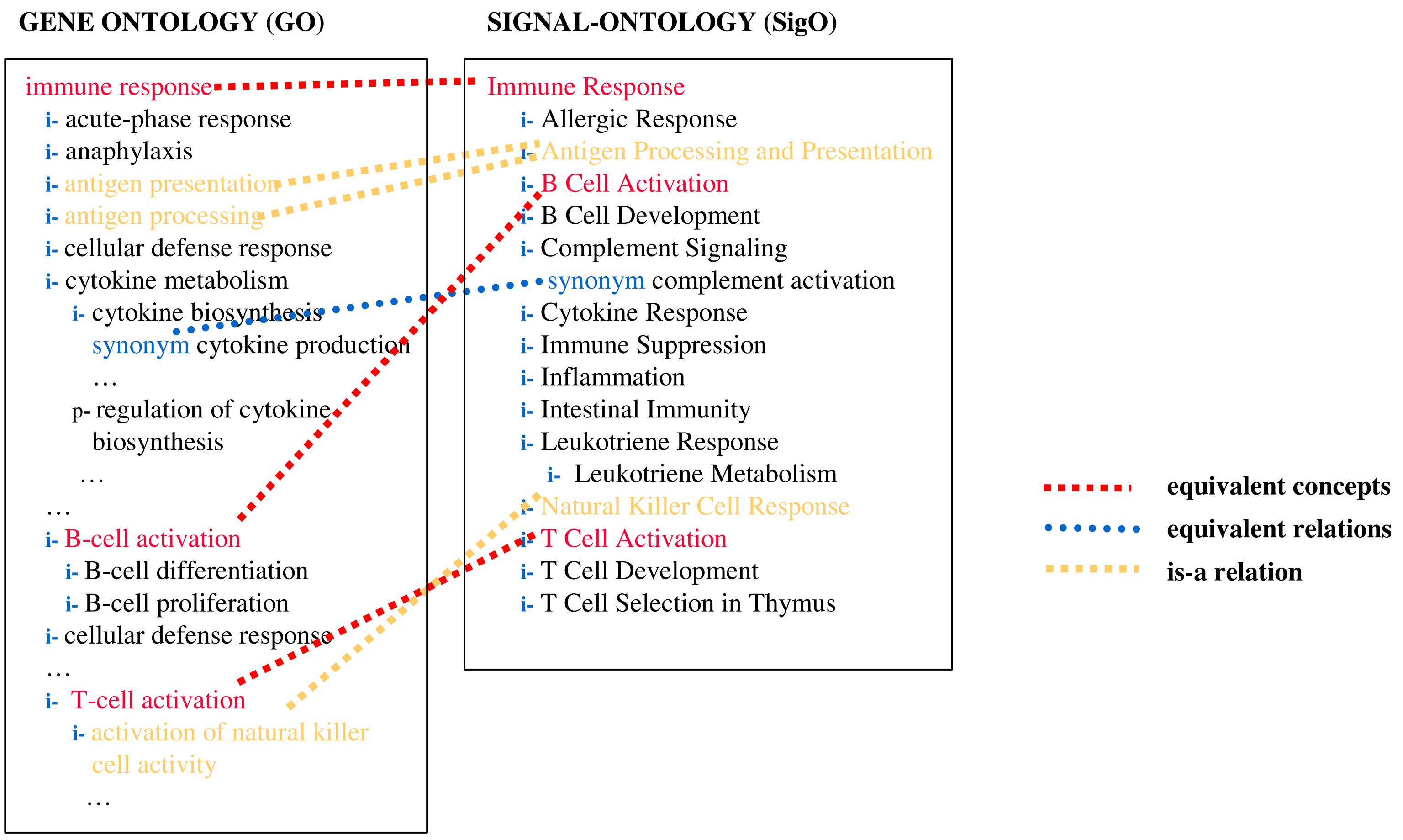
GENE ONTOLOGY (GO)

- immune response
  - **i-** acute-phase response
  - **i-** anaphylaxis
  - **i-** antigen presentation
  - **i-** antigen processing
  - **i-** cellular defense response
  - **i-** cytokine metabolism
    - **i-** cytokine biosynthesis
      synonym cytokine production
      …
    - p- regulation of cytokine biosynthesis
    - …
  - …
  - **i-** B-cell activation
    - **i-** B-cell differentiation
    - **i-** B-cell proliferation
  - **i-** cellular defense response
  - …
  - **i-** T-cell activation
    - **i-** activation of natural killer cell activity
    - …

SIGNAL-ONTOLOGY (SigO)

- Immune Response
  - **i-** Allergic Response
  - **i-** Antigen Processing and Presentation
  - **i-** B Cell Activation
  - **i-** B Cell Development
  - **i-** Complement Signaling
    synonym complement activation
  - **i-** Cytokine Response
  - **i-** Immune Suppression
  - **i-** Inflammation
  - **i-** Intestinal Immunity
  - **i-** Leukotriene Response
    - **i-** Leukotriene Metabolism
  - **i-** Natural Killer Cell Response
  - **i-** T Cell Activation
  - **i-** T Cell Development
  - **i-** T Cell Selection in Thymus

Ontology Alignment

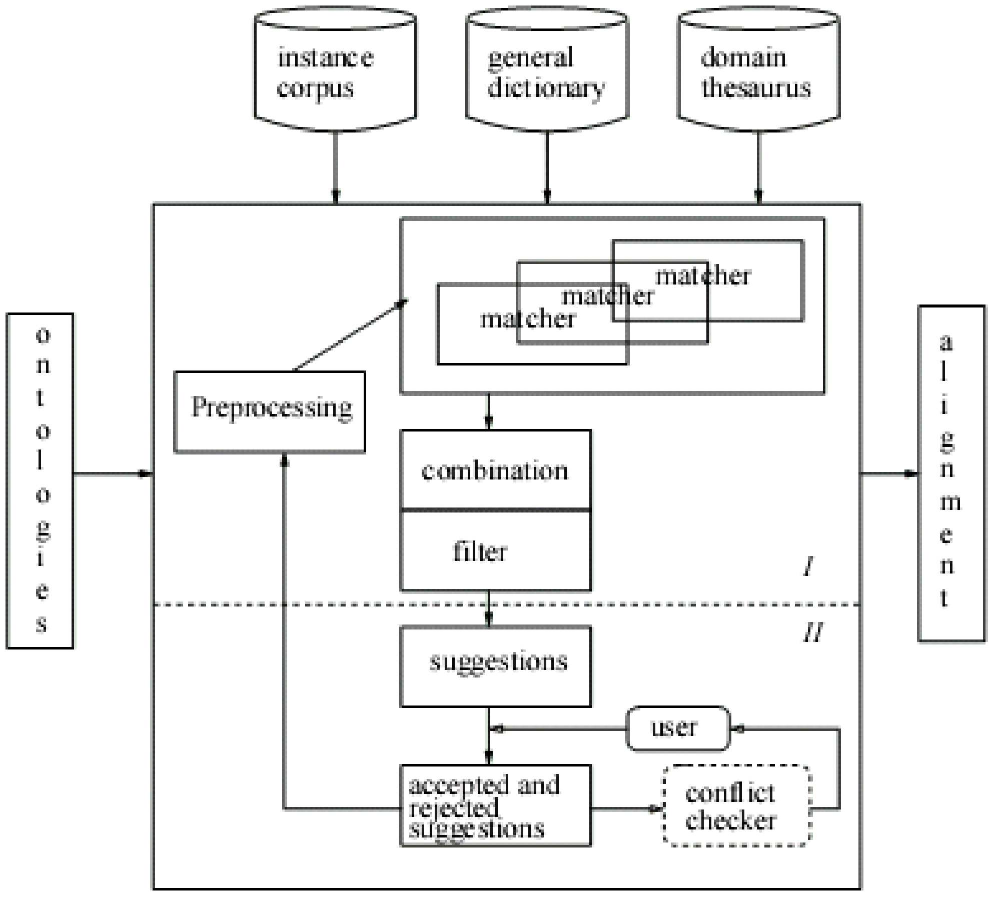


Defining the relations between the terms in different ontologies

Ontology Alignment

- Ontology alignment
- Ontology alignment strategies
- Evaluation of ontology alignment strategies
- Current issues
- Ontology-based literature search

An Alignment Framework



Preprocessing

Preprocessing

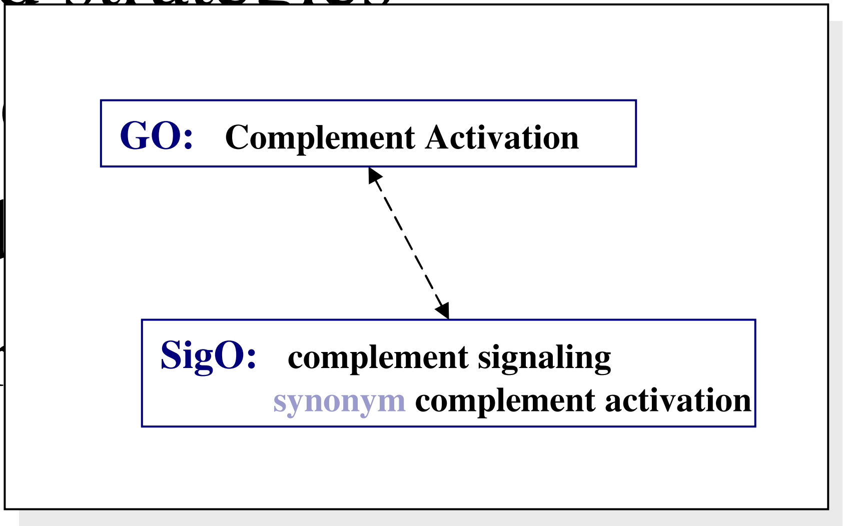
For example,

- Selection of features
- Selection of search space

Matchers

Matcher Strategies

- Strategies based on linguistic matching
- Structure-based strategies
- Constraint-bas
- Instance-based
- Use of auxiliar



GO: Complement Activation

SigO: complement signaling
synonym complement activation

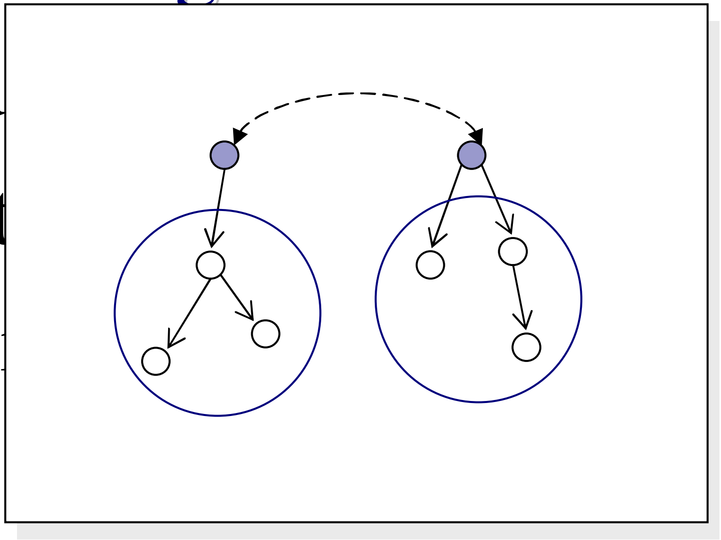
Example matchers

- Edit distance
  - Number of deletions, insertions, substitutions required to transform one string into another
  - aaaa baab: edit distance 2

- N-gram
  - N-gram : N consecutive characters in a string
  - Similarity based on set comparison of n-grams
  - aaaa : {aa, aa, aa}; baab : {ba, aa, ab}

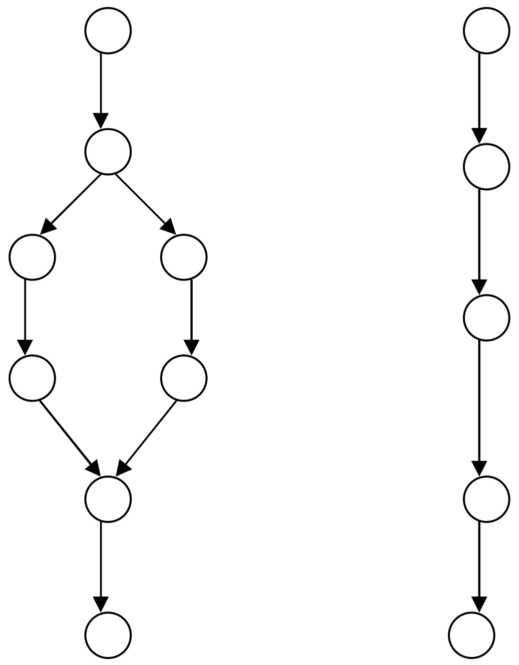
Matcher Strategies

- Strategies based on linguistic matching
- **Structure-based strategies**
- Constraint-based
- Instance-based st
- Use of auxiliary



