CellML Evolution

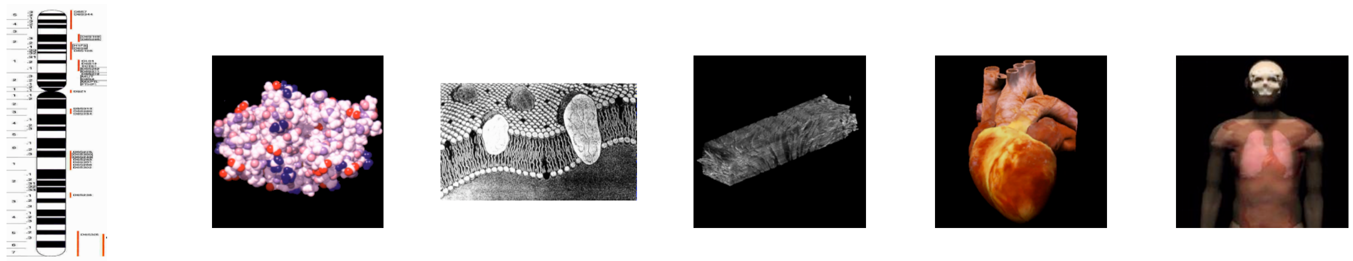
Poul Nielsen, Matt Halstead, Autumn Cuellar,
Michael Dunstan, David Bullivant, Peter Hunter.

Bioengineering Institute, University of Auckland, New Zealand.

Overview

- Coupled models integrate different processes over a wide range of spatial and temporal scales.

Atom	Protein	Cell	Tissue	Organ	Organism & organ system
10^{-12}m	10^{-9}m	10^{-6}m	10^{-3}m	10^{0}m	



Gene networks	Pathway models	Stochastic models	ODEs	Continuum models (PDEs)	Systems models

10^{-6}s	10^{-3}s	10^{0}s	10^{3}s	10^{6}s	10^{9}s

Molecular events (ion channel gating)

