Computing on BioPAX

Ontology of Networks Workshop
Newark, March 2008

Andrea Splendiani
Medical Informatics
University of Rennes 1, FR
Index (three examples)

- Abstraction of pathways as interaction networks
- From pathways to causal networks
- Extended scoring functions on ontologies (in progress)
Abstraction of Pathways as interaction networks

- **Input:** a set of ontologies represented in BioPAX (Reactome Home Sapience in PathwayCommons)

- **Output:** a representation of these ontologies as a network of interactions between physical elements

- **Method:** define the intended semantics (here from biopax), rely on OWL reasoning, and standard Semantic Web technologies.
Abstraction of Pathways as interaction networks

Reactome HomoSapiens BioPAX export

~27Mb
~270k triples
~27k individuals

Input
Abstraction of Pathways as interaction networks

Semantics

• The meaning of “interaction”.
• Description in formal terms (OWL)
Abstraction of Pathways as interaction networks

Reactome
HomoSapiens
BioPAX

[Skip-Context1:
(?i an:directp ?p) <- (?i bp:PARTICIPANT ?p)
(?p rdf:type bp:entity)]

[Skip-Context2:
(?i an:directp ?p) <- (?i bp:PARTICIPANTS ?pep)
(?pep bp:PHYSICAL'ENTITY ?p)
(?p rdf:type bp:entity)]

[Interacts:
(?x an:interacts ?y) <- (?z rdf:type bp:interaction)
(?z and:directp ?x) (?z an:directp ?y)]

Additional logic as rules

- To skip physical-entity-participants
- To represent interactions as a property, instead of a class
Abstraction of Pathways as interaction networks

Reactome
HomoSapiens
BioPAX

Reasoning
http://bioinformatics-org/rdfscape/

based on: Pellet, Jena, Cytoscape

- Computes standard entailments from OWL definition of terms (Pellet, other reasoners)
- Computes additional logic through application of rules
Abstraction of Pathways as interaction networks

Reactome
HomoSapiens
BioPAX


[Interacts: (?x an:interacts ?y) <- (?z rdf:type bp:interaction) (?z and:directp ?x) (?z and:directp ?y)]
Abstraction of Pathways as interaction networks

SPARQL

SPARQL query

RESULT

re:ribonucleotide_reductase_M2B__TP53_inducible__variant___TP53
re:ATP__ChEBI_15422__TP53
re:UniProt_P04637_1_Cellular_tumor_antigen_p53___Tumor_su__TP53
re:ADP__ChEBI_16761__TP53
re:53BP1_H2AX_complex_at_site_of_double_strand_break___TP53BP1
re:UniProt_Q12888_Tumor_suppressor_p53__binding_protein___TP53BP1
re:gamma_H2AX_MDC1_NERD1_complex_at_site_of_DNA_dou__TP53BP1
Abstraction of Pathways as interaction networks

<table>
<thead>
<tr>
<th>Task</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loading and parsing (38Mb, 260k statements)</td>
<td>10-20</td>
</tr>
<tr>
<td>OWL reasoning (Pellet) + query: retrieve all interactions</td>
<td>60/12</td>
</tr>
<tr>
<td>(partial) OWL reasoning (Jena) + query: retrieve all interactions</td>
<td>5/5</td>
</tr>
</tbody>
</table>

Performances
From pathways to causal networks

Same approach can be used to make computations on pathways, in the limits of expressivity of OWL

Example: from Pathway data to cause-effect networks

One possible definition of cause-effect:

```
[affects: (?x an:affects ?y) <- (?r1 bp:LEFT ?xpep)
 (?r1 rdf:type bp:biochemicalReaction)(?xpep bp:PHYSICAL-ENTITY ?x)
 (?r1 bp:RIGHT ?zpep1) (?zpep1 bp:PHYSICAL-ENTITY ?z)
 (?zpep2 bp:RIGHT ?ypep) (?ypep bp:PHYSICAL-ENTITY ?y)]

[transitivity: (?x an:affects ?y) <- (?x an:affects ?z) (?z an:affects ?y)]
```
Synthesis

Knowledge base
PathwayCommons
(not a list of pathways)

Ontology

Derived
information

Additional logic:
• non expressible in owl
• instance based ontologies

RDFScape: Semantic Web Meets Systems Biology, BMC Bioinformatics, in press
Computing on biopax (other experiences)


(similar approach to the one presented)

Alan Ruttenberg and Jeremy Zucker: debugging the bug (http://bio.freelogy.org/wiki/Debugging_the_bug)

(benefits of a class based representation)
Computing on biopax (limitations)

- Lack of use of URIs requires extra logic if multiple sources have to be integrated.
- Relations such as “part-of” “participates-in” are not directly represented (can be derived).
- Instance based representations don’t benefit directly from DL expressivity.
- Limits in the expressivity of OWL-DL to represent pathways.
Computing on biopax
(on the role of BioPAX as an Ontology)

Knowledge base
PathwayCommons

Derived information

Interpretation of biopax content
• fit to visualization
• maybe limited for other purposes
• (note: lacks basics as process/event in BFO)

Additional logic:
• non expressible in owl
• instance based ontologies
Extended scoring functions on ontologies (work in progress)

- Objective: to extend terms scoring approaches used in functional evaluation using additional information from ontologies (relations, types)

- In particular, to improve pathway scoring methods leveraging on ontologies

- Builds on top of what presented before, hence on deducted information from ontologies. Without lacking in generality, and simple pathway ontology and not BioPAX will be presented
Extended scoring functions on ontologies (work in progress)

Data for each entry (eventually derived)

Fisher’s Test.....

List of terms associated to each entry

Ranking of relevant terms (pathways are terms)

Classical Approach
Extended scoring functions on ontologies (work in progress)

- Data as a network
- Fisher’s Test.....
- List of terms associated to each entry
- Ranking of relevant terms (pathways are terms)
Extended scoring functions on ontologies (work in progress)

Data as a network

Fisher’s Test.....

Ranking of relevant terms (pathways are terms)

Ontology
Extended scoring functions on ontologies (work in progress)
Extended scoring functions on ontologies (work in progress)

Data as a network → Functions defined on RDF graphs → Score per class, property or instance

Ontology
Extended scoring functions on ontologies  (work in progress)
Extended scoring functions on ontologies (work in progress)
Extended scoring functions on ontologies (work in progress)

After finding a correspondence between the two graphs, values are associated to the ontology
Extended scoring functions on ontologies (work in progress)

\[ f(?x_2) = \frac{\text{SUM}(\text{var}(?x_1))}{\text{COUNT}(?x_1)} \]
Thanks