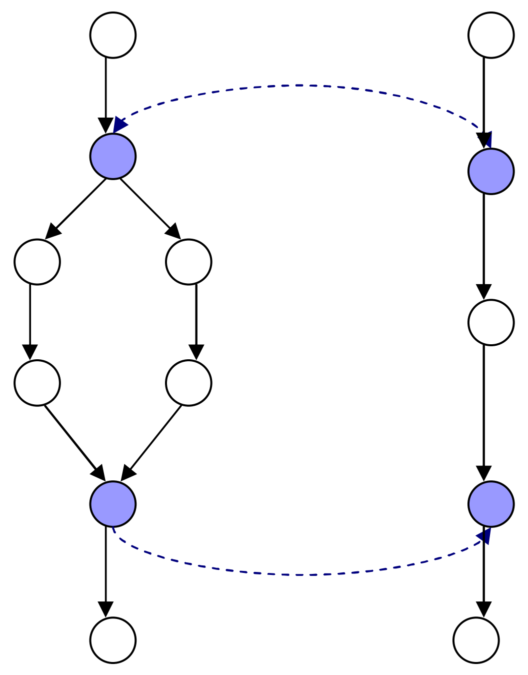
Example matchers

- Propagation of similarity values
- Anchored matching



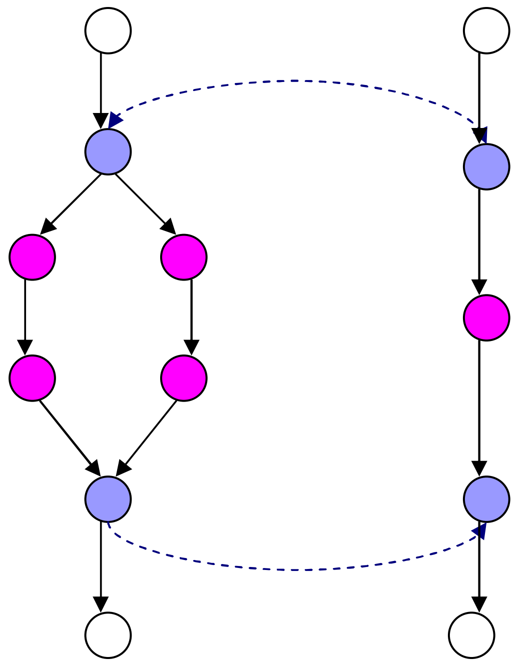
Example matchers

- Propagation of similarity values
- Anchored matching



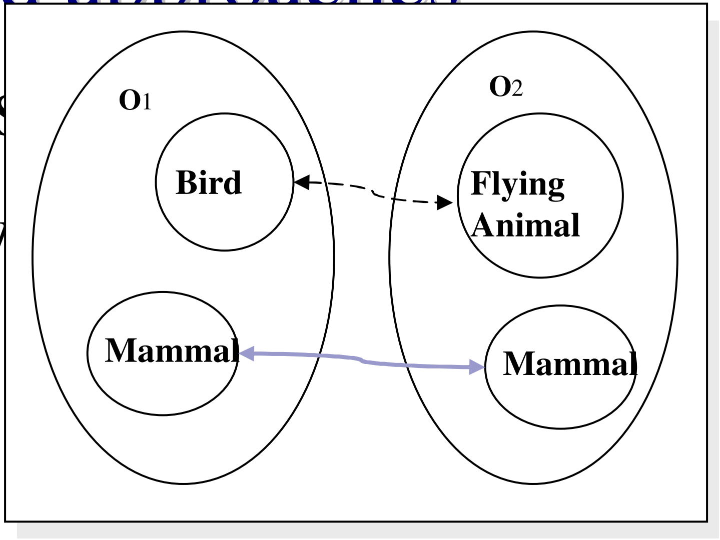
Example matchers

- Propagation of similarity values
- Anchored matching



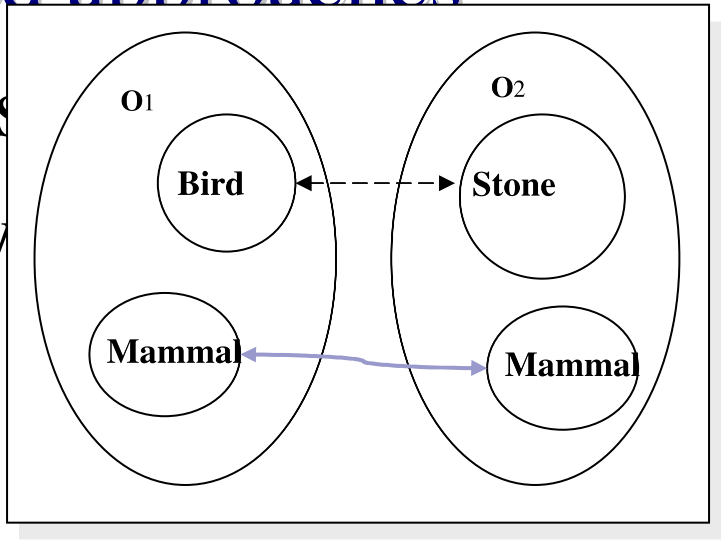
Matcher Strategies

- Strategies based on linguistic matching
- Structure-based strategies
- Constraint-based approaches
- Instance-based s
- Use of auxiliary



Matcher Strategies

- Strategies based on linguistic matching
- Structure-based strategies
- Constraint-based approaches
- Instance-based s
- Use of auxiliary

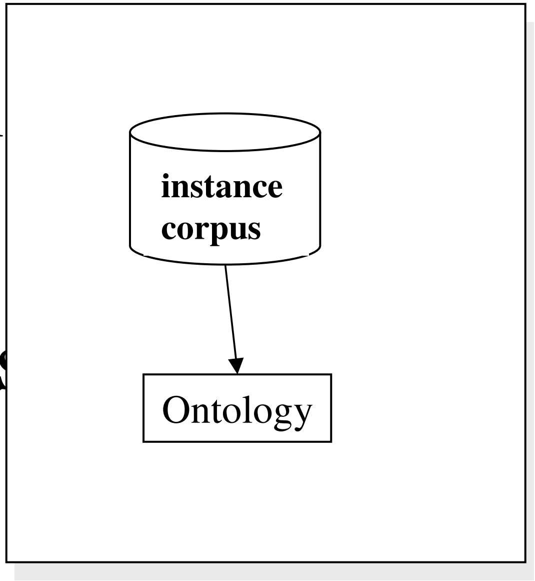


Example matchers

- Similarities between data types
- Similarities based on cardinalities

Matcher Strategies

- Strategies based on linguisti
- Structure-based strategies
- Constraint-based approaches
- Instance-based strategies
- Use of auxiliary information



Example matchers

- Instance-based
- Use life science literature as instances

Learning matchers – instance-based strategies

n Basic intuition

A similarity measure between concepts can be computed based on the probability that documents about one concept are also about the other concept and vice versa.

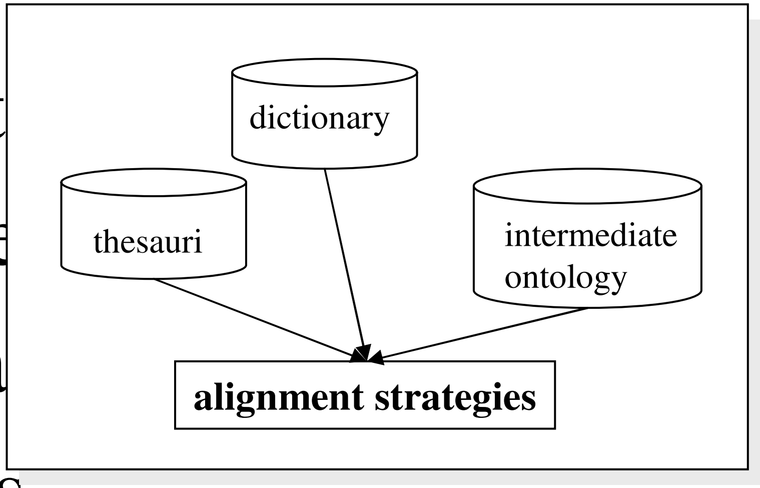
Basic Naïve Bayes matcher

- Generate corpora
  - Use concept as query term in PubMed
  - Retrieve most recent PubMed abstracts
- Generate classifiers
  - Naive Bayes classifiers, one per ontology
- Classification
  - Abstracts related to one ontology are classified to the concept in the other ontology with highest posterior probability P(C|d)
- Calculate similarities

$$sim(C_1, C_2) = \frac{n_{NBC2}(C_1, C_2) + n_{NBC1}(C_2, C_1)}{n_D(C_1) + n_D(C_2)}$$

Matcher Strategies

- Strategies based linguist
- Structure-based strategie
- Constraint-based approa
- Instance-based strategies
- Use of auxiliary information



Example matchers

- Use of WordNet
  - Use WordNet to find synonyms
  - Use WordNet to find ancestors and descendants in the is-a hierarchy
- Use of Unified Medical Language System (UMLS)
  - Includes many ontologies
  - Includes many mappings (not complete)
  - Use UMLS mappings in the computation of the similarity values

Ontology Alignment and Mergning Systems

	linguistic	structure	constraints	instances	auxiliary
ArtGen	name	parents, children		domain specific documents	WordNet
ASCO	name, label description	parents, children, siblings, path from root			WordNet
Chimaera	name	parents, children			
FCA-Merge	name			domain specific documents	
FOAM	name, label	parents, children	equivalence		
GLUE	name	neighborhood		instances	
HCONE	name	parents, children			WordNet
IF-Map				instances	a reference ontology
iMapper		leaf, non-leaf, children, related node	domain, range	instances	WordNet
OntoMapper		parents, children		documents	
(Anchor-) PROMPT	name	direct graphs			
SAMBO	name, synonym	is-a and part-of, descendants and ancestors		domain specific documents	WordNet, UMLS
S-Match	label	path from root	semantic relations codified in labels		WordNet

Combinations

Combination Strategies

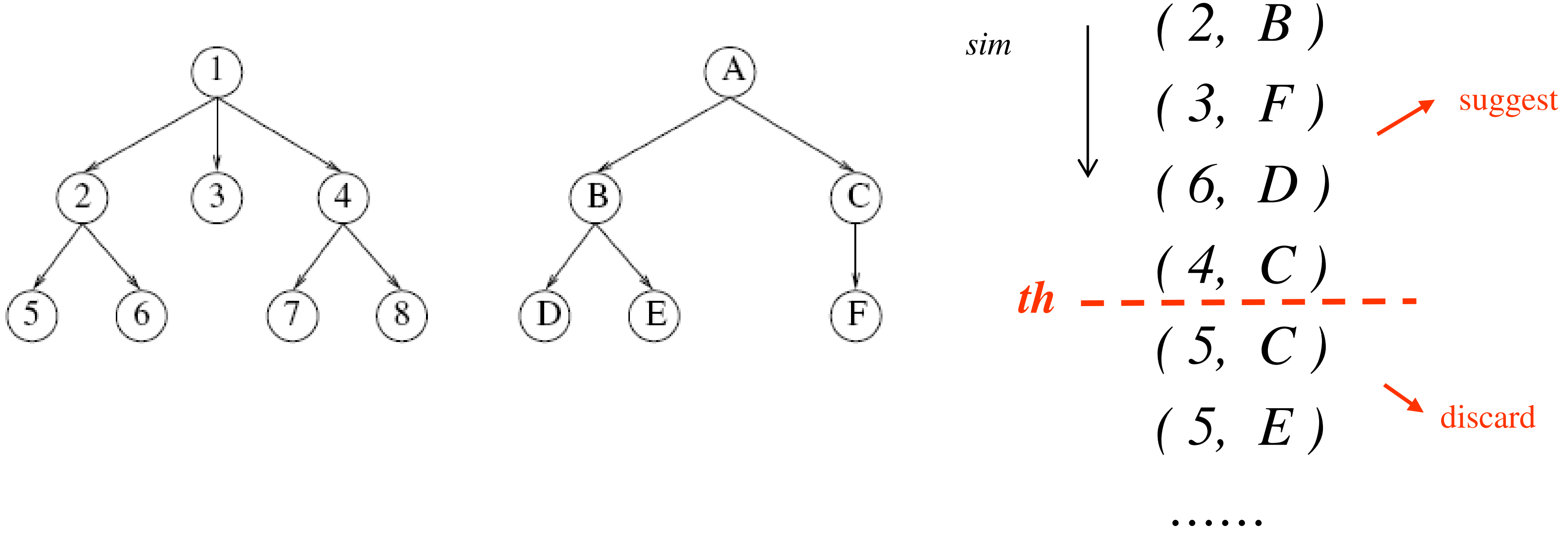
- Usually weighted sum of similarity values of different matchers
- Maximum of similarity values of different matchers