Diffusion Cell signalling

Motility

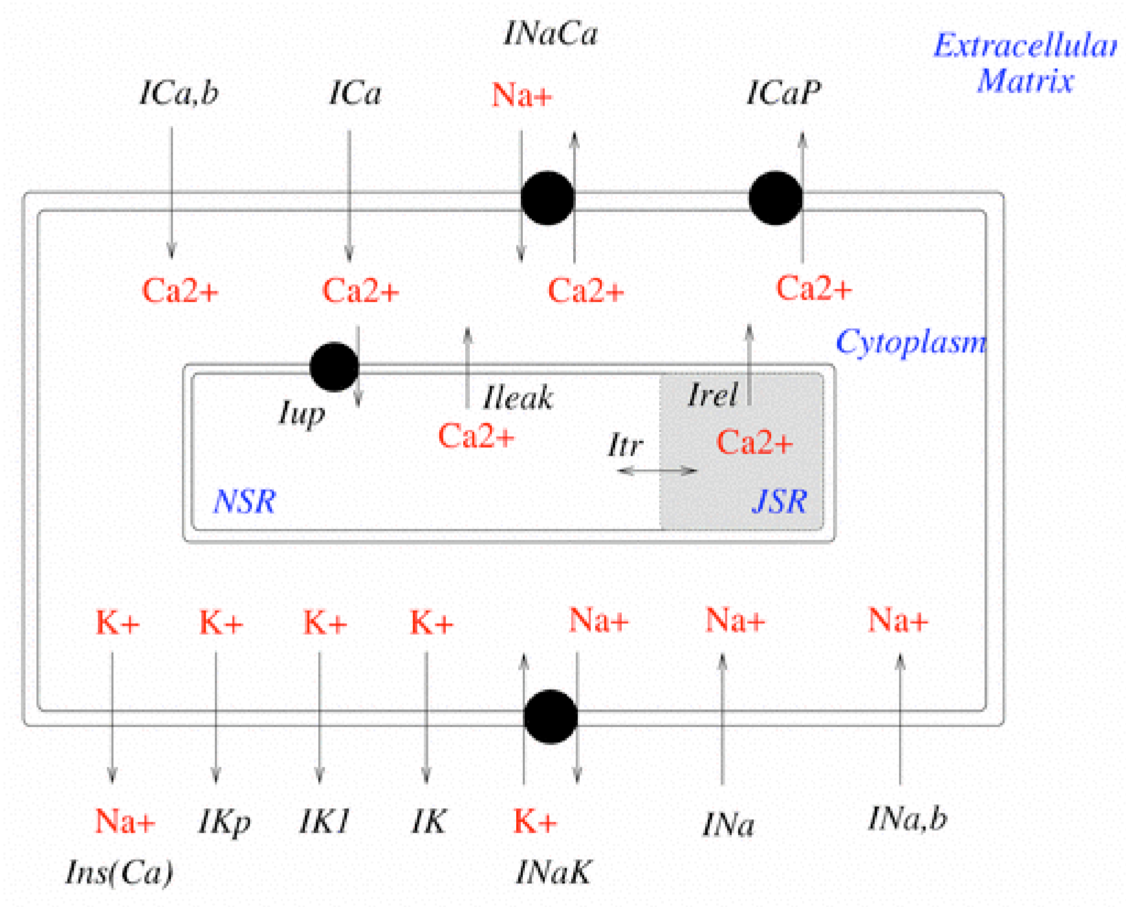
Mitosis

Protein turnover

Human lifetime

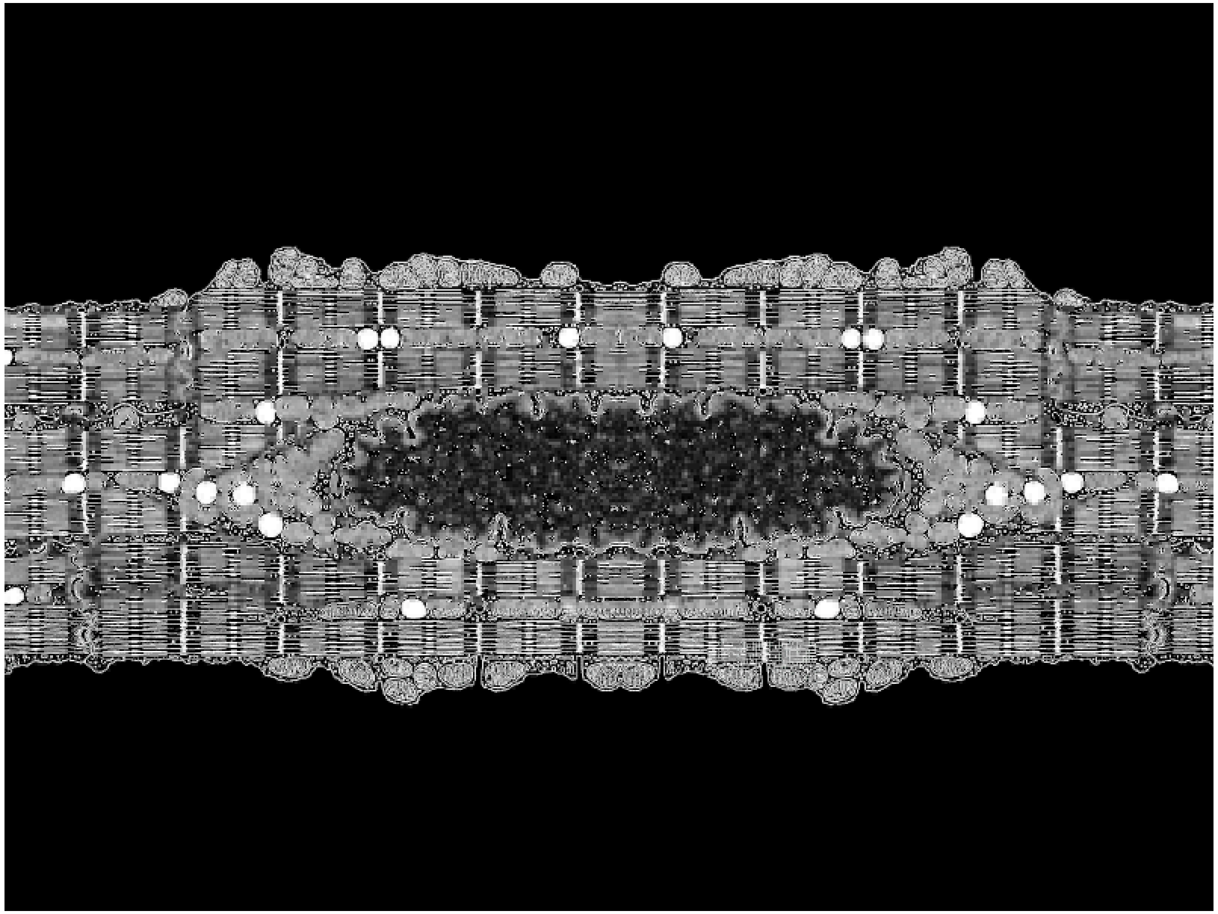
Cellular models…

- e.g. Luo-Rudy II (1994) model of myocardial action potential (~20 coupled nonlinear o.d.e.s)



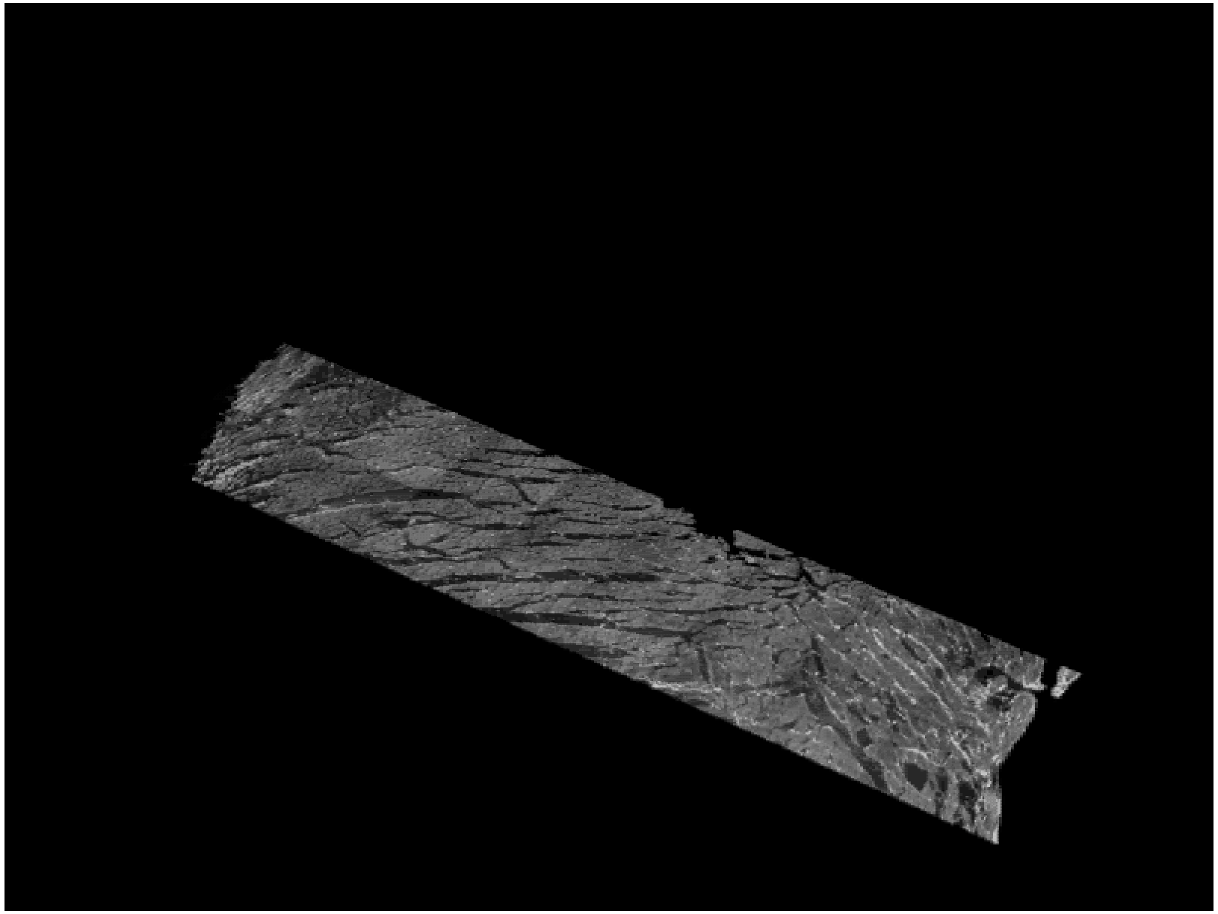
…Cellular models

- e.g. cellular level models of excitation/contraction coupling (small spatial/temoral scales).



