Software Tools & Techniques Cell/Biodynamics Simulation Project of Kyoto University

SHIMAYOSHI Takao

Project Introduction

Cell/Biodynamics Simulation Project

www.biosim.med.kyoto-u.ac.jp

- Leader: Prof. Noma
- Main Targets
  - Development of a comprehensive ventricular cell model: Kyoto Model
    - Membrane excitation, excitation contraction coupling, volume regulation, beta signalling, energy metabolism, etc.
  - Simulation of cardiac tissue & heart
    - Excitation propagation, mechanics, circulation dynamics, etc.

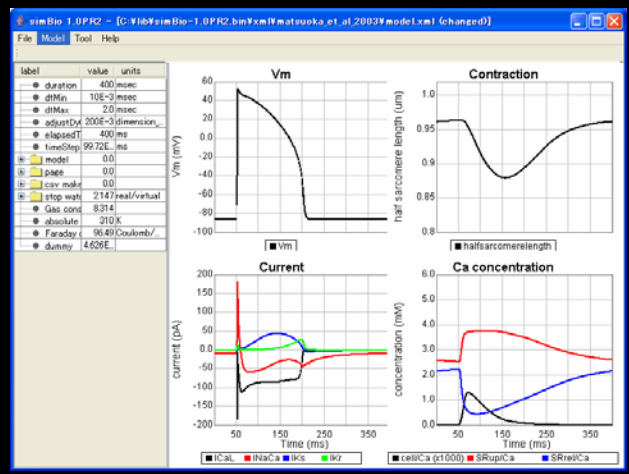
Software Packages

- **simBio** by Dr. Sarai
  - Cell model simulator in use
- **DynaBioS** by Dr. Hori, Dr. Lu
  - Platform for biosimulator in use
- *Cell modelling environment*
  - Editor and simulator under development, to be the next-generation system