Filtering

Filtering techniques

- Threshold filtering

  Pairs of concepts with similarity higher or equal than threshold are mapping suggestions



sim

(*2*, *B*)

(*3*, *F*) suggest

(*6*, *D*)

(*4*, *C*)

th

(*5*, *C*)

(*5*, *E*) discard

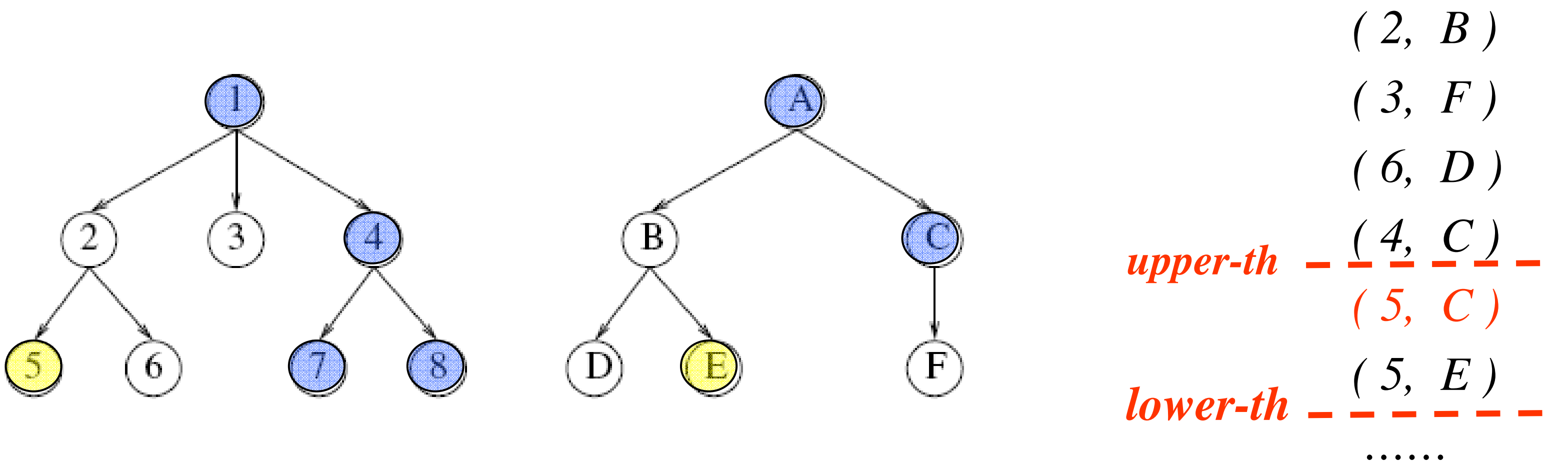
……

Filtering techniques

- Double threshold filtering

  (1) Pairs of concepts with similarity higher than or equal to **upper** threshold are mapping suggestions

  (2) Pairs of concepts with similarity between **lower** and **upper** thresholds are mapping suggestions if they make sense with respect to the structure of the ontologies and the suggestions according to (1)



(2, B)

(3, F)

(6, D)

(4, C)

upper-th

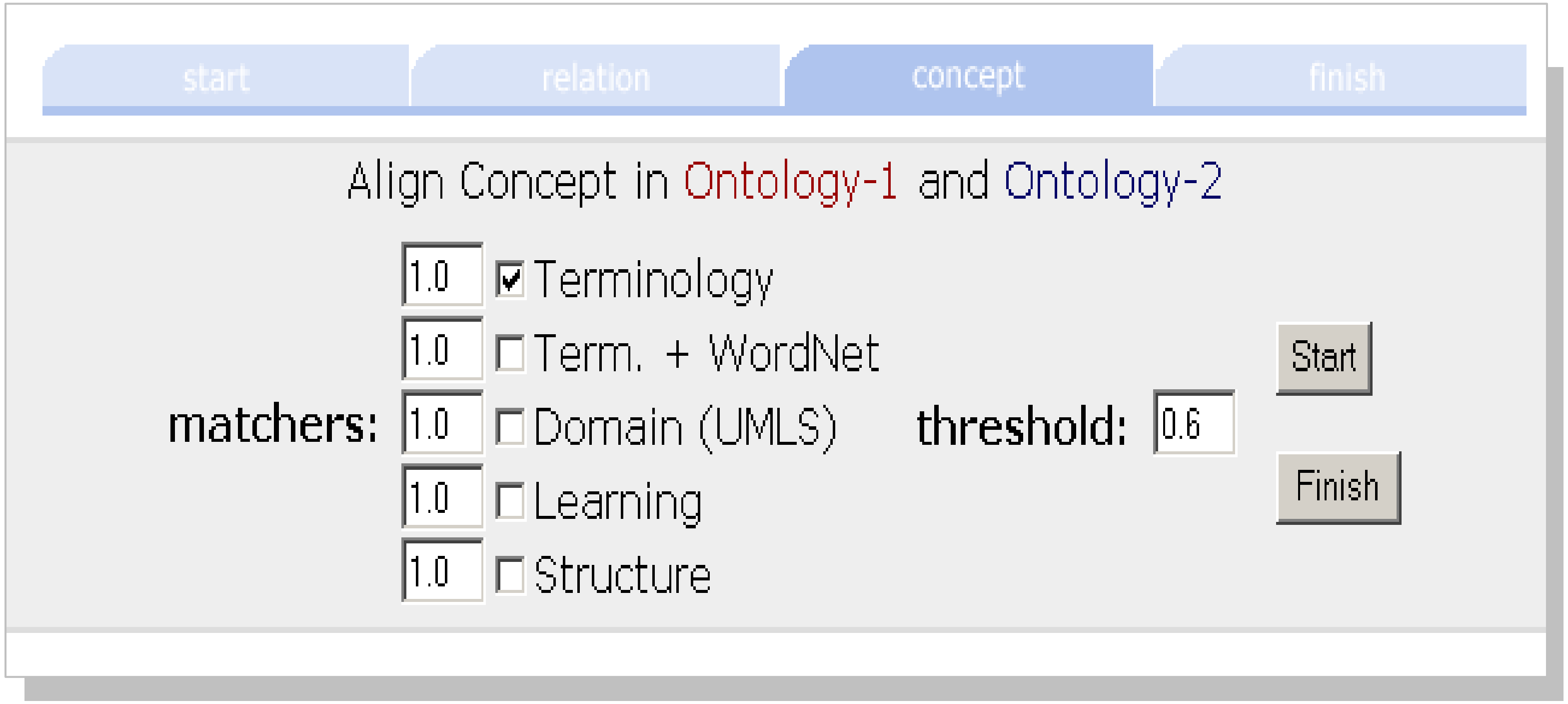
(5, C)

(5, E)

lower-th

……

Example alignment system SAMBO – preprocessing, matchers, combination, filter



Align Concept in Ontology-1 and Ontology-2

matchers:

1.0 ☑ Terminology

1.0 □ Term. + WordNet

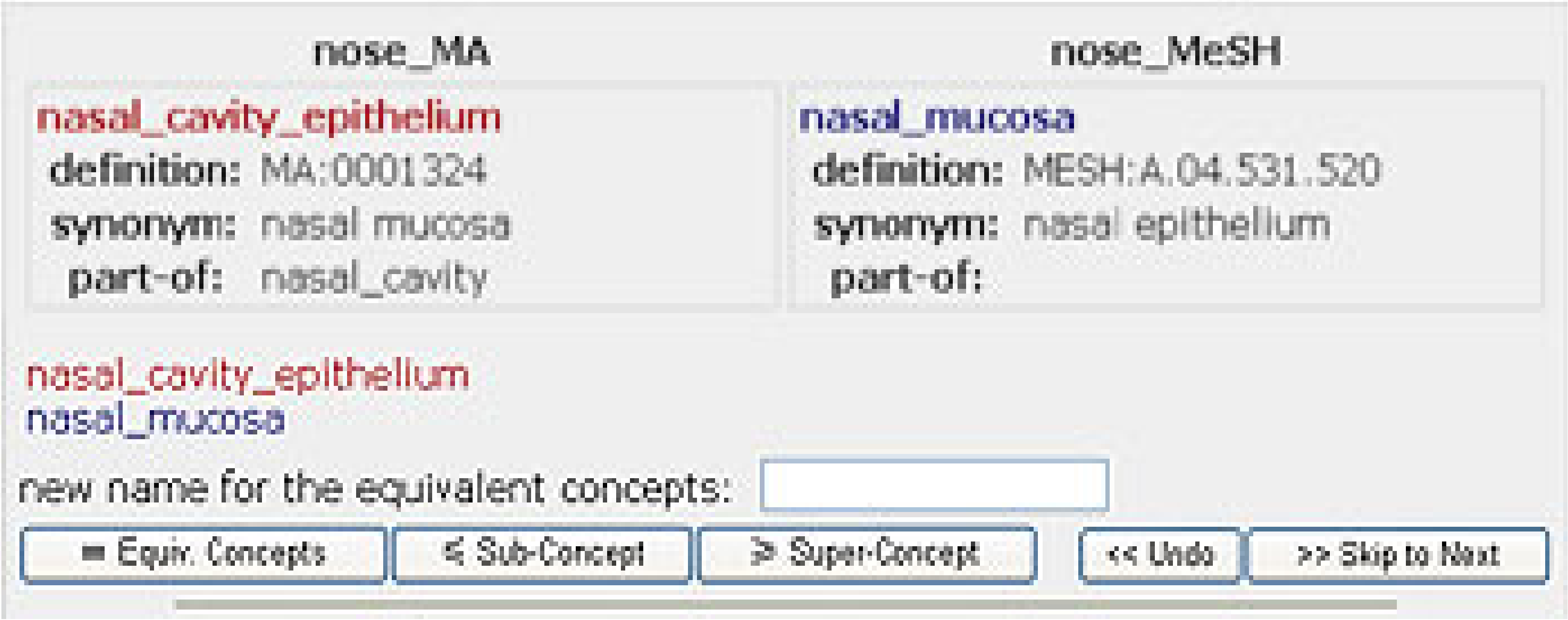
1.0 □ Domain (UMLS)

1.0 □ Learning

1.0 □ Structure

threshold: 0.6

Example alignment system SAMBO – **suggestion mode**



nose_MA

nasal_cavity_epithelium
definition: MA:0001324
synonym: nasal mucosa
part-of: nasal_cavity