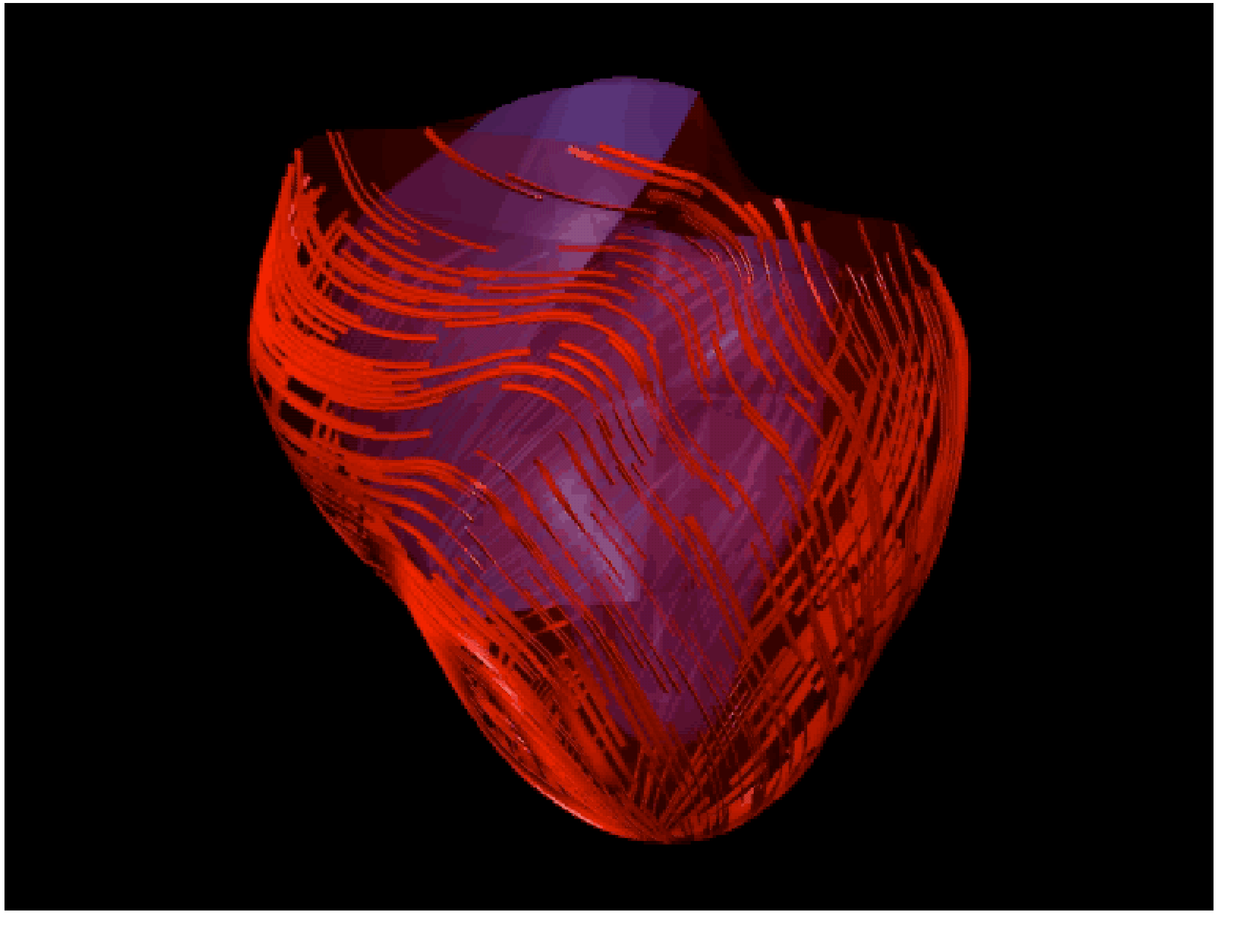
Architectural models…

- e.g. models of myocardial architecture and activation.



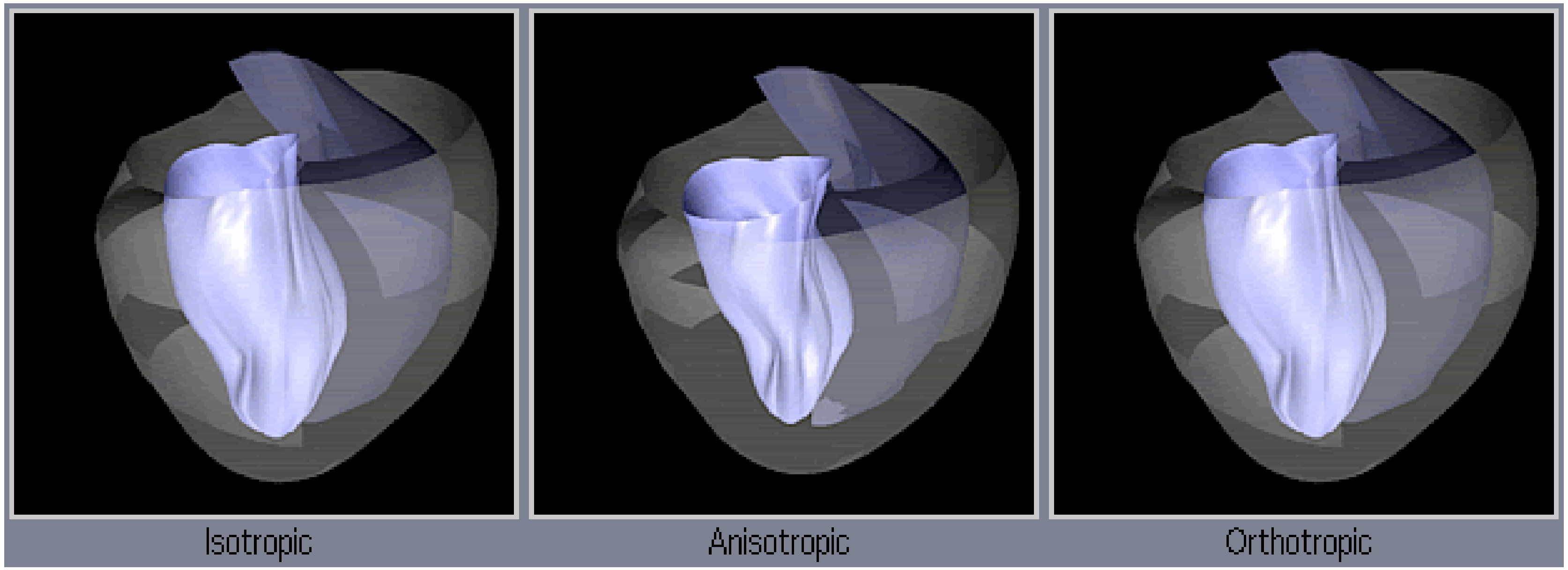
…Architectural models

- e.g. model of myocardial fibre orientation.