simBio

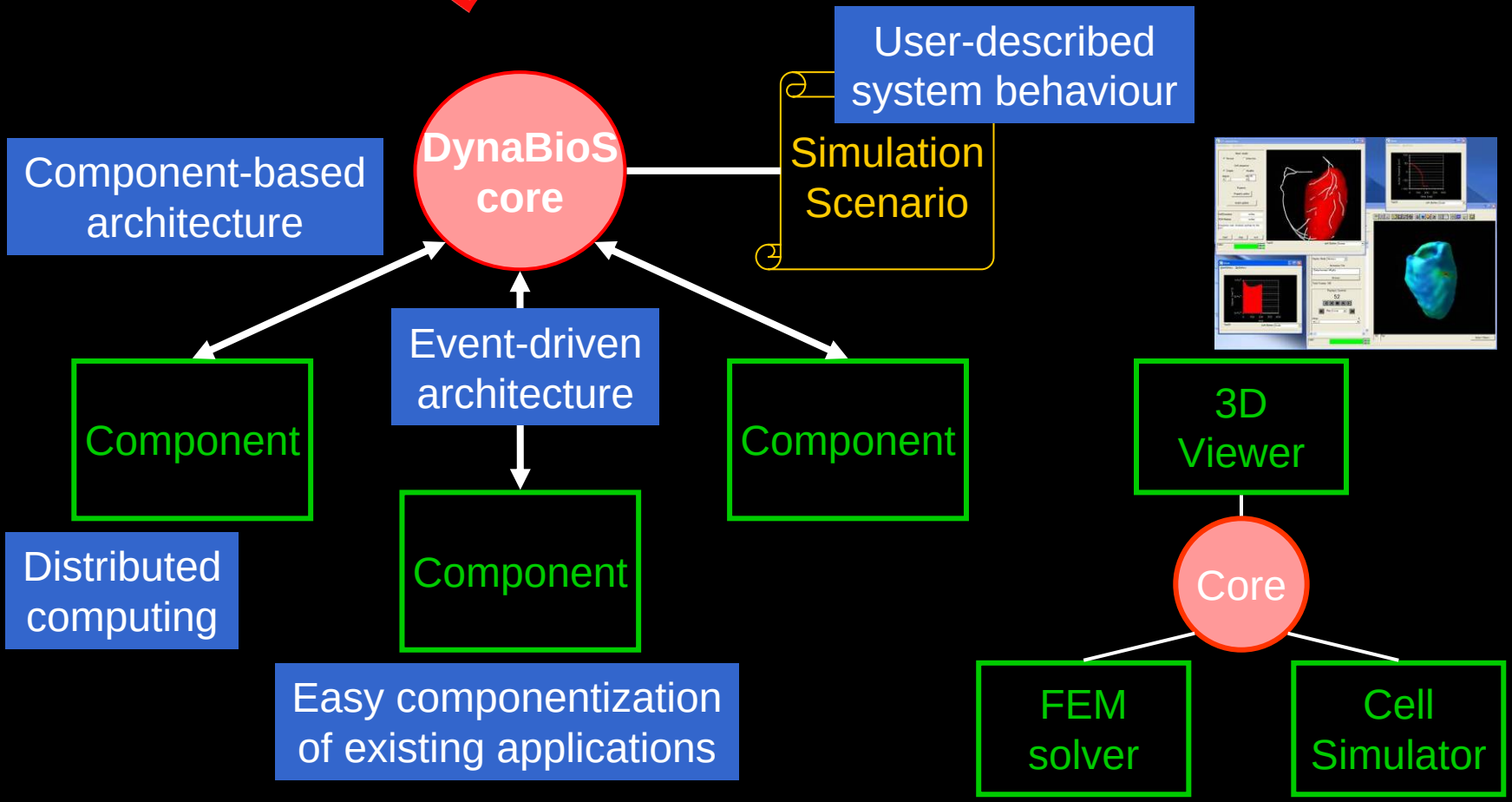
www.sim-bio.org

- Java package for biological simulation
  - Solver of ordinary differential equations
- Object-oriented model composition similar to CellML
  - Each model component is coded as a class Reactor (component in CellML)
- COR can convert CellML files into simBio codes



