

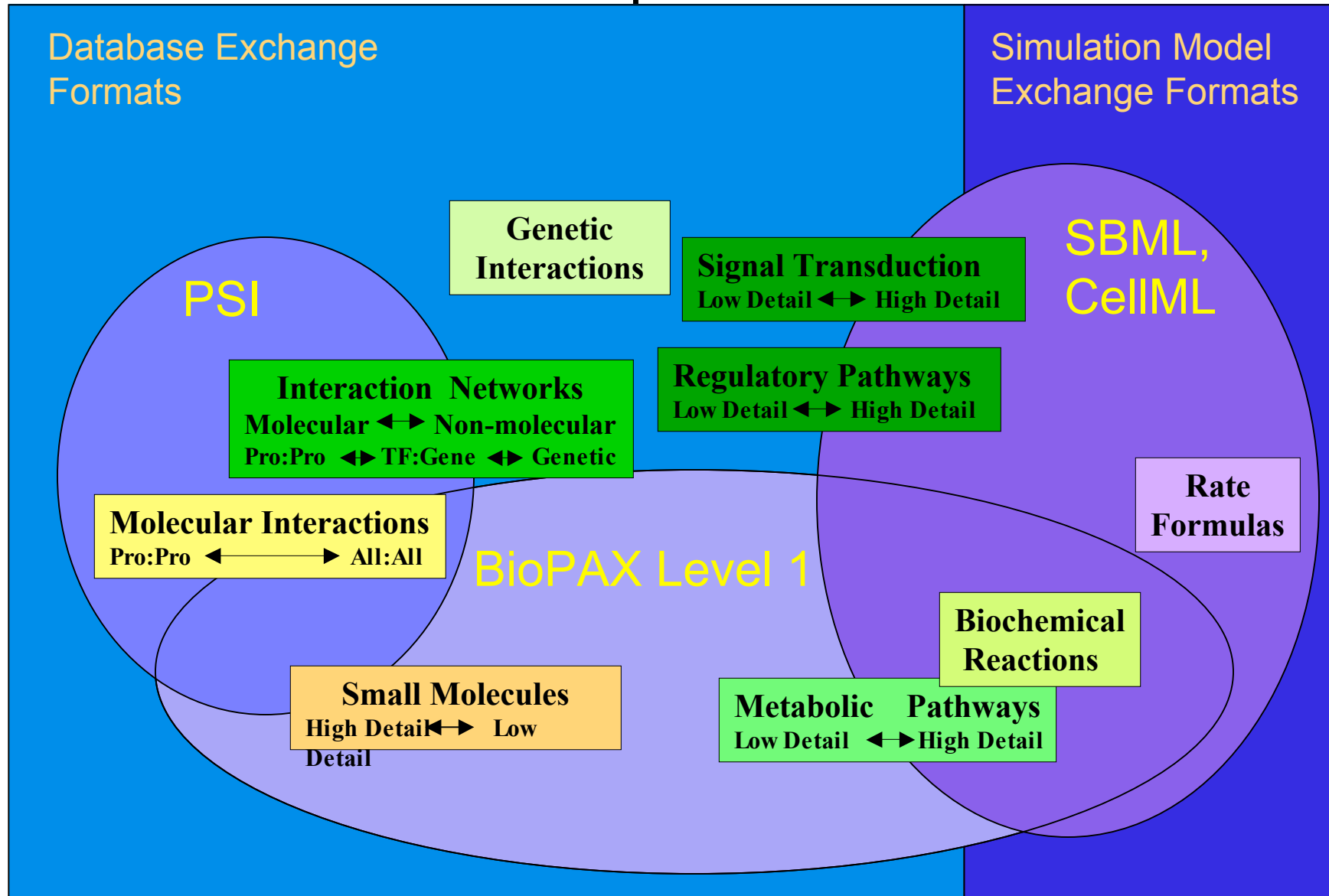
Model Integration in SBML Using the BioPAX Ontology