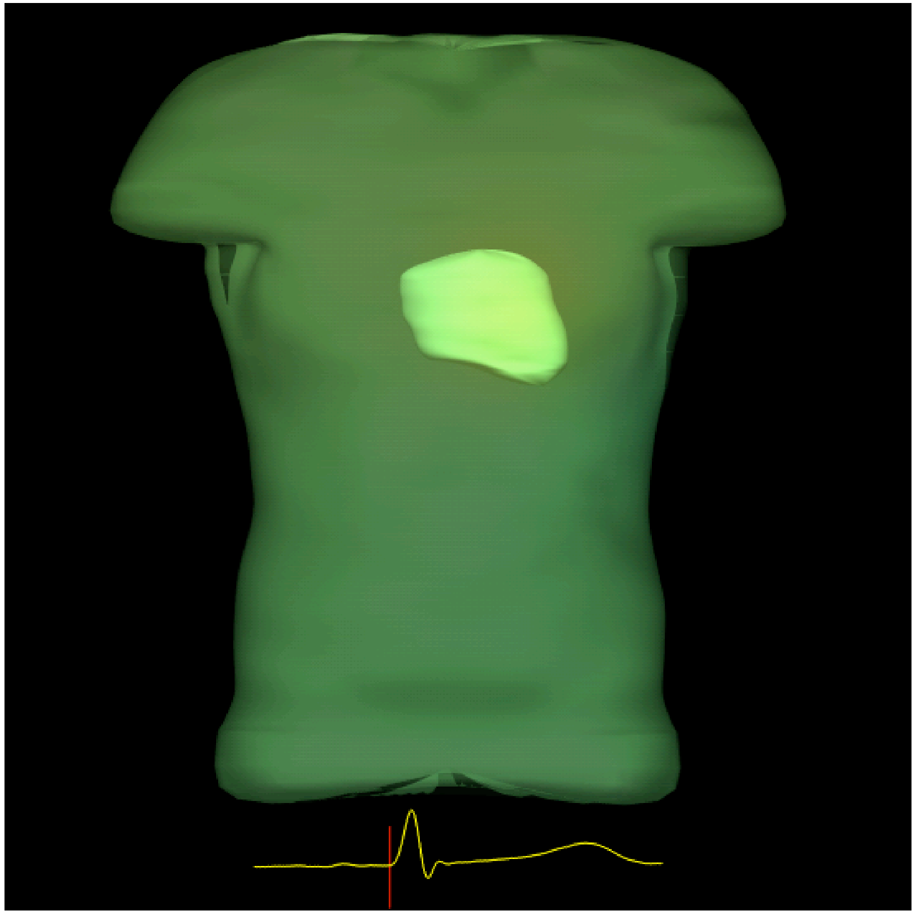
Coupled processes

- e.g. cellular-level activation processes integrated into architectural model to compare patterns of activation through the myocardium with various degrees of muscle orthotropy.



Integration to large-scale model

- e.g. activation at cardiac level inducing electric fields on skin.

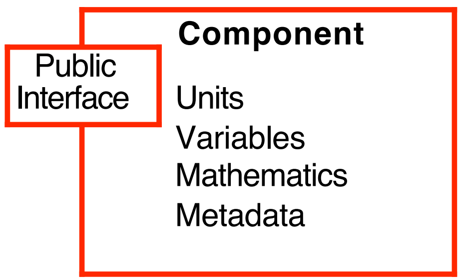


Complexity

- All of these models are complicated, involving many degrees of freedom, coupled physical processes, and wide variations of spatial and temporal scales.
- The choice of what components are primitive is arbitrary (what is primitive for one modeller may be a high level of abstraction for another).
- Any method for representing models should thus enable:
  - *Abstraction* (ignore inessential details);
  - *Encapsulation* (compartmentalize elements for information hiding);
  - *Modularity* (decompose into a set of loosely coupled modules);
  - *Hierarchy* (ranking or ordering of abstractions through class structures).

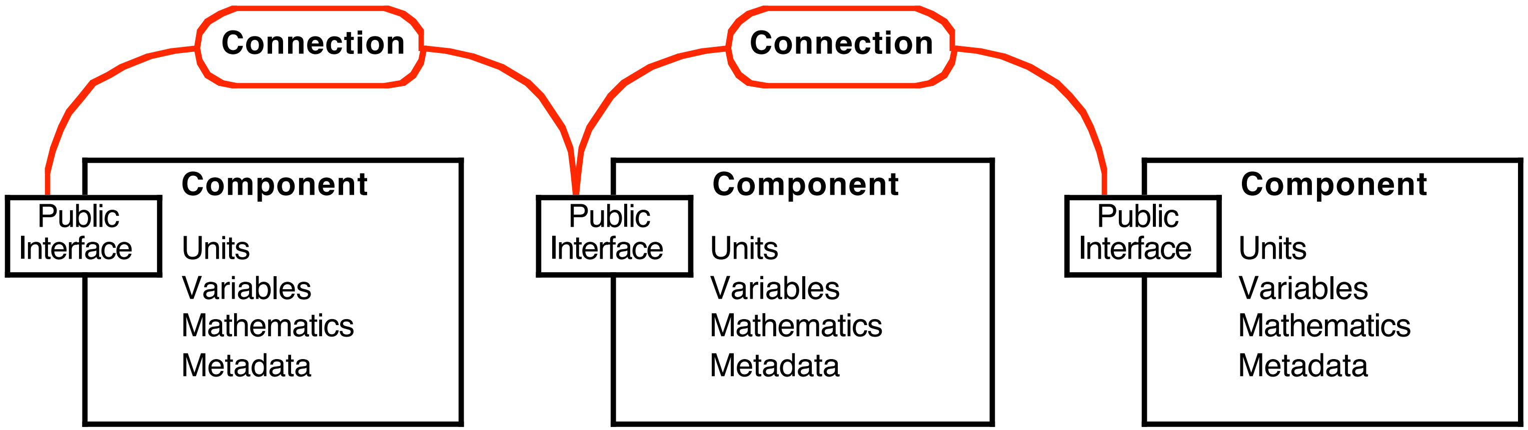
CellML for describing systems…

- CellML has a simple structure based upon connected *components*.
- Components abstract concepts by providing well-defined interfaces to other components.
- Components encapsulate concepts by hiding details from other components.