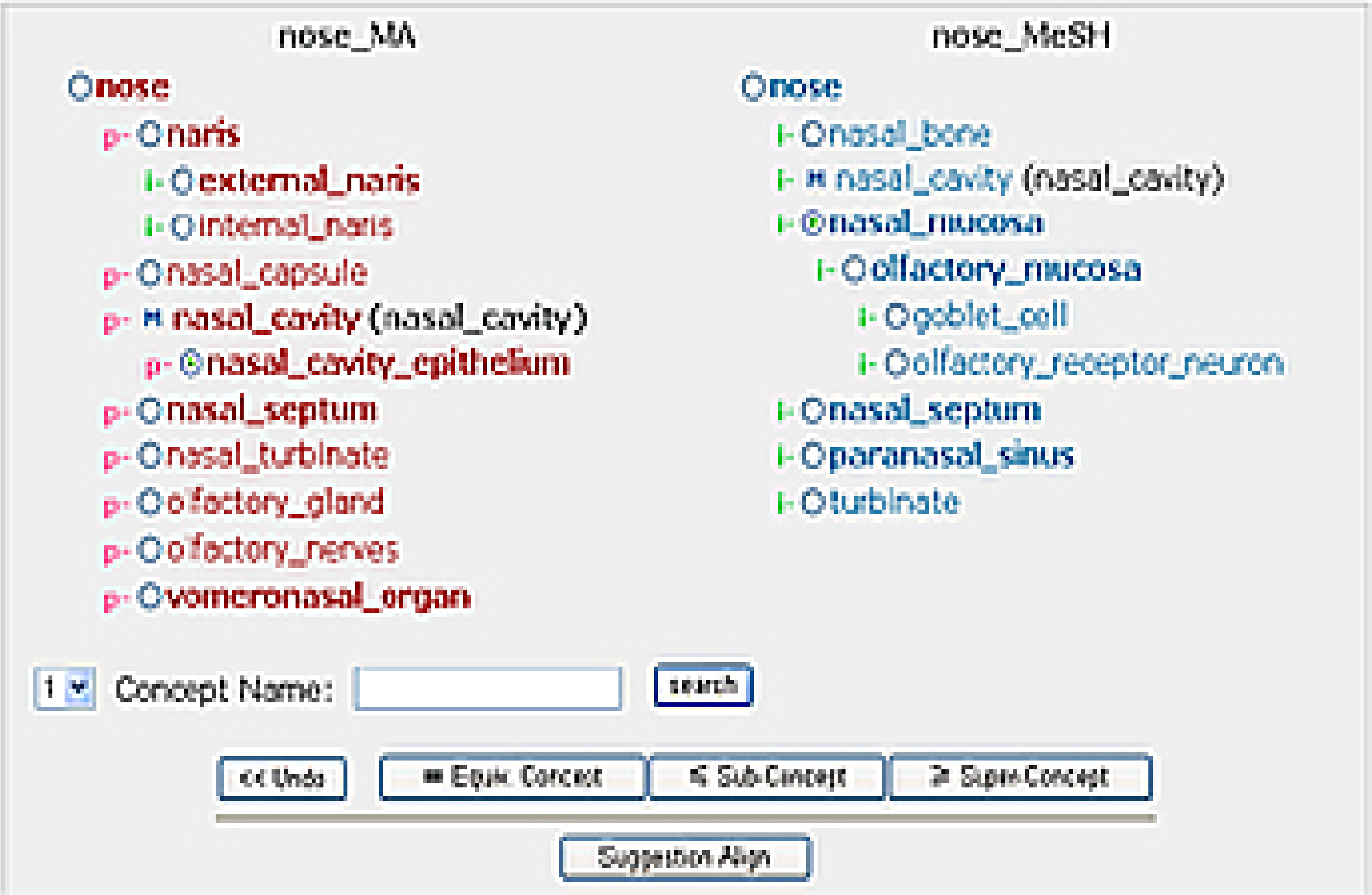
nose_MeSH

nasal_mucosa
definition: MESH:A.04.531.520
synonym: nasal epithelium
part-of:

nasal_cavity_epithelium
nasal_mucosa

new name for the equivalent concepts:

Example alignment system SAMBO – **manual mode**



nose_MA

nose
- p- naris
  - i- external_naris
  - i- internal_naris
- p- nasal_capsule
- p- nasal_cavity (nasal_cavity)
  - p- nasal_cavity_epithelium
- p- nasal_septum
- p- nasal_turbinate
- p- olfactory_gland
- p- olfactory_nerves
- p- vomeronasal_organ

nose_MeSH

nose
- i- nasal_bone
- i- nasal_cavity (nasal_cavity)
- i- nasal_mucosa
  - i- olfactory_mucosa
    - i- goblet_cell
    - i- olfactory_receptor_neuron
- i- nasal_septum
- i- paranasal_sinus
- i- turbinate

1 Concept Name:

Ontology Alignment

- Ontology alignment
- Ontology alignment strategies
- Evaluation of ontology alignment strategies
- Current issues
- Ontology-based literature search

Evaluation measures

- Precision:

$$\frac{\text{\# correct suggested mappings}}{\text{\# suggested mappings}}$$

- Recall:

$$\frac{\text{\# correct suggested mappings}}{\text{\# correct mappings}}$$

- F-measure: combination of precision and recall

Ontology Alignment Evaluation Initiative

OAEI

- Since 2004
- Evaluation of systems
- Different tracks
  - comparison: benchmark (open)
  - expressive: anatomy (blind), fisheries (expert)
  - directories and thesauri: directory, library, crosslingual resources (blind)
  - consensus: conference

OAEI 2007

- 17 systems participated
  - benchmark (13)
    - ASMOV: p = 0.95, r = 0.90
  - anatomy (11)
    - AOAS: f = 0.86, r+ = 0.50
    - SAMBO: f =0.81, r+ = 0.58
  - library (3)
    - Thesaurus merging: FALCON: p = 0.97, r = 0.87
    - Annotation scenario:
      - FALCON: pb =0.65, rb = 0.49, pa = 0.52, ra = 0.36, Ja = 0.30
      - Silas: pb = 0.66, rb= 0.47, pa = 0.53, ra = 0.35, Ja = 0.29
  - directory (9), food (6), environment (2), conference (6)

OAEI 2008 – anatomy track

- Align
  - Mouse anatomy: 2744 terms
  - NCI-anatomy: 3304 terms
  - Mappings: 1544 (of which 934 ‘trivial’)
- Tasks
  - 1. Align and optimize f
  - 2-3. Align and optimize p / r
  - 4. Align when partial reference alignment is given and optimize f

OAEI 2008 – anatomy track#1

- 9 systems participated
- SAMBO
  - p=0.869, r=0.836, r+=0.586, f=0.852
- SAMBOdtf
  - p=0.831, r=0.833, r+=0.579, f=0.832
- Use of TermWN and UMLS

OAEI 2008 – anatomy track#1

Is background knowledge (BK) needed?

Of the non-trivial mappings:

- Ca 50% found by systems using BK and systems not using BK
- Ca 13% found only by systems using BK
- Ca 13% found only by systems not using BK
- Ca 25% not found

Processing time:

hours with BK, minutes without BK

OAEI 2008 – anatomy track#4

Can we use given mappings when computing suggestions?

partial reference alignment given with all trivial and 50 non-trivial mappings

- SAMBO
  - p=0.636 0.660, r=0.626 0.624, f=0.631 0.642
- SAMBOdtf
  - p=0.563 0.603, r=0.622 0.630, f=0.591 0.616

(measures computed on non-given part of the reference alignment)

OAEI 2007-2008

- Systems can use only one combination of strategies per task

  systems use similar strategies

  - text: string matching, tf-idf
  - structure: propagation of similarity to ancestors and/or descendants
  - thesaurus (WordNet)
  - domain knowledge important for anatomy task?

Ontology Alignment

- Ontology alignment
- Ontology alignment strategies
- Evaluation of ontology alignment strategies
- Current Issues
- Ontology-based literature search

Current issues

- Systems and algorithms
  - Complex ontologies
  - Use of instance-based techniques
  - Alignment types (equivalence, is-a, …)
  - Complex mappings (1-n, m-n)
  - Connection ontology types – alignment strategies
- Evaluation
  - SEALS – Semantic Evaluation At Large Scale

Current issues

- Recommending ’best’ alignment strategies
- Use of Partial Reference Alignment

- Integration of ontology alignment and repair of the structure of ontologies

Ontology Alignment

- Ontology alignment
- Ontology alignment strategies
- Evaluation of ontology alignment strategies
- Current issues
- Ontology-based literature search

Literature search

- Huge amount of scientific literature.
- Need to integrate a spectrum of information to perform a task.

Literature search

- How to know what is in the repository
  - Lack of knowledge of the domain
- How to compose an expressive query
  - Lack of knowledge of search technology

Example scenario

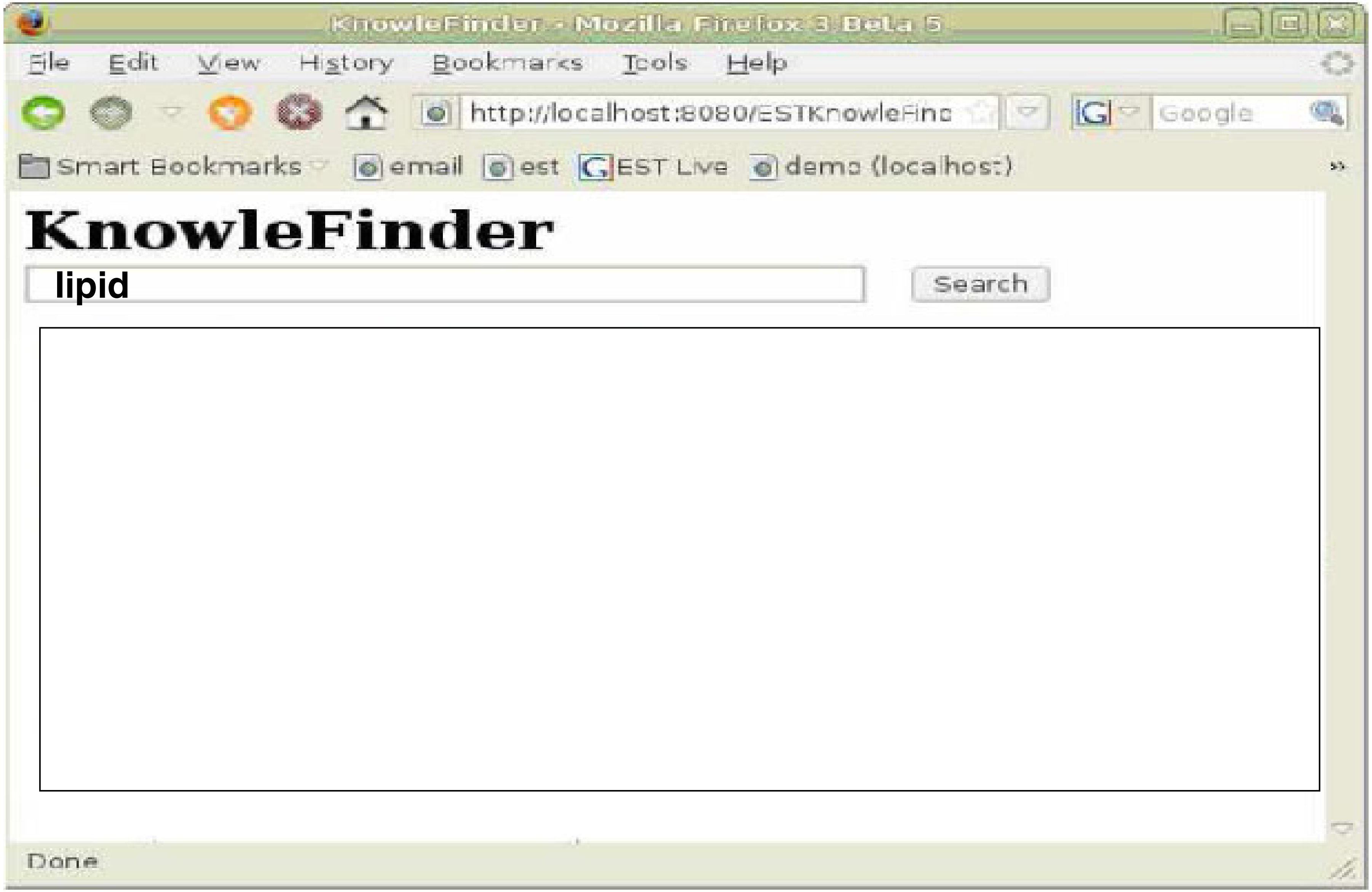
“Lipid”

- Keyword search returns all documents containing lipid.
  - No knowledge; terminology problem
- Relationships: use of multiple keywords with/without boolean operators,
  e.g. *lipid and disease*

