Gaps in Pathway Representation

Alan Ruttenberg

Science Commons

Overall Design

- Seamless participation in OBO Foundry: **BioPAX-OBO**
  - Reuse/Advance: RO, ChEBI, GO, ECO, OBI, FMA, PATO, MP, ENVO
  - Fostering working with projects like BioCYC
- Identity – New entity versus same entity with new quality
- Have a better story about what a “Pathway” is (e.g. A selective view of a process and its participants. A query result.)
- Clearly separate user interface/ data entry/ exchange/ computational access vs. **representation of biology**
- Declarative representations wherever possible
- “Good enough to be wrong” (use a reasoner)

Entities

- Rid us of “small molecule” crutch
  - Represent energy, electrons, voltage, light
- Reaction granularity – Reaction participants have atomic level changes, kinetics, messaging function. (occurrents)
- Material granularity – from electrons to ion channels, molecules and mixtures. (continuants)
- Better ability to talk about parts as participants (e.g nicked DNA, domains, axons)
- Experiments / generalizations about processes (type level) – When to represent each?
- Clarify important processes/relationships e.g. “Regulates”, “Causes”, “Next”

Gaps (from discussions)

- Do we benefit from representing physical interaction of regulatory interactions (Richard)
- Generalization versus Uncertainty (Nancy)
- What is real about pathways? (Emek, Barry, Nigam) What are the boundaries (Richard)
- Understanding distinctions between function/role (Chris, Richard)
- INOH / BioPAX mapping (Richard)
- Often/Sometimes (Erick)
- Is inconsistency incompatible with Ontology (Andrey)
- The ontology of old stuff that is understood to be wrong (Andrey).
- Parasite/Host, developmental processes, disease processes, cellular differentiation… (Gopal)
- Importing databases require human intervention (knowledge gap) (Gopal)
- Every reaction in a pathway? (Gopal)

Gaps (from discussions)

- Pathways are models. But in what formalism? (Nigam)
- Unique identifiers for the entities in question (Nigam)
- “Reference Network” (Nigam)
- BioPAX/community collaboration (Chris)
- Reconciling information from different databases (Chris)
- How to update information (Chris)
- Capture of information at source (Journals) (Chris)

Three levels of representing scientific knowledge

- *Record level*: Represent database records. Inconsistent if two sources disagree about contents of a field.
- *Statement level*: Represent what researchers say. Inconsistent if two people disagree about what a paper said
- *Domain level*: OBO Foundry approach. Represent your best understanding of consensus. Inconsistent if facts contradict.
- We need all three (but make clear which is which)

Don't inadvertently mix levels