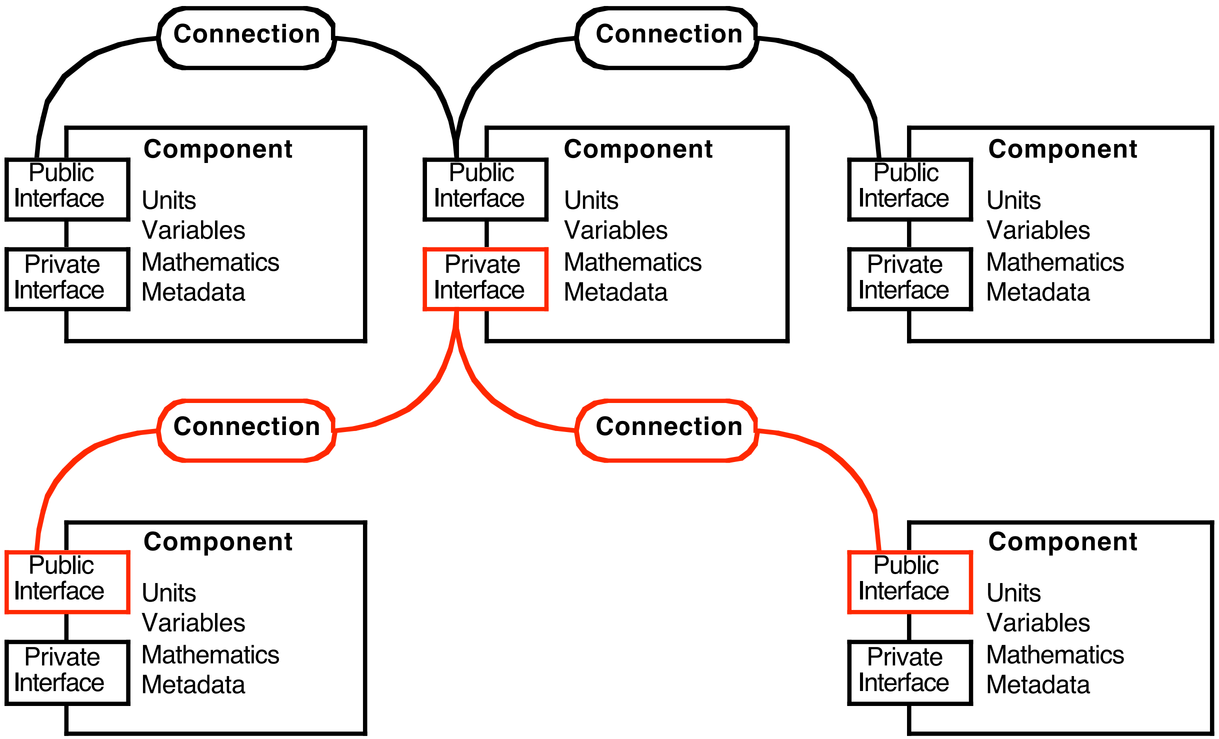
…CellML for describing systems…

- *Connections* provide the means for sharing information by associating variables visible in the interface of one component with those in the interface of another component.
- Consistency is enforced by requiring that all variables be assigned appropriate physical *units*.



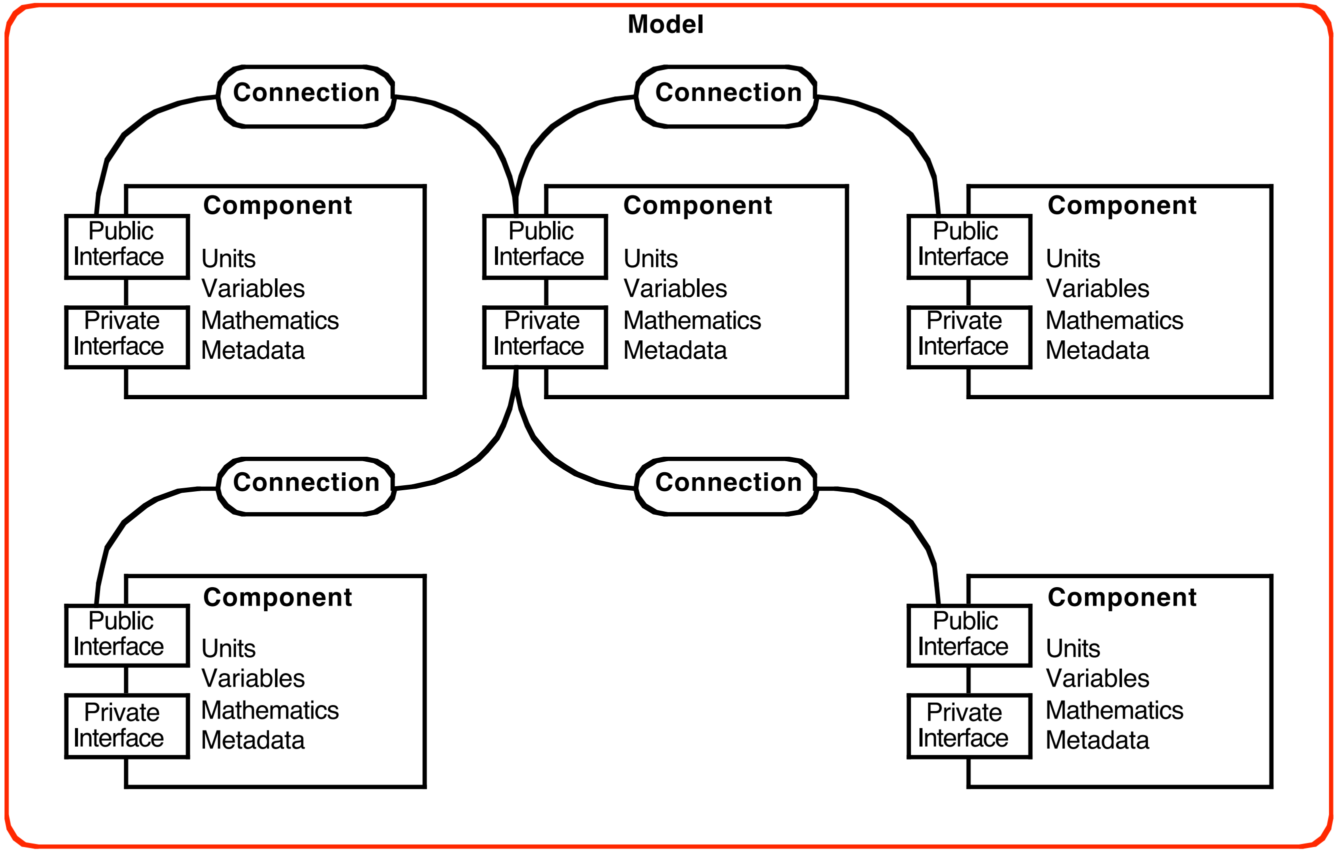
…CellML for describing systems…

- Encapsulation hierarchies are enabled using *private interfaces*.