DynaBioS

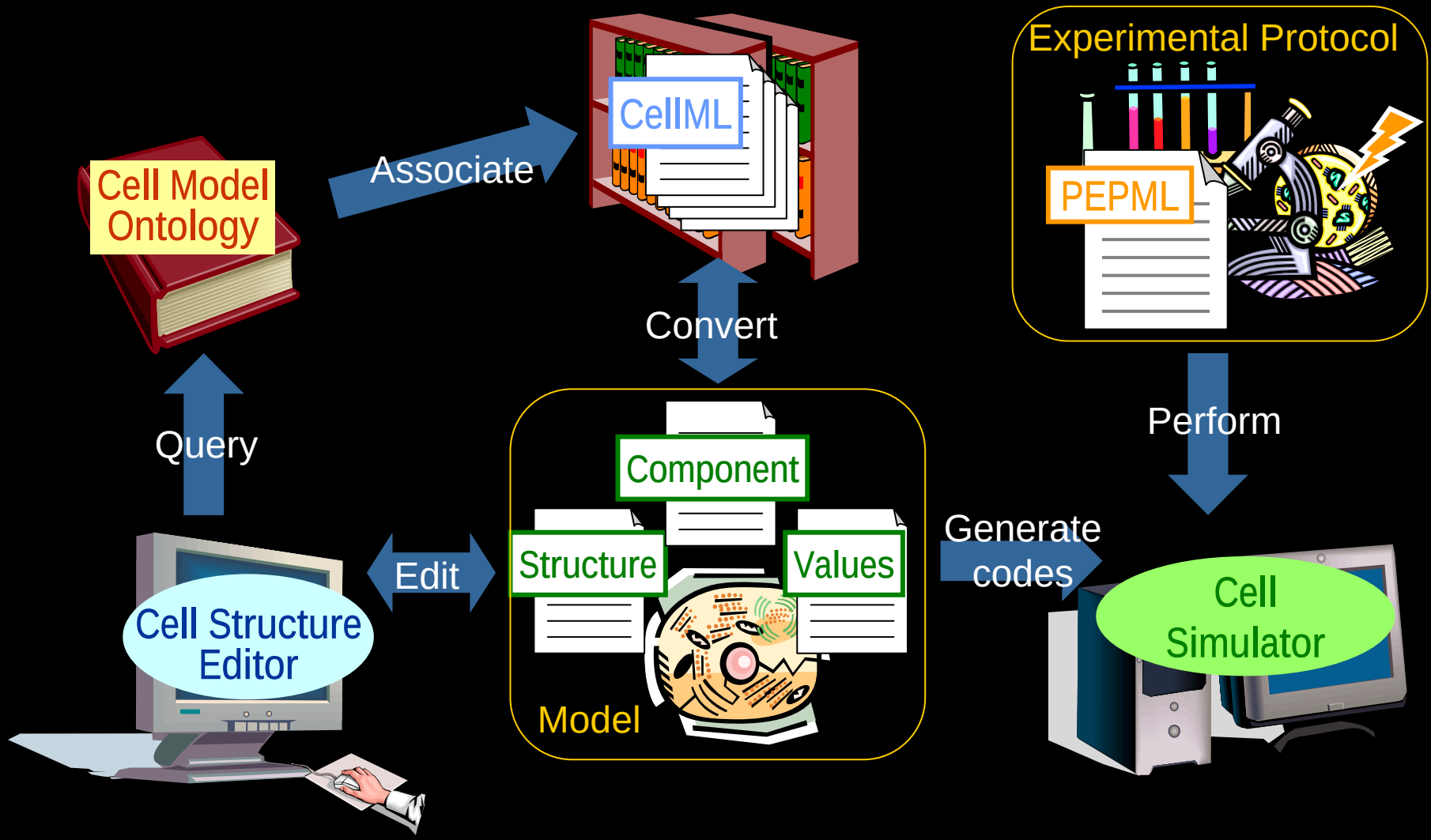
www.dynabios.org



LV motion simulator

Cell Modelling Environment

System Schema

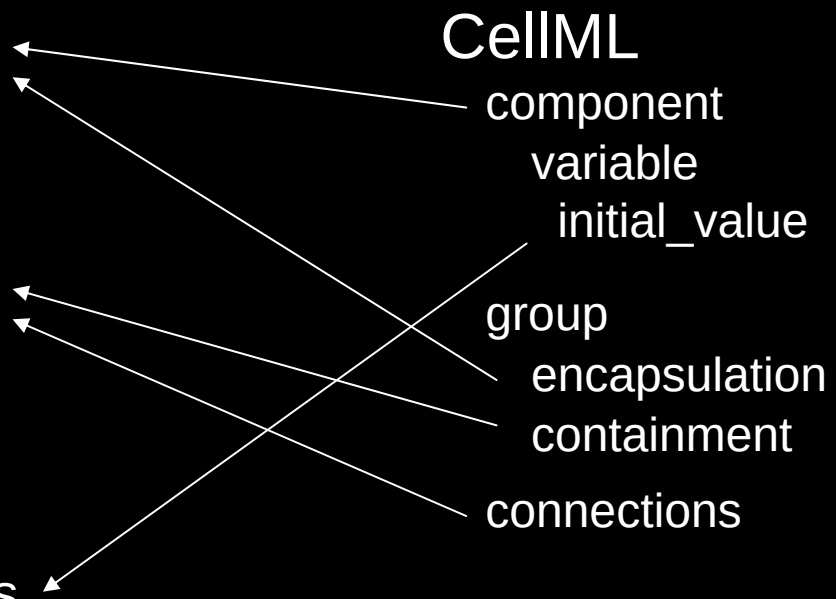


Concept

- Purpose
  - Efficient development and utilization of cell physiology models
- Flexibility
  - Functionally separated tools & formats
- Usability
  - Abstract & semantic representation
  - Intelligent assistance of user operation
- Compatibility
  - Convertible formats from/to CellML files
  - Use of CellML repository as a model library

Model Representation

- Three functionally separated formats
  - Component file
    - Mathematical declaration of a model component
  - Structure file
    - Anatomical hierarchy & composition of a model
  - Values file
    - Values of model variables