Example scenario

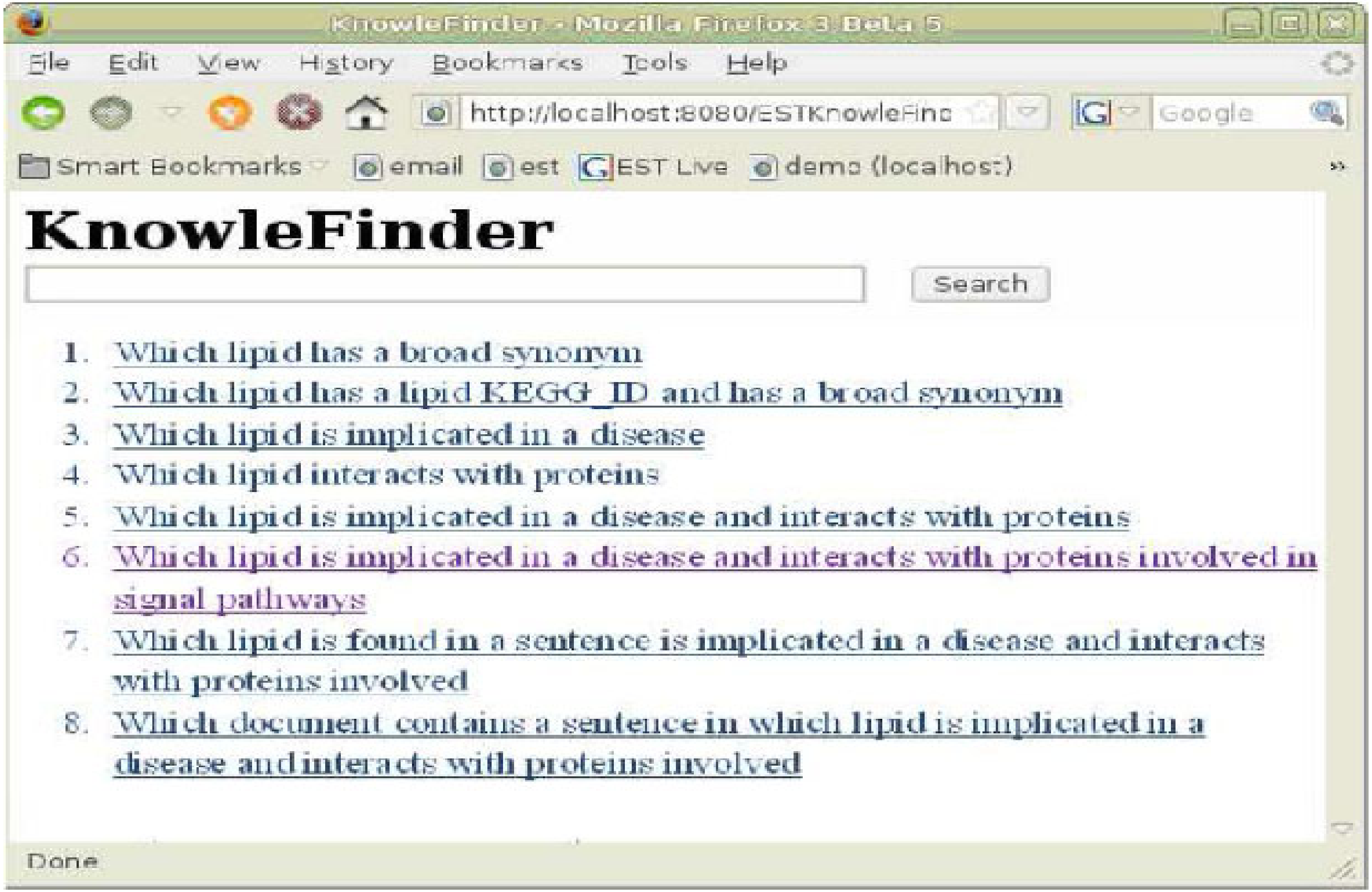
“Lipid”

- Keyword search returns a list of relevant questions concerning lipid. User selects question and retrieves knowledge and provenance documents.
- Multiple search terms: requirement that there are relevant connections between the keywords.



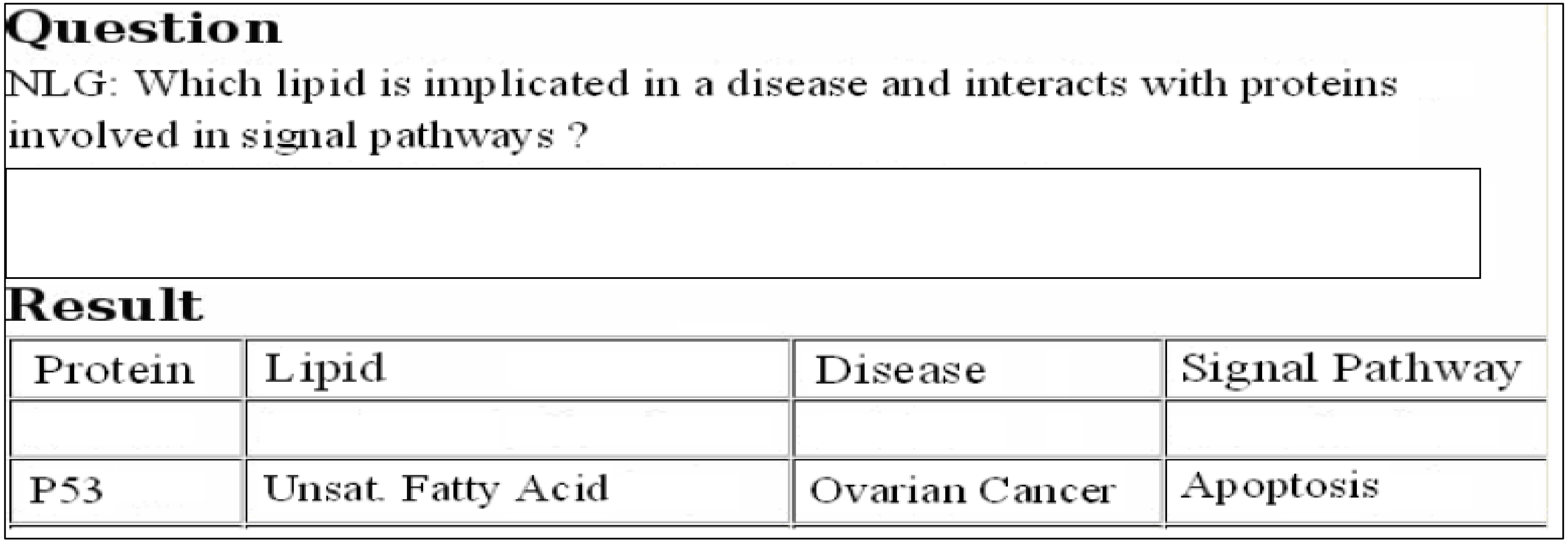
KnowleFinder

lipid


KnowleFinder - Mozilla Firefox 3 Beta 5
File Edit View History Bookmarks Tools Help
http://localhost:8080/ESTKnowleFinc
Google
Smart Bookmarks
email
est
EST Live
demo (localhost:)
Search
Done

KnowleFinder

1. Which lipid has a broad synonym
2. Which lipid has a lipid KEGG_ID and has a broad synonym
3. Which lipid is implicated in a disease
4. Which lipid interacts with proteins
5. Which lipid is implicated in a disease and interacts with proteins
6. Which lipid is implicated in a disease and interacts with proteins involved in signal pathways
7. Which lipid is found in a sentence is implicated in a disease and interacts with proteins involved
8. Which document contains a sentence in which lipid is implicated in a disease and interacts with proteins involved



Question

NLG: Which lipid is implicated in a disease and interacts with proteins involved in signal pathways ?

Result

Protein	Lipid	Disease	Signal Pathway
P53	Unsat. Fatty Acid	Ovarian Cancer	Apoptosis

Relevant queries

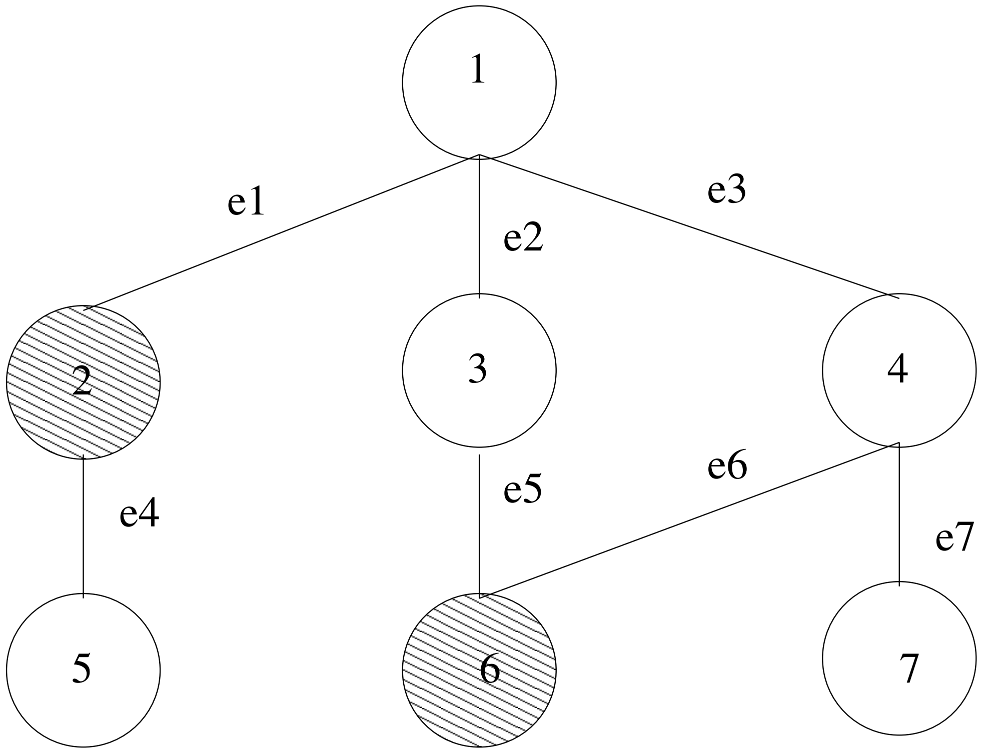
- Relevant query including a number of concepts and relations from an ontology

  connected sub-graph of the ontology that includes the concepts and relations.

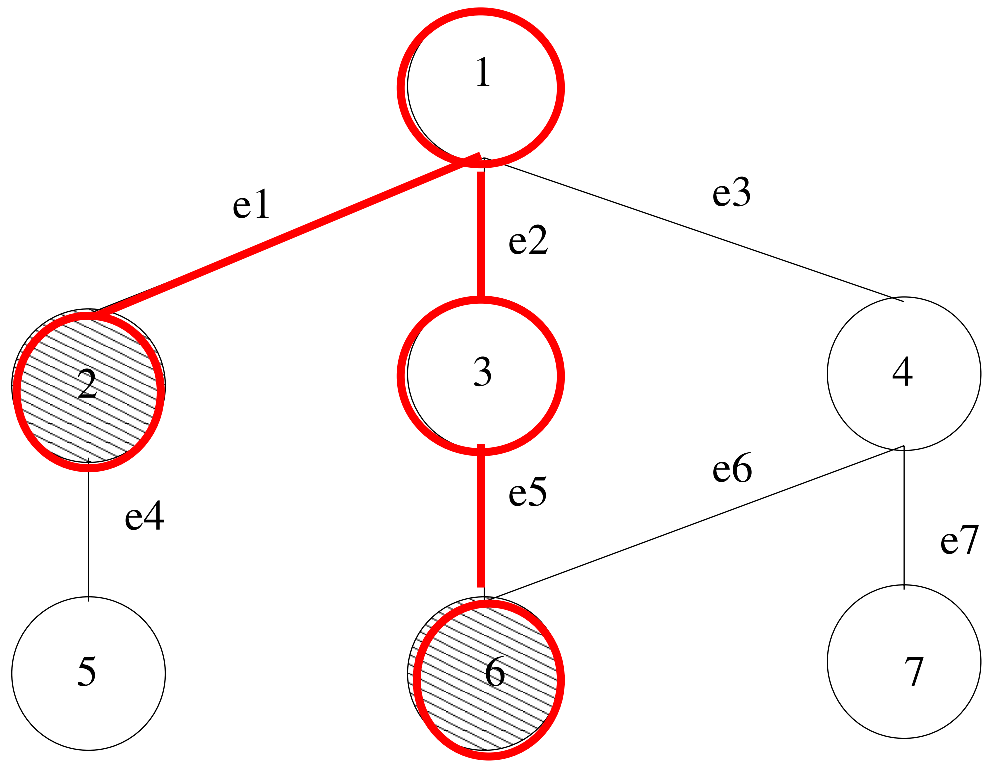
  (query graph based on the concepts and relations;

  slice is set of all query graphs based on the concepts and relations)

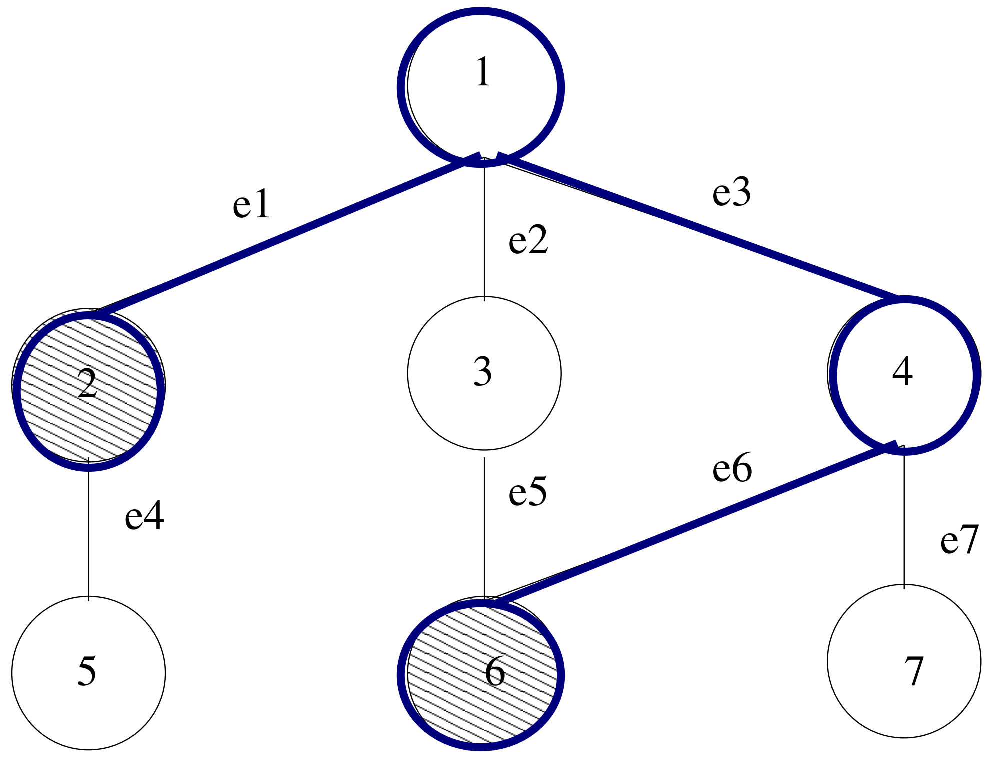
Query graph



Query graph



Query graph



Special cases

- No relations, several concepts
  - Relevant queries regarding concepts; relations are suggested by the system.
  - Difference with traditional techniques: extra requirement that search terms need to be connected in the ontology.
- No relations, one concept
  - Relevant queries including a specific query term.
  - Computes the ontological environment of the query term.