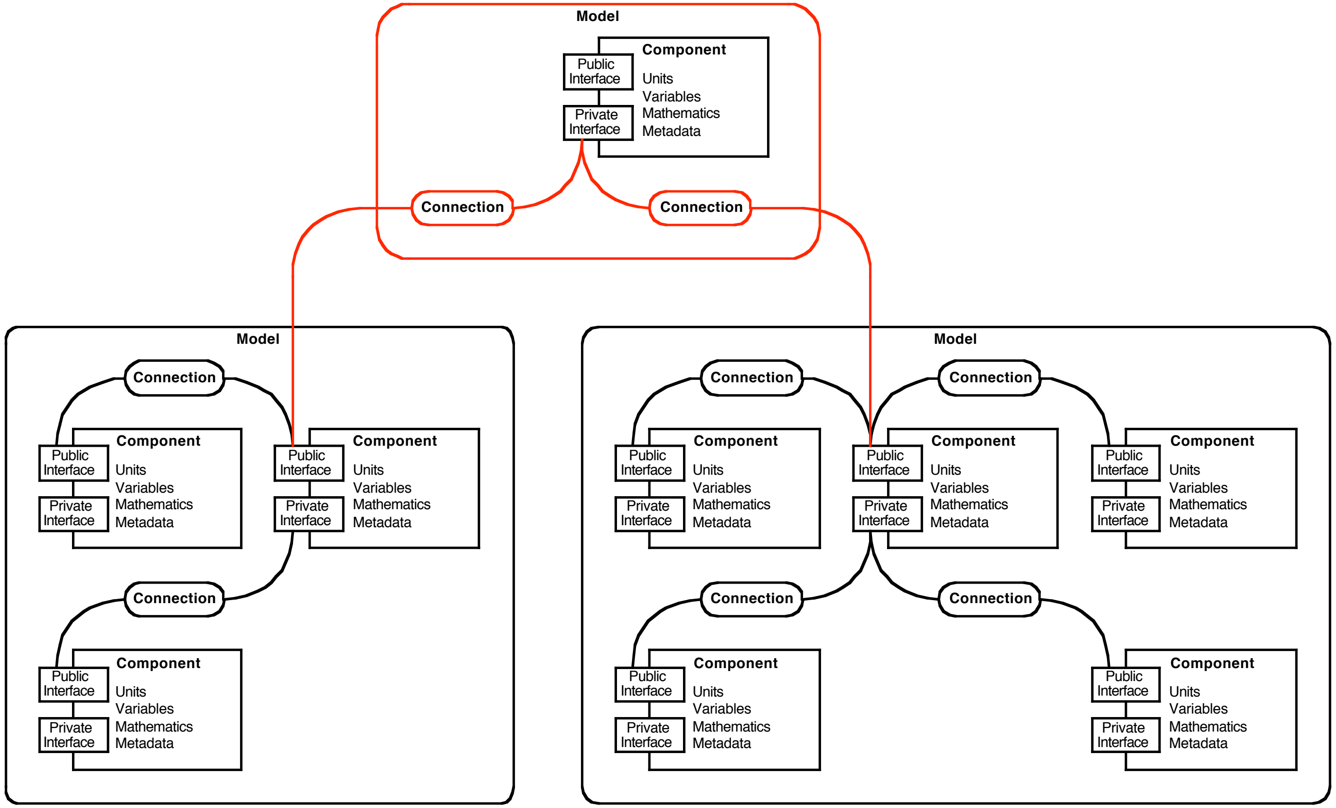
…CellML for describing systems…

- A *model* is the root element for a CellML document. It is a container for components, connections, units, and metadata.



…CellML for describing systems…

- Model reuse is enabled by the *import_model* element.
- New models may thus be constructed by combining existing models into model hierarchies.



…CellML for describing systems

- Geometric inclusion is handled by the use of *containment* hierarchies.
- The *reaction* element is used for both qualitative and quantitative representations of biochemical pathways.
- CellML builds upon widely accepted standards (e.g. XML, MathML, RDF, XLink), defining new concepts only when no others exist (e.g. *metadata* and *units*).
- CellML model repository consists of over 200 models of electrophysiological, mechanical, signal transduction, and metabolic pathway processes.
- **http://www.esc.auckland.ac.nz/www/sites/physiome/cellml/**
  or
  http://www.cellml.org/

Knowledge

- However CellML, like SBML, is a low-level language for describing systems. There are no constructs for typing variables and components or creating class hierarchies.
- Models often carry much implicit information about the processes being described.
- How can we make such knowledge explicit?
- We require mechanisms for defining controlled vocabularies (so we can agree upon terms) and rules for defining allowable behaviour between the things of interest.
- This is the realm of *ontologies* and *description logic*.

Ontology…

- An ontology is a specification of a conceptualization.
- Ontologies are used for the purpose of enabling knowledge sharing and reuse.
- Ontology descriptions, based upon Open Knowledge Base Connectivity, provide a uniform conceptualization of classes, individuals, slots, facets, and inheritance. OKBC protocol transparently supports networked as well as direct access to knowledge representation systems and knowledge bases.

…Ontology

- Ontologies provide:
  - *Classing mechanisims* (multiple inheritance, subclassing, domain, range,…);
  - *Class expressions* (union, intersection, complement,…);
  - *Class axioms* (one of, disjoint with, equivalence,…);
  - *Property characteristics* (transitive, symmetric, functional,…);
  - *Cardinality* (minimum, maximum, exactly);
  - *Rich set of primitive data types* (string, boolean, integer, real, datetime, URI,…);
  - *Management* (imports, versions, compatibility,…).
- Many tools are available to create, modify, make inferences and check assertions about ontologies (e.g. Protégé, OILed, Ontolingua, …)

How are we using ontologies?

- We are creating ontologies for:
  - Formal description of reactions;
  - Specification of CellML (minus units);
  - Partial anatomical and functional characterisation of
    - Organ systems
    - Organs
    - Tissues
    - Cells
    - Organelles
    - Cell function
    - Proteins
    - Protein domains
  - **http://www.bioeng.auckland.ac.nz/physiome/physiome.php**

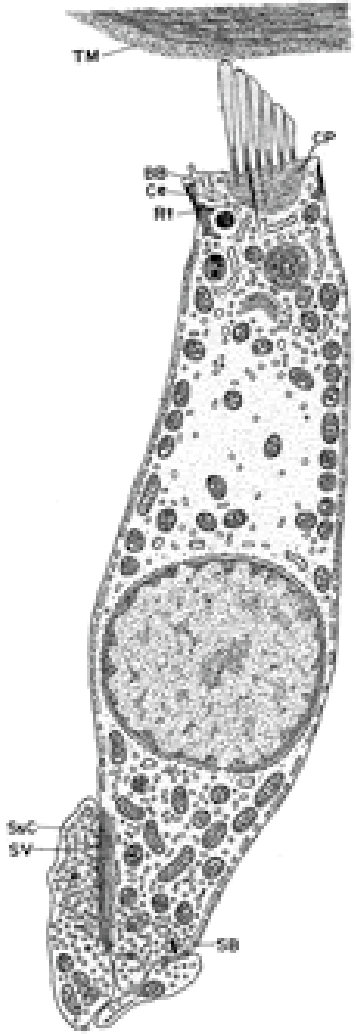
e.g. Cell and organelle ontology interfaces

Physiome Project
Quick Reference (Tree)
Quick Reference (Alphabetical)