Jeremy Zucker and the rest of the
BioPAX Workgroup

BioPAX Goals

- BioPAX = Biological Pathway Exchange
- Ontology for representing pathway knowledge:
 - Metabolic pathways
 - Signaling pathways
 - Protein-protein interactions
 - Genetic regulatory pathways
- Abstract data model for integrating databases such as BioCyc, BIND, WIT, aMAZE, KEGG, etc. over 120 pathway databases
- Interoperate with existing standards (RDF/OWL, LSID, **SBML**, PSI, CellML Metadata Standard)

Scope of BioPAX Level 1 in the Pathway Data Space



How can BioPAX help SBML?

Addresses nasty data integration issues

- Different data types, same representation
- Same data, different representations
- External references...
- Synonyms...
- Provenance...

How can SBML help BioPAX?

Modeling and Simulation with BioPAX data

- Represent Kinetic Parameters
- Systems Biology Ontology (SBO)
- Wide software support
- Large user community

BioPAX:XML hacks resolved

- Sid, name, rdfs:label, rdf:id
- Everything is in a global namespace
- Sbml merge is automagically accomplished by any compliant rdf tool
- SBML UML model directly translates to OWL.
- Automatic code generation from spec.

Different data types, same representation

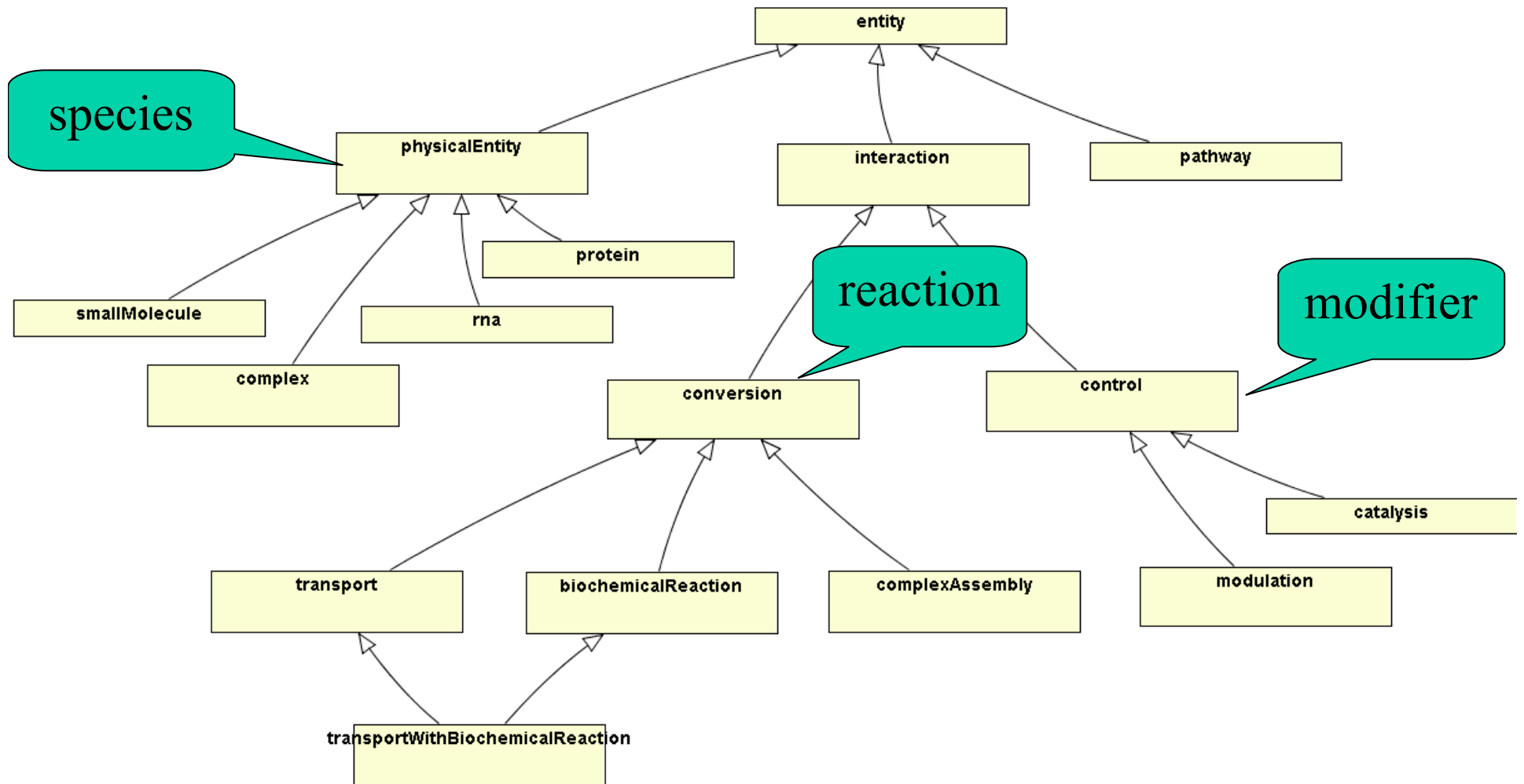
Protein-Protein Interaction

```
<reaction
  id="pyruvate_dehydrogenase_cplx
"/>
<listOfReactants>
  <speciesRef species="PdhA"/>
  <speciesRef species="PdhB"/>
</listOfReactants>
<listOfProducts>
  <speciesRef
    species="Pyruvate_dehydrogenase
      _E1"/>
</listOfProducts>
</reaction>
```