- Conversion between a set of three files and a CellML file possible

Developing Methods

- Cell Model Ontology:
  ontology on physiology models
  - For integrated processing of models
- PEPML:
  Physiology Experimental Protocol ML
  - Generic representation with ontology
- Simulation method of model equations
  - Analysis and optimization of calculation procedure with graph theory

Cell Model Ontology

Ontology for Physiology Models

- Existing ontologies: GO, BioPAX, etc.
  - Knowledge about substances
    - Anatomical classification
    - Genome information
    - Proteome information
- Cell Model Ontology (CMO)
  - Knowledge about cellular functions
    - Functional dependence
    - Relationships between substances and functions

Why is CMO needed?

- Cell physiology models refer to the same cellular component or function with different names;
  - (Physiological) function name
    vs. (Biochemical) substance name
  - Historical aliases & abbreviations
- Model components and variables have general functional relationships.

→ For integrated and semantic processing of physiology models, an ontology is needed.

Ontology

- Function: rapid component of delayed rectifier potassium current
  - Generator substance: hERG channel
    - hERG → Abbr. human ether-a-go-go related gene
  - Transports: potassium
  - Symbol: I_{Kr}
- Function: sarcolemmal calcium pump current
  - Generator substance: plasma membrane Ca-ATPase
    - Abbr. ← PMCA
  - Depends on: internal calcium concentration
  - …

Utilities using CMO

- Identification of CellML with the ontology
  - Assign an ID of CMO to each component and variable in CellML files.
- Cell structure editor
  - Edit the composition of a model with intelligent assistance

DOI: 10.2997/ipsjdc.2.726

Identification of CellML

- Append a cmo:id attribute to a CellML element

```
<component name="fast_sodium_current" cmo:id="520">
  <variable name="Nai" cmo:id="211" …/>
```

- Estimation method
  - Lexical keyword matching of the name
  - Analysis of inclusion relationships
- Results
  - Achieves about 80% correct estimations
- Future work
  - Analysis of anatomical locations & mathematical equations

Cell Structure Editor

- Graphical editor of a model structure with importing components from existing CellML files
- Intelligent assist using CMO
  - restricted allocation to the anatomical hierarchy
  - intelligent addition of required variables
  - automatic connection of components and variables
  - extraction of focused components

Current Status of CMO

- Specification:
  - not fixed yet
- Data:
  - Only several entities and limited attributes
- Users and Applications:
  - Finding

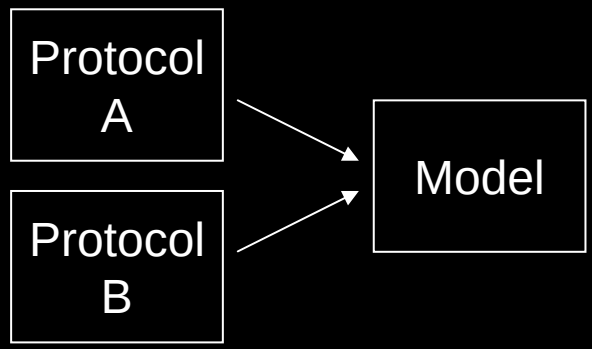
- Collaborators WANTED!

PEPML
Physiology Experimental Protocol Markup Language

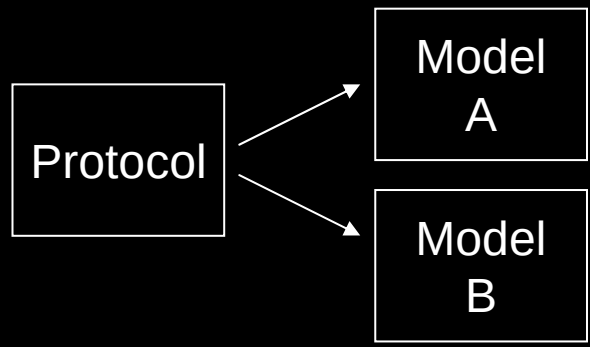
Experimental Protocol

- All physiology experiments are performed according to experimental protocols
- Application of protocol

