CellML 2010 Workshop
Model annotation:
the requirements of a CellML 1.1 model

Poul Nielsen

Modelling complex systems

- Fundamental ways we cope with complexity:

- Abstraction

  - …denotes essential characteristics of an object that distinguishes it from other kinds of objects, providing well-defined conceptual boundaries.

- Encapsulation

  - …is the process of compartmentalising the elements of an abstraction that characterise its structure and behaviour.

  - Separates the interface of an abstraction from its implementation.

- Modularity

  - …is the property of a system that has been decomposed into a set of cohesive and loosely coupled modules.

- Hierarchy

  - …is a ranking or ordering of abstractions.

CellML 1.0 capabilities

- Design goal of CellML was to aid abstraction, encapsulation, modularity, and hierarchy through the language elements.

- The elements used by CellML 1.0 for this include:
  - The use of **MathML** to represent mathematical relations (high-level abstraction of mathematical concepts, independent of algorithms);
  - **Units** are required for all variables (encapsulation of units and variables enables independent construction of components);
  - **Components** organise concepts into manageable units (functional abstraction);
  - **Interfaces** provide well-defined communication to external components (encapsulation via contractual interface and information hiding);
  - **Public/private** interfaces enable organised structuring of information (hierarchy of components).
  - **Models** provide a mechanism to specify a unique universe.

AUCKLAND
BIOENGINEERING INSTITUTE
THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

CellML 1.1 capabilities

- CellML 1.1 added capabilities for organising hierarchies of models by introducing an **import** element.

- Any model, or part of a model, can now be imported into a new model and reused as a template.

- Name clashes are avoided by enabling renaming of components and units.

- The same model may be imported many times to create independent instances.

- A powerful mechanism for reusing models as library modules.

- How can CellML elements be most effectively used to create reusable models?

Structuring models

- There are many ways that models can be organised.

- Process-oriented approaches are useful when we are attempting to understand causal links between entities and processes (e.g. signalling cascades, reaction pathways).

- Entity-oriented approaches are useful when we are interested in the participation of particular entities in multiple processes (e.g. ATP interacting with energy producing/consuming reactions).

- A tension exists between models structured to hide information (encapsulation by minimising coupling) and those organised to expose entities (providing flexibility through complex interface).

Biological models…

- Biological systems are complex.

- At most scales there is a clear compartmentalisation of anatomical structures, providing a well-defined separation of physical entities.

- Within compartments, there are generally few barriers to interaction between entities.

- There is no distinction between the same species involved in different functions within the same compartment.

- At this level, the process-oriented organisation of models reflects a human-centric view of organising complexity.
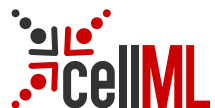
…Biological models

- Because entities provide the common unit of interaction within a compartment, and between compartments, entity-oriented models present a natural way of organising biological systems in a reusable form.

- In order to be reusable the entities at the interface of a component should be tagged with sufficient information to identify the specialisation appropriate to the level of abstraction represented by that component.

- Apart from units, in CellML there are few rigorous mechanisms to tag entities with such information.

- Annotating entities with metadata seems to offer some hope.

Annotating models…

- Because unambiguous species identification and anatomical placement are fundamental to determine whether entities are functionally equivalent, or can physically interact, we must start annotating models with such information to facilitate reusability.

- Species information can be provided by linking entities with protein/gene ontologies/databases, such as PDB and SBO.

- Anatomical information can be provided by linking entities with anatomical/functional ontologies, such as FMA, SBO, and GO.

…Annotating models

- If we are to create libraries of reusable models it is vital to:
  - Decompose existing models into modules of appropriate abstraction;
  - Annotate interface entities with appropriate species and anatomical ontologies.
- This is a difficult and time-consuming task that requires a thorough understanding of the models being decomposed:
  - Model phylogeny;
  - Experimental conditions;
  - Species/age/health/…
- There will be a tension between abstraction and specificity (as currently exists with units).
- How we complete this task will require community support. How can this be organised and managed?