- Cells
  - Blood cells
    - Agranulocyte
    - Erythrocyte
    - Erythropoietic cells
    - Granulocytes
    - Granulopoietic cells
    - Haemopoietic stem cell
    - Platelets
  - Bone cells
    - Osteoblast
    - Osteoclast
    - Osteocyte
    - Osteoprogenitor cell
  - Cardiac cells
  - Cartilage cells
  - CNS cells
  - Epidermal cells
  - Gastrointestinal cells
  - Generic cells
  - Immune cells
  - Neural cells
  - Liver cells
  - Pancreatic cells
  - Reproductive cells
  - Respiratory cells
  - Sensory system cells

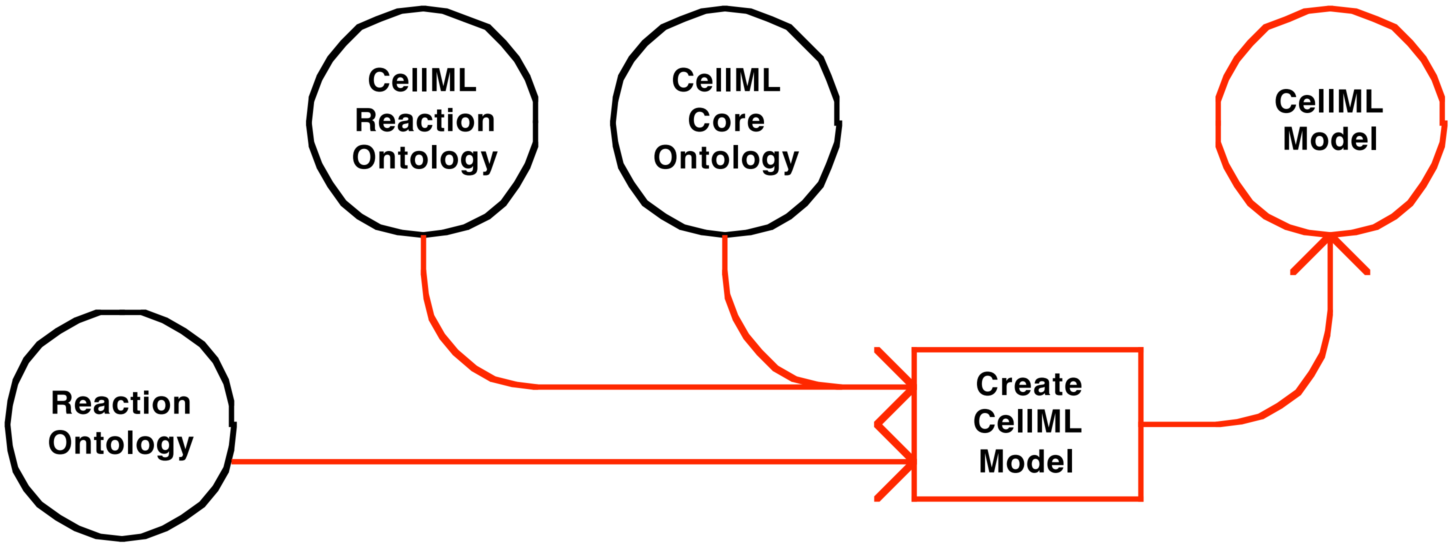
Physiome Project
Quick Reference (Tree)
Quick Reference (Alphabetical)

- Organelles
  - Cytoplasmic organelles
    - Membraneous organelles
      - Cell memebrane
        - Receptors
        - Ion channels
        - Pumps
        - Exchangers
      - Nucleus
      - Mitochondria
      - Rough endoplasmic reticulum
      - Smooth endoplasmic reticulum
      - Golgi apparatus
      - Secretory vesicles (granules)
      - Coated vesicles
      - Lysosomes
      - Endosomes
      - Peroxisomes
    - Nonmembraneous organelles
  - Cytoplasmic inclusions
  - Cytoplasmic matrix (cytosol)