Model verifications & applications



Model comparisons & tests



PEPML

- A generic representation format of experimental protocols
  - Separate from models
    - Multiple protocols – single model
  - Independent of models by using CMO
    - Single protocol – multiple models
  - Procedural

    ↔ Declarative, CellML

Structure of PEPML

```
<protocol>
 <event id="event1">
  <condition>
   <and>
    <ge>
     <time />
     <literal value="10.0"
              units="ms" />
    </ge>
    <eq>
     <variable ref="cmo:x" />
     <variable ref="cmo:m" />
    </eq>
   </and>
  </condition>
```

$(t < 10.0)\ \&\&\ (x == m)$

```
  <action>
   <set_value>
    <variable ref="cmo:y" />
     <add>
      <literal value="1.0" …/>
      <sin><time /></sin>
     </add>
    </set_value>
    <add_value>
     <variable ref="cmo:z" />
     <literal value="4.0" …/>
    </add_value>
   </action>
  </event>
</protocol>
```

$y = 1.0 + \sin(t)$

$z \mathrel{+}= 4.0$

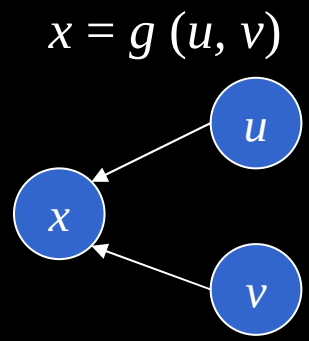
Simulation Method

Model Equations

- Formulation of physiology models
  - Differential Algebraic Equations
    - Differential equation: $dy/dt = f(x)$
    - Algebraic equation: $x = g(p)$
  - Include simultaneous algebraic equations:
    - Chemical & dynamical equilibrium
    - Conservation
- Generic simulation of model equations
  → analysis of calculation procedure

Analysis of Calculation Procedure

- Purpose:
  - Extraction of simultaneous equations
  - Determination of calculation sequence
- Method:
  - Structure analysis of equations with graph theory [Murota 1980]
    - Vertex: variable, Directed-edge: dependence
  - Adapting for physiology models



Optimization of Equations

- General equation forms of models can be optimized

Original		Optimized
$[\mathrm{Ca}]+[\mathrm{CaX}]=[\mathrm{Ca}]_t$ $[\mathrm{X}]+[\mathrm{CaX}]=[\mathrm{X}]_t$ $[\mathrm{Ca}]\cdot[\mathrm{X}]=K_m[\mathrm{CaX}]$	$\Leftrightarrow$	$[\mathrm{Ca}]=[\mathrm{Ca}]_t-[\mathrm{CaX}]$ $[\mathrm{X}]=[\mathrm{X}]_t-[\mathrm{CaX}]$ $0=[\mathrm{Ca}]\cdot[\mathrm{X}]-K_m[\mathrm{CaX}]$
Three dimensions		*One dimensions*

- Develop a method to search equation transformations on the graph of model equations

Requests for CellML

Software Independency

- Keep CellML software independent
  - Functions for particular software to be optional, supplemental and separable
  - Not interleave software dependents into CellML, But import CellML into software specific formats
- Because
  - CellML repository can be a model library for general use (including other utilities than simulator)

Declaration of Published Model

- Describe models as-is in public CellML files
  - Model description to be declarative (not procedural)
  - Without any transformations of equation for numerical calculation

E.G. Original

$$[\mathrm{B}]\cdot[\mathrm{X}] = K_m[\mathrm{BX}]$$

$$[\mathrm{B}]+[\mathrm{BX}] = [\mathrm{B}]_t, \quad [\mathrm{X}]+[\mathrm{BX}] = [\mathrm{X}]_t$$

Transformed

$$[\mathrm{BX}] = \left(-b + \sqrt{b^2 - 4c}\right)/2$$

- Because
  - Transformations cause semantic information lost
    - Transformed equation: no more than quadratic formula
  - Extensions / Imports get impossible

Summary

- simBio
- DynaBioS
- Cell Modelling Environment
  - Model Representation Formats
  - Cell Model Ontology
  - PEPML
  - Method to analyze calculation procedure

Thank you!