Biochemical Reaction

```
<reaction
  id="pyruvate_dehydrogenase_rxn"/>
<listOfReactants>
  <speciesRef species="NADP+"/>
  <speciesRef species="CoA"/>
  <speciesRef species="pyruvate"/>
</listOfReactants>
<listOfProducts>
  <speciesRef species="NADPH"/>
  <speciesRef species="acetyl-CoA"/>
  <speciesRef species="CO2"/>
</listOfProducts>
<listOfModifiers>
  <modifierSpeciesRef
    species="pyruvate_dehydrogenase_E1"
  />
</listOfModifiers>
</reaction>
```

BioPAX Ontology: Overview



Level 1 v1.0 (July 7th, 2004)

Linking SBML to BioPAX

```
<sbml xmlns:bp="http://www.biopax.org/release1/biopax-
  release1.owl"
      xmlns:owl="http://www.w3.org/2002/07/owl#"
      xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
<!DOCTYPE rdf:RDF [
  <!ENTITY biopax "http://www.biopax.org/examples/glycolysis#">]>

<listOfSpecies>
  <species id="PdhA" definitionURL="&biopax;PdhA" />
  <species id="NADP" definitionURL="&biopax;NADP" />
</listOfSpecies>
<listOfReactions>
  <reaction id="pyruvate_dehydrogenase_cplx"
definitionURL="&biopax;pyruvate_dehydrogenase_cplx" />
  </reaction>
  <reaction id="pyruvate_dehydrogenase_rxn"
definitionURL="&biopax;pyruvate_dehydrogenase_rxn" />
```

Linking SBML to BioPAX:

Protein-Protein Interaction

```
<reaction
  id="pyruvate_dehydrogenase_cplx"
  definitionURL="&biopax;pyruvate_dehydrogenase_cplx"/>
<listOfReactants>
  <speciesRef species="PdhA"
    definitionURL="&biopax;PdhA"/>
  <speciesRef species="PdhB"
    definitionURL="&biopax;PdhB"/>
</listOfReactants>
<listOfProducts>
  <speciesRef
    species="Pyruvate_dehydrogenase_E1"

    definitionURL="&biopax;pyruvate_dehydrogenase_E1"/>
</listOfProducts>
</reaction>
```

