

BioPAX Group
October 3, 2003 Meeting Minutes

Location: Hilton Garden Inn, Denver, CO.

Participants: Gary Bader, Michael Cary, Peter Karp, Joanne Luciano, Suzanne Paley, Imran Shah

Note: The majority of this meeting was dedicated to discussing the features to be included in each level (with the main focus being Level 1) of BioPAX. The product of the discussion was a roadmap, a summary of which is available here:

http://www.biopax.org/Docs/BioPAX_Roadmap.html

The complete roadmap (Excel spreadsheet) is available here:

http://cvs.sourceforge.net/viewcvs.py/biopax/biopax/doc/BioPAX_Roadmap.xls

Overview:

- I. Discussed open issues
 - a. Representing physical attributes on entities
 - b. Representing complex processes
- II. Discussed scope of Level 1
 - a. Minimally: metabolic data sets
 - b. Include proteomics data in Level 1?
- III. Developed roadmap
 - a. More examples, with better documentation
 - b. GKB → OWL converter
 - c. Formal specification document
 - d. Data converters for a number of DBs
- IV. Next meeting
 - a. Tentative: Mid-February in San Francisco

Summary:

The meeting began with a discussion of two major open ontology issues: a) How do we represent (often variable) attributes on physical entities, such as location and post-translational modifications, and b) How do we represent common complex processes (e.g. transcription, translation).

[Note: Full descriptions of these issues can be found on the BioPAX SourceForge Bug Tracker web page: http://sourceforge.net/tracker/?group_id=85345&atid=575904]

The discussion on these issues revealed the existence of a larger, more pressing issue: the core group lacked consensus on the scope (i.e. range of data coverage) of BioPAX Level

1. We therefore decided to dedicate most of the remainder of the meeting to reaching such a consensus and focusing on the steps needed to generate the Level 1 release.

The group quickly agreed that, at the minimum, Level 1 should be able to represent the majority of the data in the major metabolic databases (e.g. BioCyc, KEGG, WIT). However, whether or not proteomics data, (e.g. protein-protein interactions from DIP, BIND, or MINT) should also be included in Level 1 was much more difficult to agree upon. For the time being, we decided to work toward finishing a version of Level 1 that does not include protein-protein interactions; Mike and Gary will work toward incorporating these after the other major tasks for Level 1 have been completed. It remains to be seen whether proteomics data will be part of Level 1 or will need to be postponed to a later release.

After agreeing on the minimal scope of Level 1, we discussed the concrete steps we need to take in order to generate the Level 1 release. We agreed that we should make better use of examples, done in GKB, to guide the development. Joanne and Mike will translate the set of examples done in Protégé into GKB. These will then be used to evaluate and refine a pre-release (version 0.9).

Peter agreed to work on a GKB → OWL converter, so that in the future users could use either ontology editor (Protégé or GKB). Imran and Gary agreed to begin working on a formal specification document (in the W3C style), which will accompany the Level 1 release. We also decided to develop data converters for, minimally, BioCyc and WIT by the time Level 1 is released. This will provide a concrete demonstration of the utility of BioPAX.

The above decisions were incorporated into a document that we called the BioPAX Roadmap; this document will be revisited and revised (if necessary) during future meetings and conference calls. We tentatively agreed to hold the next meeting in mid-February at SRI in San Francisco.