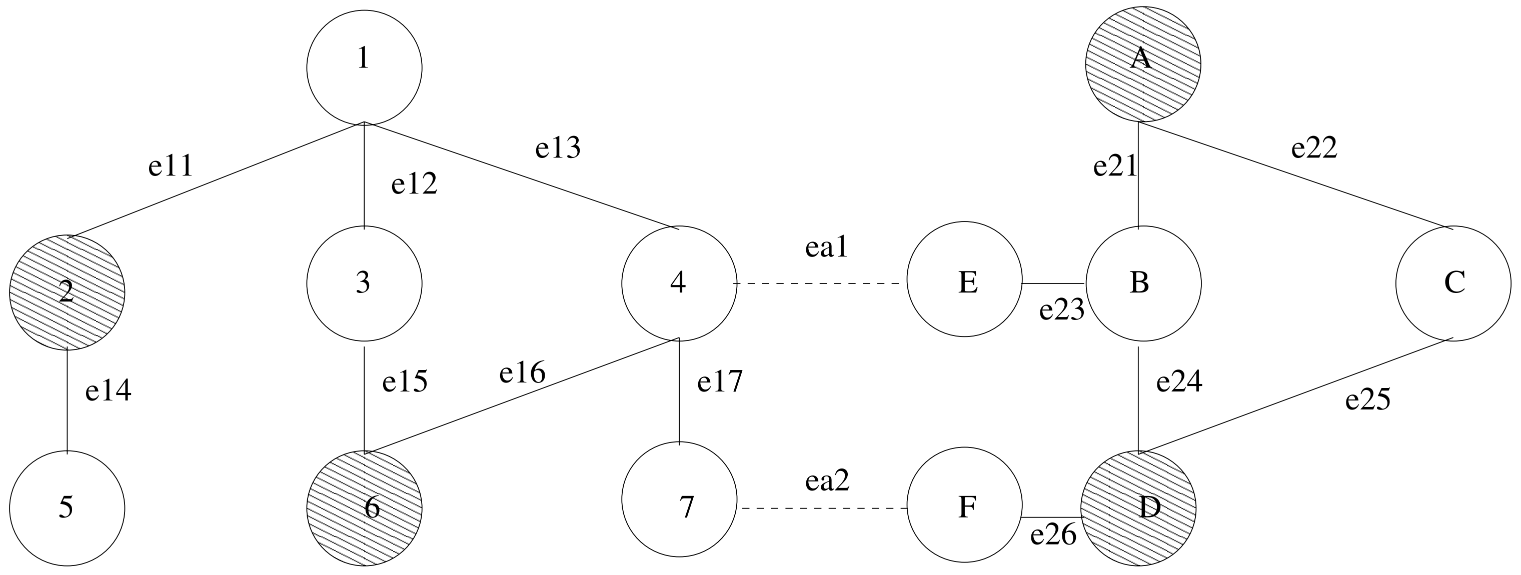
Relevant queries – multiple ontologies

- Relevant query including a number of concepts and relations from multiple ontologies

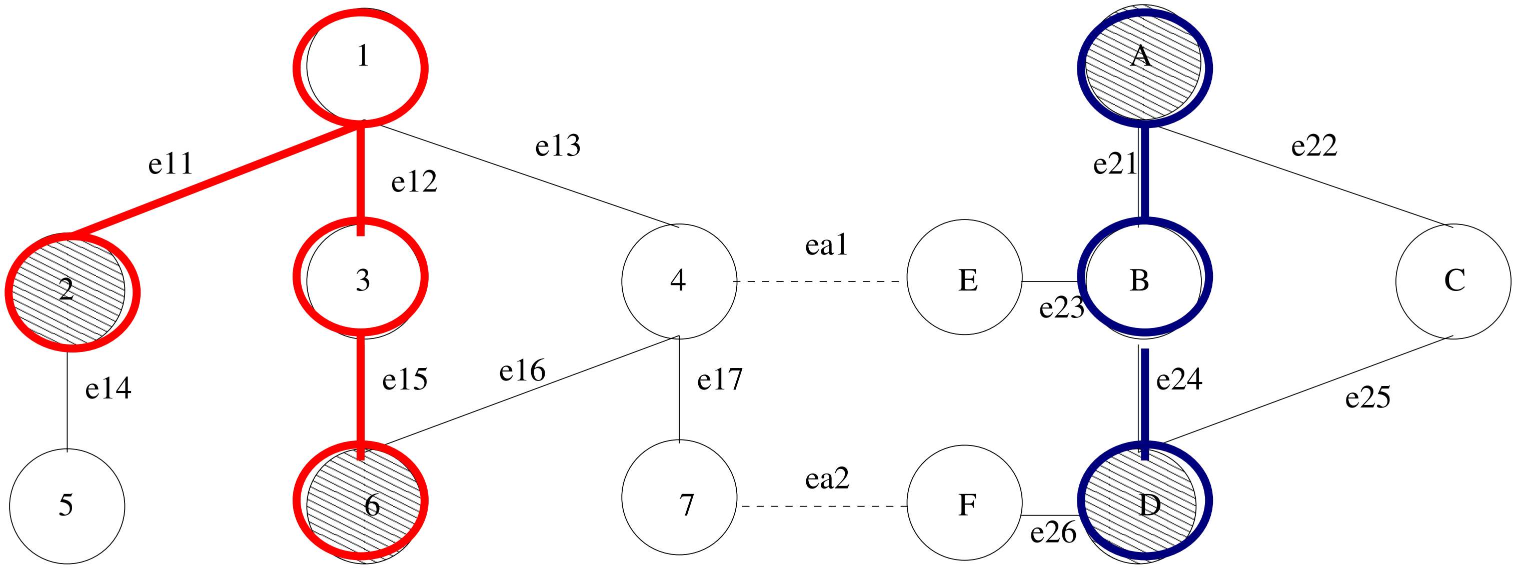
  Query graphs connected by a path going through a mapping in the alignment.

  (aligned query graph based on query graphs;
  aligned slice is set of all aligned query graphs based on the query graphs)