Biochemical Reaction

```
<reaction
  id="pyruvate_dehydrogenase_rxn"/>
<listOfReactants>
  <speciesRef species="NADP+"/>
  <speciesRef species="CoA"/>
  <speciesRef species="pyruvate"/>
</listOfReactants>
<listOfProducts>
  <speciesRef species="NADPH"/>
  <speciesRef species="acetyl-CoA"/>
  <speciesRef species="CO2"/>
</listOfProducts>
<listOfModifiers>
  <modifierSpeciesRef
    species="pyruvate_dehydrogenase_E1"
  />
</listOfModifiers>
</reaction>
```


BioPAX Ontology: Root

- Root class: Entity
 - Any concept referred to as a discrete biological unit when describing pathways. This is the root class for all biological concepts in the ontology, which include pathways, interactions and physical entities

entity	
0	SYNONYMS
0	COMMENT
0	DATA-SOURCE
0	SHORT-NAME
0	AVAILABILITY
0	NAME
0	XREF

- "a decade of experience with scalable identifier systems suggests that using arbitrary strings sucks dead puppy dogs tails." - Tim Berners-Lee,
<http://lists.w3.org/Archives/Public/xml-uri/2000Jun/0804.html>

BioPAX: External References

```
<species id="pyruvate" metaid="pyruvate">
<annotation xmlns:rdf="http://www.w3.org/TR/>
  <rdf:RDF xmlns:bp="http://biopax.org/release/biopax-level1.owl">
    <bp:smallMolecule rdf:ID="#pyruvate">
      <bp:Xref>
        <bp:unificationXref rdf:ID="#unificationXref119">
          <bp:DB>LIGAND</bp:DB>
          <bp:ID>c00022</bp:ID>
        </bp:unificationXref>
      </bp:Xref>
    </bp:smallMolecule></rdf:RDF>
  </annotation>
</species>
```

BioPAX: Synonyms

```
<species id="pyruvate" metaid="pyruvate">
<annotation xmlns:bp="http://biopax.org/release1/biopax_release1.owl"/>
<rdf:RDF>
  <bp:smallMolecule rdf:ID="#pyruvate" >
    <bp:SYNONYMS>pyroracemic acid</bp:SYNONYMS>
    <bp:SYNONYMS>2-oxo-propionic acid</bp:SYNONYMS>
    <bp:SYNONYMS>alpha-ketopropionic acid</bp:SYNONYMS>
    <bp:SYNONYMS>2-oxopropanoate</bp:SYNONYMS>
    <bp:SYNONYMS>2-oxopropanoic acid</bp:SYNONYMS>
    <bp:SYNONYMS>BTS</bp:SYNONYMS>
    <bp:SYNONYMS>pyruvic acid</bp:SYNONYMS>
  </bp:smallMolecule>
</annotation>
</species>
```

Complementary approaches

- Semantic validation
- Semantic consistency checks
- Modularization of SBML
- Level 3 features
- Controlled vocabularies
- Other ontologies

Coming soon...

