Meeting Minutes 6 April 2004

Autumn A Cuellar

This was a special meeting to give the Sheffield Group a clearer picture of how CellML is developing. Dawn was interested in whether CellML can be extended to describe her research which occurs more at the tissue level, while Steve was curious about more technical issues such as change control and bug fixes. The questions raised may provide information for other interested parties. I've bolded some of the key terms and themes.

- Poul kicked things off by distributing a copy of the Road Map [../tool_development/projects.html].
- Rod Smallwood explained that they submitted a grant application for the development of cell models and **TissueML** - 1 person for 3 years. They are not yet certain if the money will come through, but if it does (by June 1), they intend for the person funded by the grant to coordinate with Team CellML.
- Poul drew attention to some of the main points discussed in the Road Map document:
    - We're moving away from the concept of CellML being limited to the description of cell models. Because of its more generalized application, **ModelML**, a modularized language, is evolving.
    - **Reactions** were included in CellML grudgingly. Reactions are a specific domain of knowledge that (like others) should be contained in **ontologies** into which ModelML can be hooked.
    - Another specific domain we've been working on is an **anatomical ontology**. Dawn: “Is that on the website because I've been trying to piece together something similar?” Matt: “No, it can only be accessed internally at the moment.”
    - SBML [http://sbml.org/] and CellML now have MathML and Metadata in common.
- Matt: “One reason we want to move the **reaction** part of CellML into **ontologies** is that other groups (such as **BioPAX** [http://www.biopax.org/]) have better descriptions of the qualitative information regarding biochemical pathways. There is no reason to recreate something that's already been done fairly well. This is where typing information comes into play.”

  “The **anatomy ontology** can be used to find mathematical models and vice versa.”

  Dawn: “Are BioPAX models freely available?”

  Matt: (explaining) “**BioPAX** describes the qualitative aspects of a pathway, information about species and reactions. It does not describe the mathematics. There is currently no centralized database of BioPAX models. You might be interested in checking out PATIKA [http://www.patika.org/], a global shared repository (can be stored locally or centrally). The PATIKA group is heavily involved in the development of BioPAX.”

  “Eventually we expect to make all of the CellML pathway models instances of BioPAX.”
- Dawn: “Where does **FieldML** [http://www.physiome.org.nz/fieldml/pages/] fit in?”

  Matt: “That's a question for Shane, but the plan is to integrate FieldML models with the **anatomical ontology**. For instance, you could query for the FieldML representation of the tibia.”
- Rod: “How do you handle **validation**?”

  Poul: “There's validation at different levels:”

- Syntax - tools [../../public/tools/index.html] available.
- Units - no tools yet.
- Model itself - tested through simulation; some tools [../../public/tools/index.html] available.
- Results - compared with the experimental results; responsibility of the model builder.
- Maths - content MathML can be translated so one-to-one comparisons could be made with the journal article.

At this point, the discussion went off on a tangent about the role and definition of models sparked by Rod's comment, “The validation of the model answers the question ‘Is this model specification doing what I specified it to do?’ *rather than* ‘Is this model a faithful representation of reality?’” Though this discussion was interesting, it wasn't entirely relevant, therefore, I didn't try to capture it in my notes.

- Steve: “How is CellML going to be managed for **change control** and **bug fixes**?”

  Poul & Matt: “CellML is at a point where it needs to be released so that community can drive its direction. 1.1 is at a stable point. We have a UML description, DTDs, and XML Schemas available. We just need to kick the process off.”

  Matt: “There will be [three] communities that overlap: The tool development community (bug trackers, cvs repository, platform testing), a community for the development of the specifications, [and the user community interested in building and running the models]. We are close to having this all in place. See the roadmap.”

- Matt: (addressing Dawn) “Regarding the modelling of physical conditions, we need to get people like you involved.” CellML core (ModelML) + ontological layer focussed on tissue-level.

- Me: “In which direction is **TissueML** headed?”

  Dawn: “No one's decided yet.”

  Poul: “We'd like to get your feedback on everything we've talked about. See if it will suit your needs.”

  Dawn: “I'm not sure if some of the **epitheliome** processes can be described by a markup language. **Describing the structures** my model generates would be a good start. The ML would describe a template for the general structure, then we could create specific instances.”

- Rod: “How did you develop CellML? How did you know what to put in?”

  Poul: “We had a set of requirements aimed at the cellular level. It turns out that the solution we came up with has a more general applicability.”

- Dawn: “I've written down the parameters I'd need for the description at the bladder tissue level (components and variables). I will go over these with Matt.”

  Matt: “An ontology is a nice place to start to define what you need to describe.”

- Rod: “The point of view from which we approach this is important. **Structure** first or **function** first?”

  Matt: “Use TissueML to describe states and EmbryoML to describe what these states mean? You might be interested to hear what UMLS [http://www.nlm.nih.gov/research/umls/] have been doing.

They are bringing together databases to address how we model developmental processes. They define two databases together and sort out conflicts. The Digital Anatomist project (describing structure/function together) sprung from this."

[Author's aside: my notes on this conversation were extremely slim (as in, key words only) because I'd run out of paper and was notating in the margins. Matt later fleshed out the missing bits in an e-mail as follows.] Rod talked about structure, function, and then transformations, such as those observed over development. The nature of my comments on structure and function were as follows...

- The Foundation Model Explorer (FMA) [http://sig.biostr.washington.edu/projects/fm/AboutFM.html] of the Digital Anatomist project started off with structure, and then bound this to a functional terminology.

- The FMA also seeks to include embryology (from the link above "We have proposed to enhance the FMA with concepts and relationships of classical embryology and developmental biology through the Anatomical Transformation (ATA) component of the FMA").

- The FMA is part of the UMLS integration project (again from the FMA link - "The Anatomy taxonomy (At) component of the FMA has been incorporated in the National Library of Medicine's Unified Medical Language System (UMLS) as one of its constituent vocabularies. In UMLS the FMA is known as the University of Washington Digital Anatomist (UWDA) vocabulary. The UWDA consist of the At and selected structural relationships (part-of, branch-of, tributary of). The UWDA can be accessed through the UMLS knowledge server").

- The UMLS group are the ones trying hard to integrate multiple databases from many areas of biology. The following is a summary of their work, "The Unified Medical Language System [http://umlsks.nlm.nih.gov] is a repository of biomedical vocabularies developed by the US National Library of Medicine. The UMLS integrates over 2 million names for some 900 000 concepts from more than 60 families of biomedical vocabularies, as well as 12 million relations among these concepts. Vocabularies integrated in the UMLS Metathesaurus include the NCBI taxonomy, Gene Ontology, the Medical Subject Headings (MeSH), OMIM and the Digital Anatomist Symbolic Knowledge Base. UMLS concepts are not only inter-related, but may also be linked to external resources such as GenBank. In addition to data, the UMLS includes tools for customizing the Metathesaurus (MetamorphoSys), for generating lexical variants of concept names (lvg) and for extracting UMLS concepts from text (MetaMap). The UMLS knowledge sources are updated quarterly. All vocabularies are available at no fee for research purposes within an institution, but UMLS users are required to sign a license agreement. The UMLS knowledge sources are distributed on CD-ROM and by FTP."