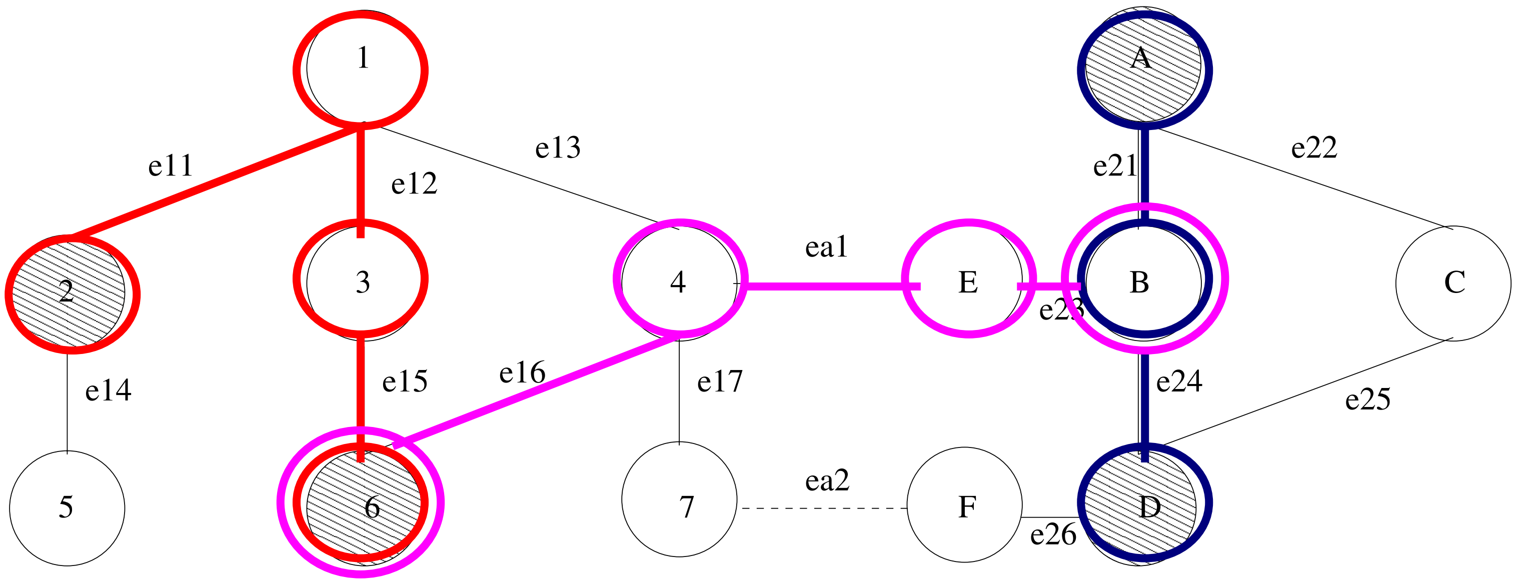
Aligned query graph



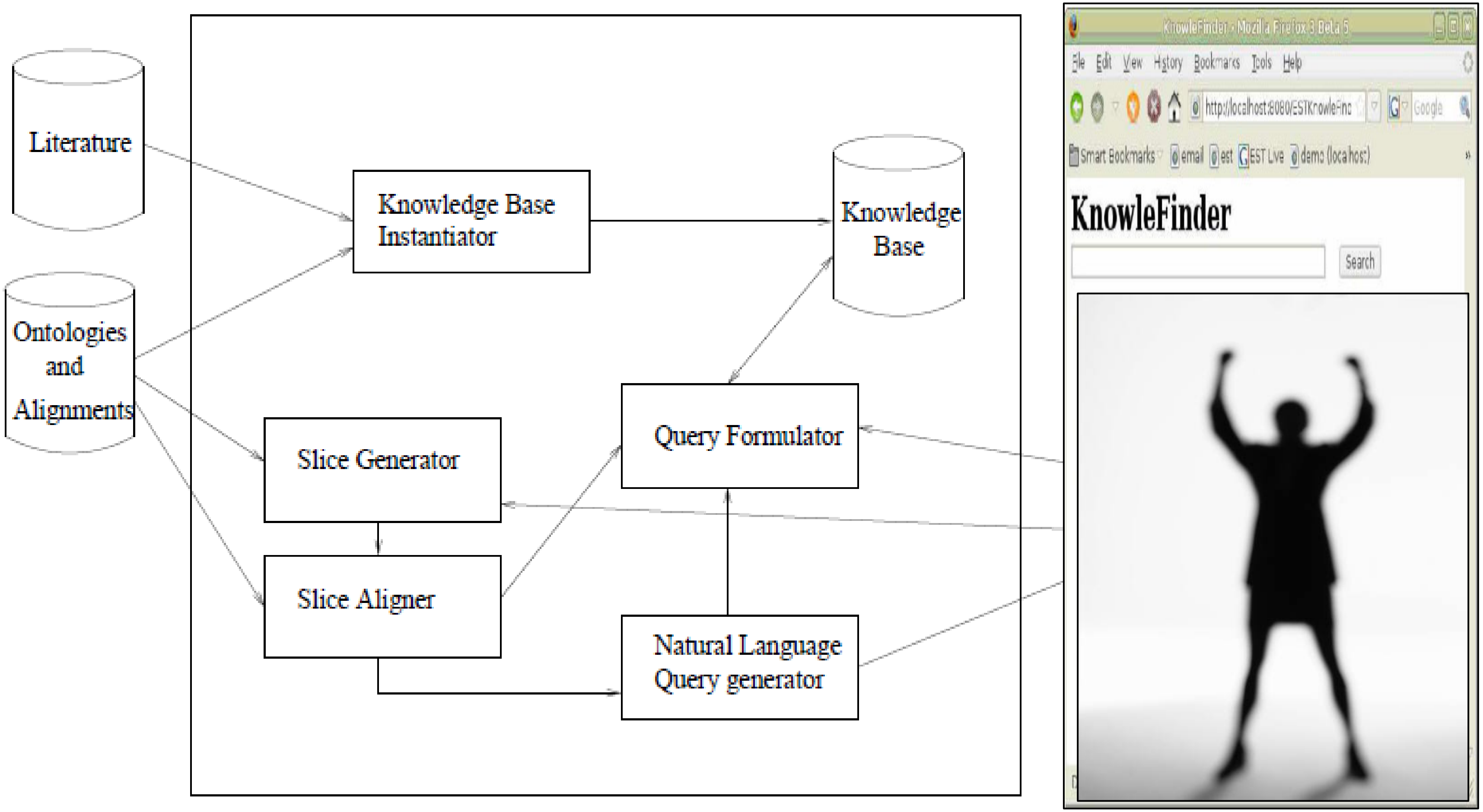
Aligned query graph



Aligned query graph



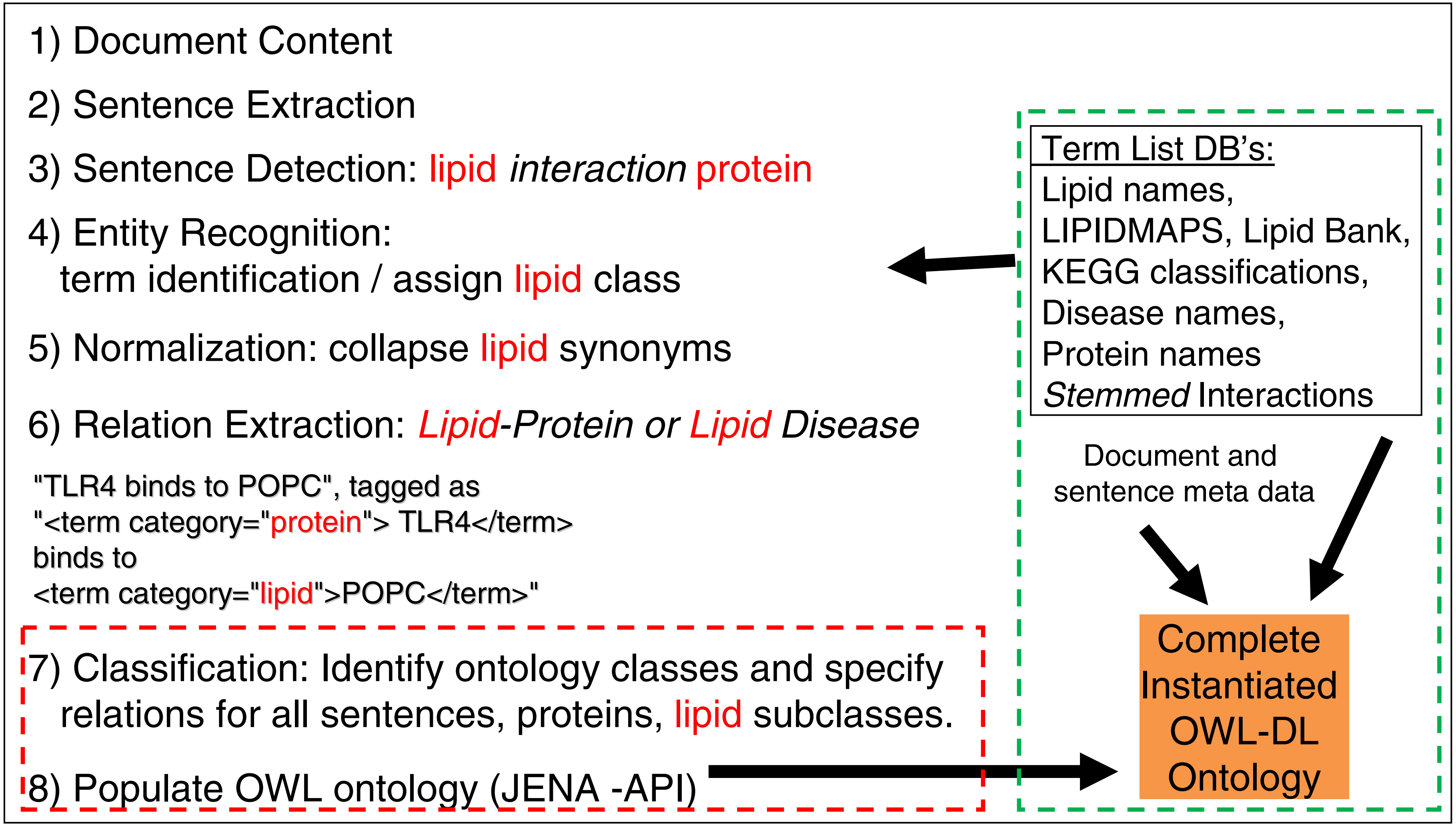
Framework


Literature
Ontologies and Alignments
Knowledge Base Instantiator
Knowledge Base
Slice Generator
Slice Aligner
Query Formulator
Natural Language Query generator
KnowleFinder - Mozilla Firefox 3 Beta 5
File Edit View History Bookmarks Tools Help
http://localhost:8080/ESTKnowleFinc
Google
Smart Bookmarks
email
est
EST Live
demo (localhost)
KnowleFinder
Search

External resources

- Literature document base
  - Generated from a collection of 7498 PubMed abstracts relevant for Ovarian Cancer. 683 papers included lipid names from which 241 full papers were downloadable.
- Ontology and ontology alignment repository
  - Lipid ontology
  - Signal ontology
  - Aligment using SAMBO

Knowledge base instantiation



1) Document Content

2) Sentence Extraction

3) Sentence Detection: lipid *interaction* protein

4) Entity Recognition:
term identification / assign lipid class

5) Normalization: collapse lipid synonyms

6) Relation Extraction: *Lipid-Protein or Lipid Disease*

"TLR4 binds to POPC", tagged as
"<term category="protein"> TLR4</term>
binds to
<term category="lipid">POPC</term>"

7) Classification: Identify ontology classes and specify relations for all sentences, proteins, lipid subclasses.

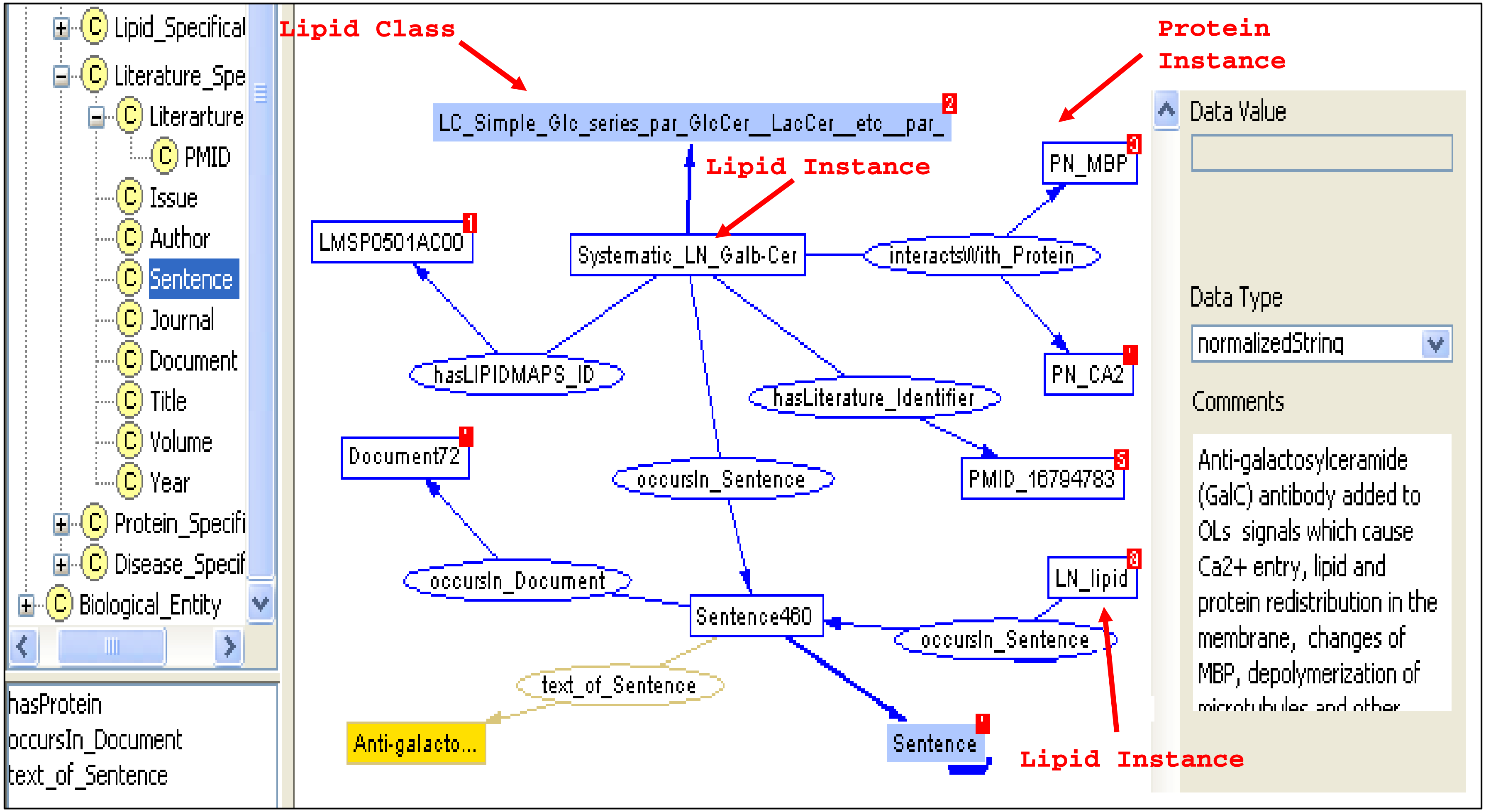
8) Populate OWL ontology (JENA -API)

Term List DB's:
Lipid names,
LIPIDMAPS, Lipid Bank,
KEGG classifications,
Disease names,
Protein names
Stemmed Interactions

Document and
sentence meta data

Complete
Instantiated
OWL-DL
Ontology

Knowledge base instantiation



Lipid Class

Lipid Instance

Protein Instance

Lipid Instance

Data Value

Data Type

normalizedString

Comments

Anti-galactosylceramide (GalC) antibody added to OLs signals which cause Ca2+ entry, lipid and protein redistribution in the membrane, changes of MBP, depolymerization of microtubules and other

Slice generation

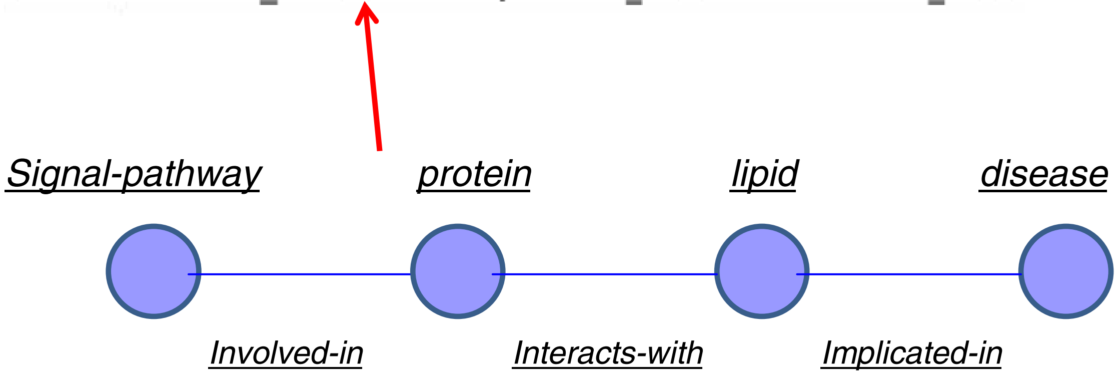
- Current implementation focuses on slices based on concepts.
- Depth-first traversal of ontology to find paths between given concepts; paths can be put together to find slices/query graphs.