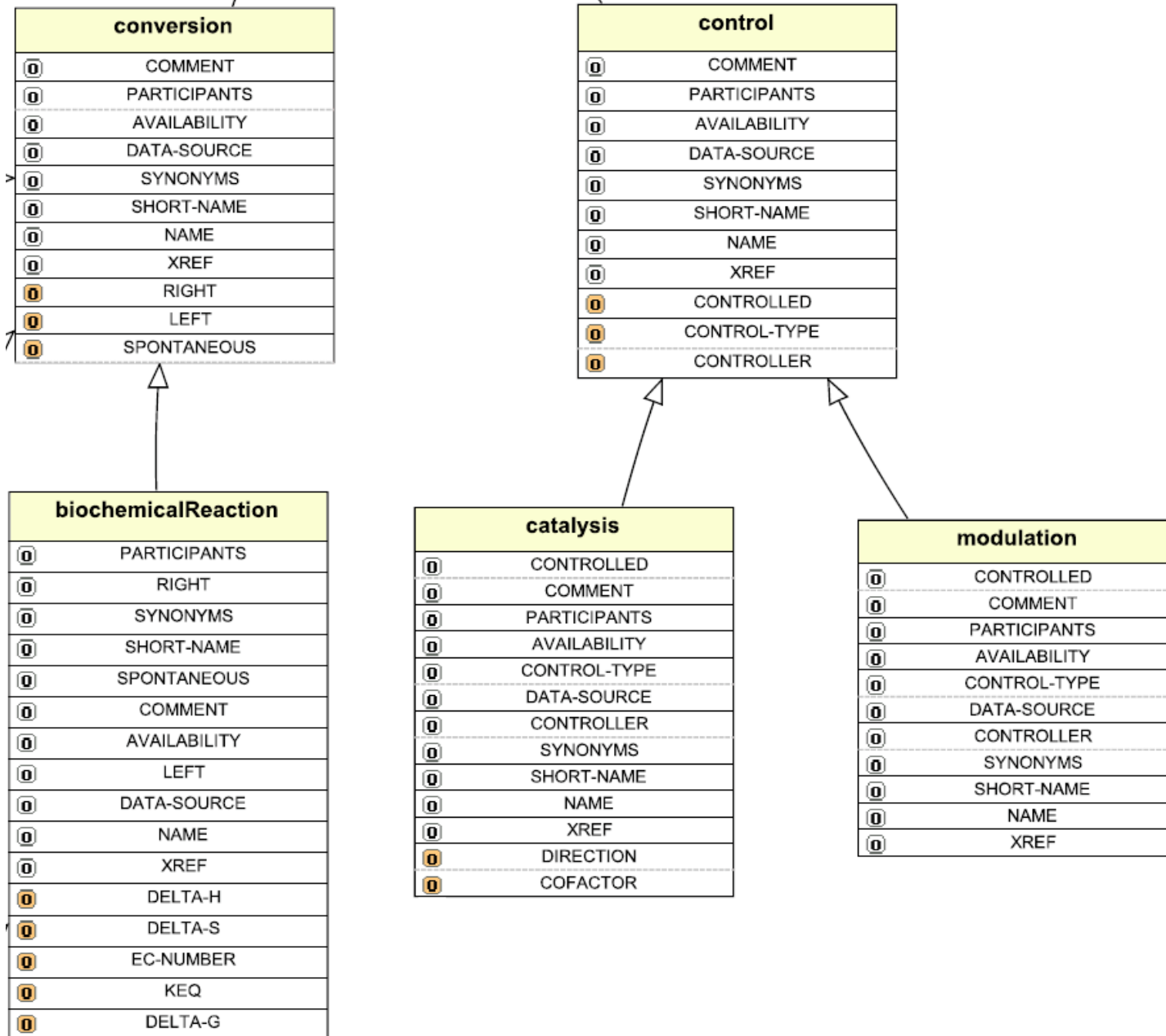
- Biocyc, KEGG, WIT, aMAZE => BioPAX
- libBioPAX (apologies to Ben)
- BioPAX 2 SBML
- BioPAX 2 BioWarehouse
- BioPAX hackathons, Forums, conference calls

To join: `biopax-discuss@biopax.org`

Lessons Learned:

- “A good representation is the key to good problem solving”—Patrick Winston, *Artificial Intelligence*
- “Standard is Better than Best”—Gerald J. Sussman
- “The great thing about standards is that there are so many from which to choose.”—Unknown
- “Six weeks in the laboratory can save you six minutes at the computer” –Tom Knight
- “Above all, one must have a feeling for the organism”—Barbara McClintock

BioPAX Ontology



Same data, Same model

- Different SBML representation
- Modules UML-based

Same model, different data

Same data, different model

- BioPAX solution:
- `<reaction id=“`

Different data, different model

- BioPAX solution