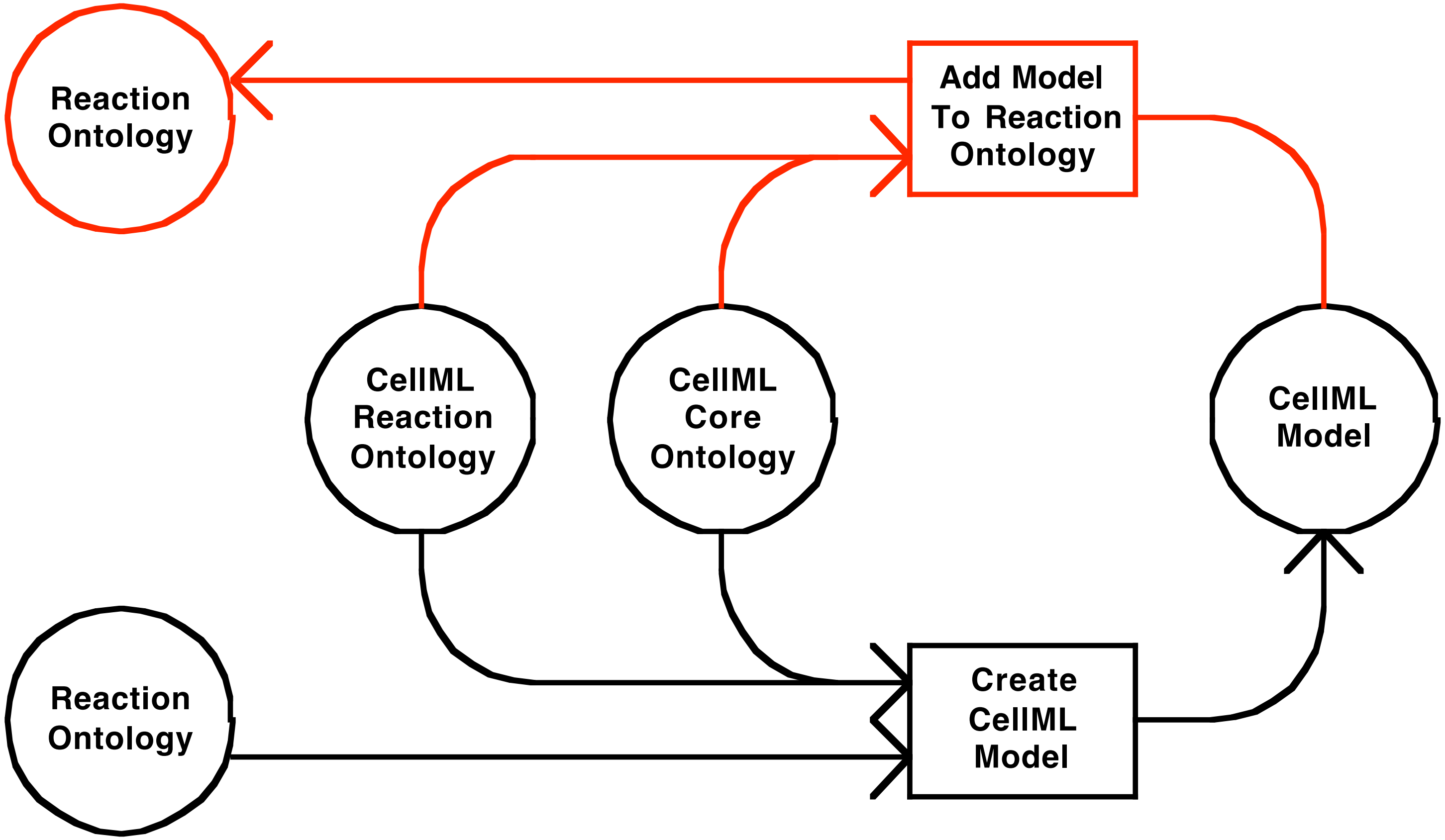
Relationship to CellML…

- A CellML reaction model may be automatically created from the reaction ontology representation using classes and rules defined in CellML ontologies.
- The same transformation, using SBML ontologies, could be used to automatically create an SBML reaction model from the reaction ontology representation.



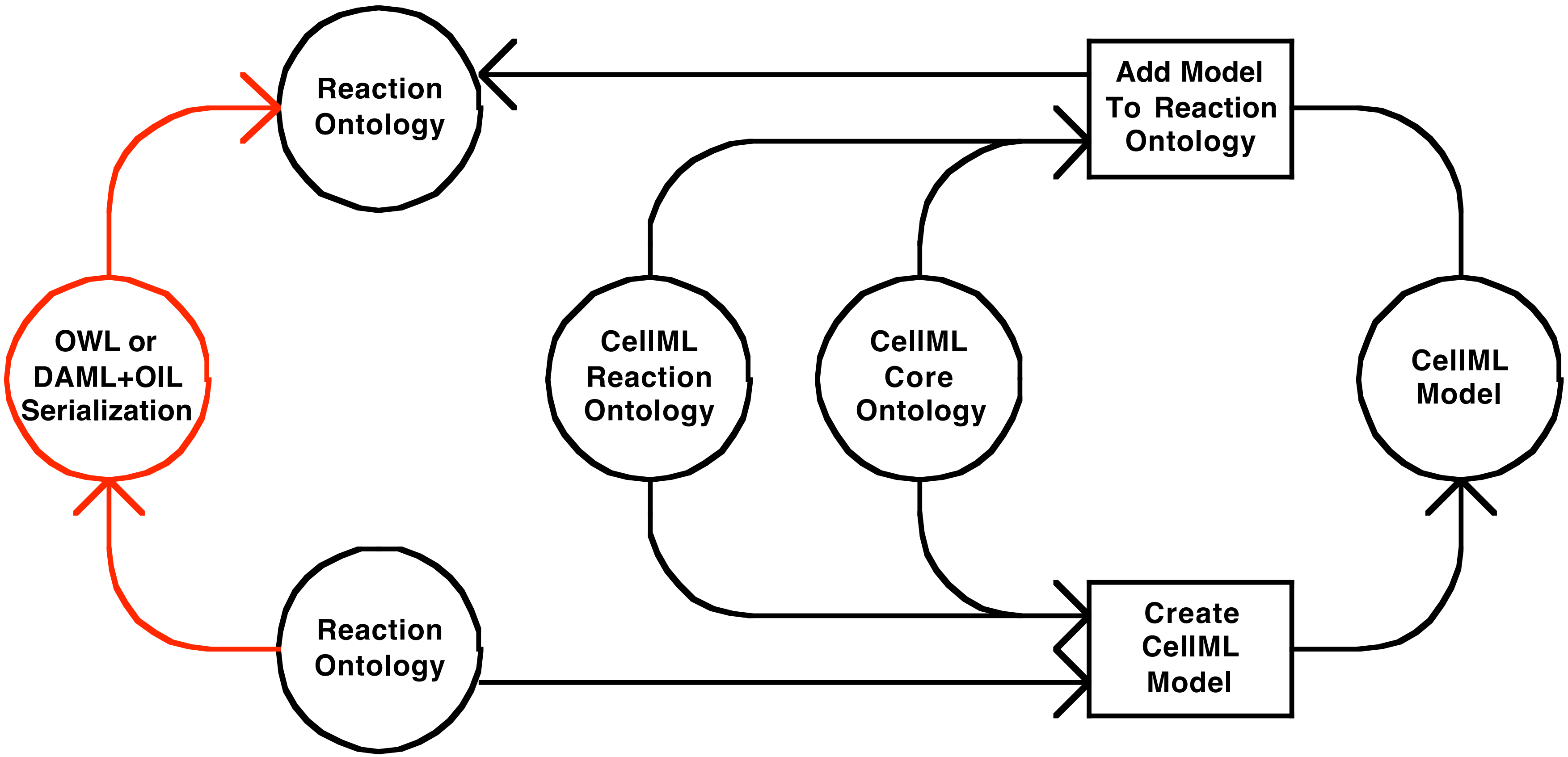
…Relationship to CellML…

- With classing information included in the CellML model, the reaction may be added to another reaction ontology using CellML as the exchange format.



…Relationship to CellML

- A more direct method is to use standard ontology exchange formats, such as OWL or DAML+OIL.



What are the toolsets we are creating?

- Ontology and model databases for concepts and modelling languages.
- Interfaces to the ontologies and models – data standards and APIs.
- Interactive tools for navigating ontology data and models, visualization of model states, and merging of different models.
- Editing tools for creating and modifying Content MathML.
- Editing tools for creating and modifying reaction models.

We use markup standards at all levels

- Our own XML based languages
  - CellML
  - FieldML
  - AnatML
- Ontology standards
  - OWL/DAML+OIL
  - RDF, RDF-Schema
- Application interfaces
  - SVG
  - MathML
  - DOM L2 events
  - XUL layout
  - GraphML, GXL

Physiome project

- The development of the ontologies, markup languages, and tools form part of the Physiome project sponsored by the *International Union of Physiological Sciences* (IUPS) and the *IEEE Engineering in Medicine and Biology Society* (EMBS).
- The Physiome project provides a framework for modelling the human body, using computational methods that incorporate biochemical, biophysical, and anatomical information on cells, tissues, and organs.
- The main project goals are to use computational modelling to analyse integrative biological function and to provide a system for hypothesis testing.

Links

- Physiome project:
  http://www.bioeng.auckland.ac.nz/physiome/physiome.php
- CellML project:
  http://www.esc.auckland.ac.nz/www/sites/physiome/cellml/
  http://www.cellml.org/
- CellML + ontologies:
  http://portal.bioengineering.elyt.ods.org/cellml/CellMLEvolutionMinutes/