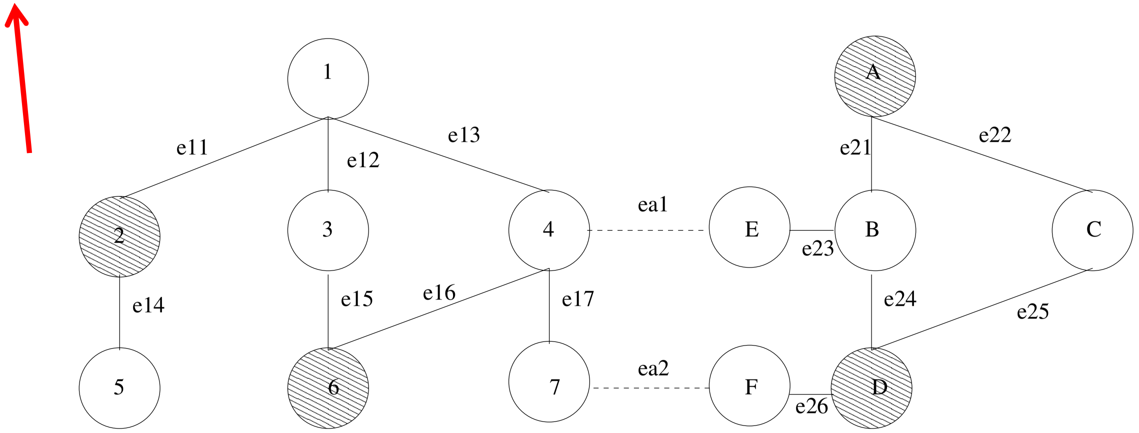
Slice alignment

- Algorithm computes subset of aligned slice.
- Assumption: shorter paths represent closer relationships.
- Algorithm connects slices using shortest paths from given concepts in one ontology to given concepts in other ontology.

Slicing through the literature

```
nRQL: (RETRIEVE (?X ?Y ?Z ?W)
(AND (?X Protein) (?Y Lipid) (?Z Disease) (?W SignalPathway)
(?X ?Y Interacts_with)(?Y ?Z Implicated_in)(?X ?W Involved_in)))
```





Natural language query generation

- Triple representation:

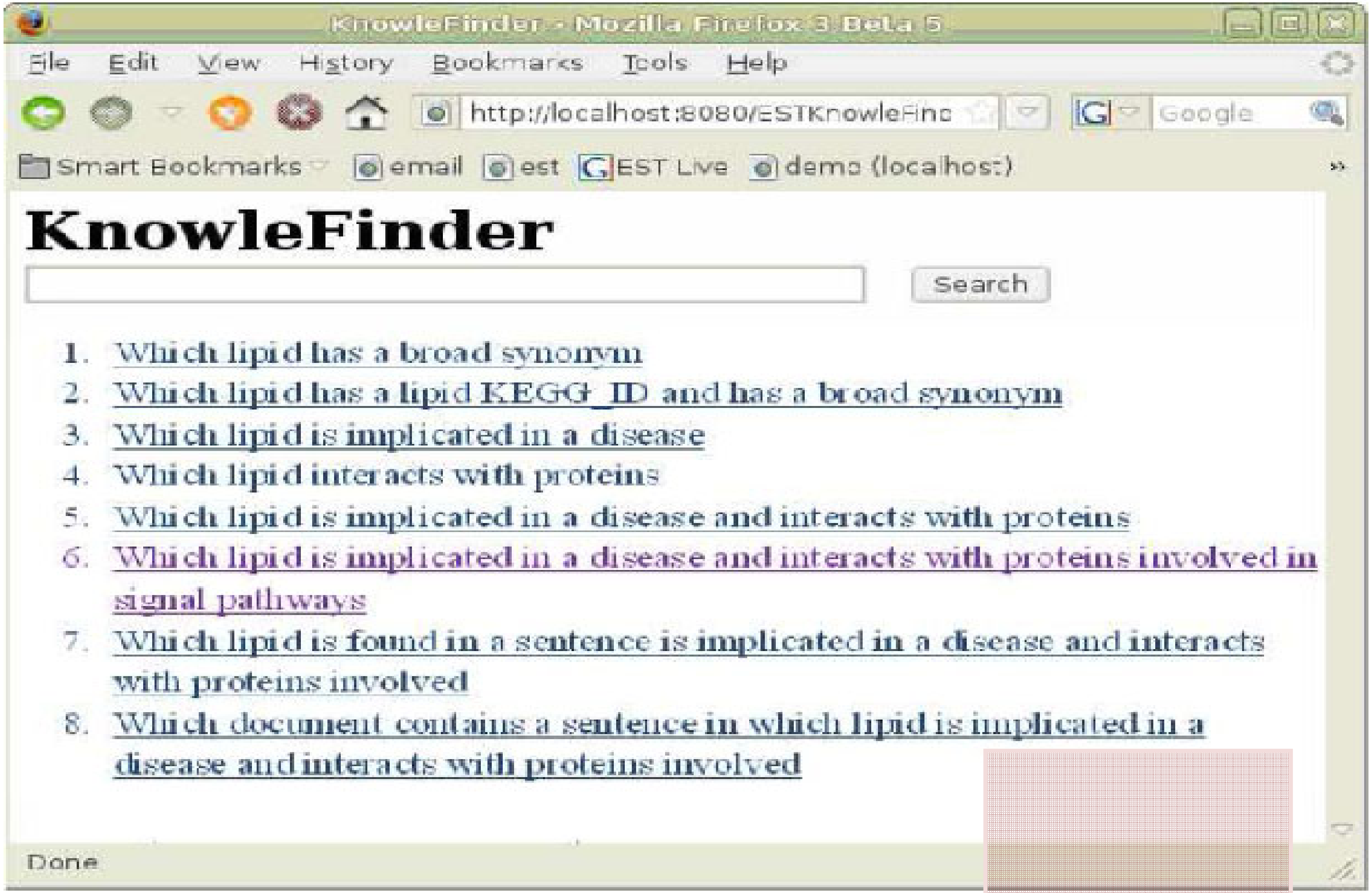
  <lipid, interacts-with, protein>

- Rule base to generate NL statements.

  What lipid interacts with proteins?

  - Learned from examples.

- Aggregation of statements from different triples, grammar checking.



KnowleFinder - Mozilla Firefox 3 Beta 5
File Edit View History Bookmarks Tools Help
http://localhost:8080/ESTKnowleFinc
Google
Smart Bookmarks
email
est
EST Live
demo (localhost)
Done

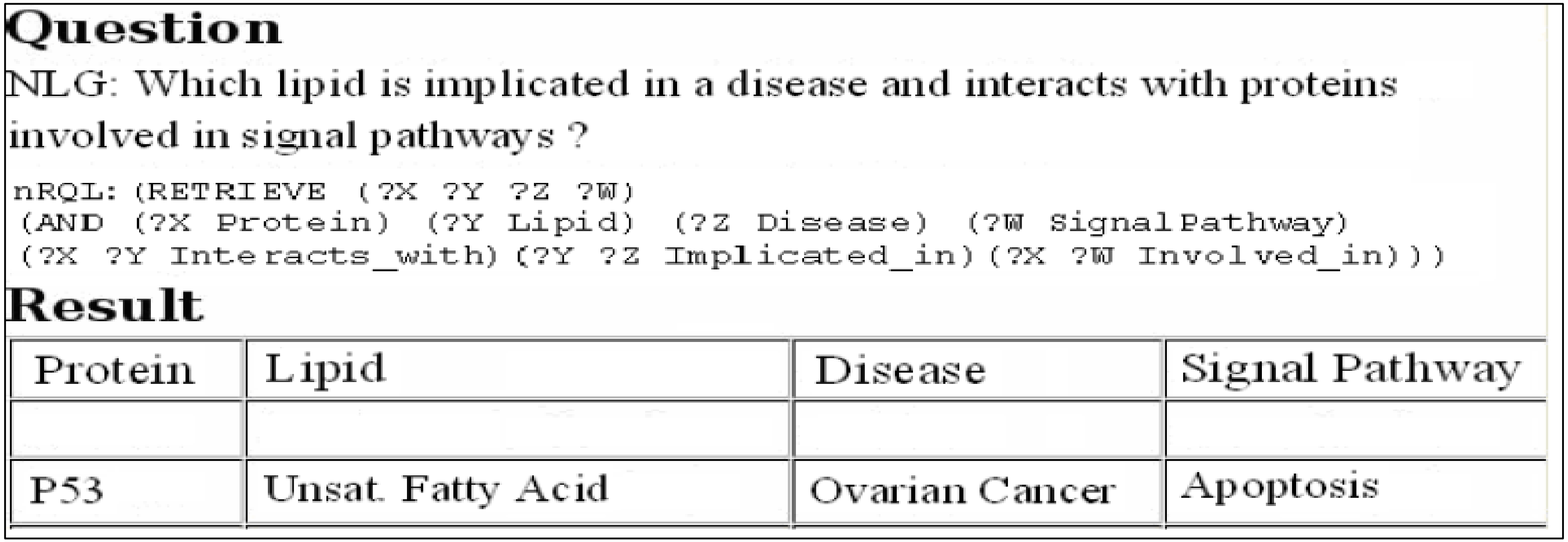
KnowleFinder

Search

1. Which lipid has a broad synonym
2. Which lipid has a lipid KEGG_ID and has a broad synonym
3. Which lipid is implicated in a disease
4. Which lipid interacts with proteins
5. Which lipid is implicated in a disease and interacts with proteins
6. Which lipid is implicated in a disease and interacts with proteins involved in signal pathways
7. Which lipid is found in a sentence is implicated in a disease and interacts with proteins involved
8. Which document contains a sentence in which lipid is implicated in a disease and interacts with proteins involved

Query

- Send nRQL query to RACER.



Question

NLG: Which lipid is implicated in a disease and interacts with proteins involved in signal pathways ?

```
nRQL: (RETRIEVE (?X ?Y ?Z ?W)
(AND (?X Protein) (?Y Lipid) (?Z Disease) (?W SignalPathway)
(?X ?Y Interacts_with)(?Y ?Z Implicated_in)(?X ?W Involved_in)))
```

Result

Protein	Lipid	Disease	Signal Pathway
P53	Unsat. Fatty Acid	Ovarian Cancer	Apoptosis

Future Work

- Tradeoff in query generation between completeness and information overload.
- Relevance measure and query ranking.
- Integrated implementation.
- Scalability testing.

Further reading

Ontology alignment - general

- http://www.ontologymatching.org

(plenty of references to articles and systems)

- Ontology alignment evaluation initiative: http://oaei.ontologymatching.org

(home page of the initiative)

- Euzenat, Shvaiko, *Ontology Matching*, Springer, 2007.

- Lambrix, Strömbäck, Tan, Information integration in bioinformatics with ontologies and standards, in Bry, Maluszynski (eds), *Semantic Techniques for the Web: The REWERSE perspective*, chapter 8, 343-376, 2009.

(contains currently largest overview of ontology alignment systems)

Further reading

Ontology alignment - systems

- Lambrix, Tan, SAMBO – a system for aligning and merging biomedical ontologies, *Journal of Web Semantics*, 4(3):196-206, 2006.

(description of the SAMBO tool and overview of evaluations of different matchers)

- Lambrix, Tan, A tool for evaluating ontology alignment strategies, *Journal on Data Semantics*, VIII:182-202, 2007.

(description of the KitAMO tool for evaluating matchers)

Further reading

Ontology alignment - recommendation of alignment strategies

- Tan, Lambrix, A method for recommending ontology alignment strategies, *International Semantic Web Conference*, 494-507, 2007.
- Ehrig, Staab, Sure, Bootstrapping ontology alignment methods with APFEL, *International Semantic Web Conference*, 186-200, 2005.
- Mochol, Jentzsch, Euzenat, Applying an analytic method for matching approach selection, *International Workshop on Ontology Matching*, 2006.

Ontology alignment - PRA in ontology alignment

- Lambrix, Liu, Using partial reference alignments to align ontologies, *European Semantic Web Conference*, 188-202, 2009.

Literature search

- Baker, Lambrix, Laurila Bergman, Kanagasabai, Ang, Slicing through the scientific literature, *Data Integration in the Life Sciences*, 127-140, 2009.

DILS 2010
7th International Conference on
Data Integration in the Life Sciences

August 25-27, Gothenburg, Sweden

paper submission deadline in April