Model Composition for Biological Mathematical Systems

Mandeep Gill\textsuperscript{1}, Steve McKeever\textsuperscript{2} and David Gavaghan\textsuperscript{1}
\textsuperscript{1}Department of Computer Science, University of Oxford, Parks Road, Oxford, U.K.
\textsuperscript{2}Department of Informatics and Media Studies, Uppsala University, Uppsala, Sweden

Keywords: Modules, Collaborative Modelling, Biological Mathematical Systems, Heart Models, Generics.

Abstract: Mathematical models are frequently used to model biological process, such as cardiac electrophysiological systems. In order to separate the models from the implementations, and to facilitate curation, domain specific languages (DSLs) have become a popular and effective means of specifying models (Lloyd et al., 2004; Hucka et al., 2004). In previous papers (Gill et al., 2012a; Gill et al., 2012b; McKeever et al., 2013) we have argued for including parameterised modules as part of such DSLs. We presented our prototype Ode language and showed how models could be created in a generic fashion. In this paper we extend our work with concrete examples and simulation results. We show how complex heart models can be constructed by aggregation, encapsulation and subtyping. Our use-case retraces the steps taken by (Niederer et al., 2009), which investigated the common history between cardiac models, and shows how they can be cast in our language to be reused and extended. Our DSL enables ‘physiological model engineering’ through the development of generic modules exploiting high cohesion and low coupling.

1 INTRODUCTION

We are investigating the application of system engineering methods to mathematical continuous models of biological systems, where the complexity of such models prohibits the use of discrete modelling and simulation methods. The mathematics within a model lends itself to a textual representation, traditionally this has been accomplished in an operational manner using numerical languages such as MATLAB and Python. Research has also resulted in the development of the Systems Biology Markup Language (SBML) (Hucka et al., 2004) and CellML (Lloyd et al., 2004) domain specific formats for common interchange and archiving. This modelling approach has a successful history (Noble and Rudy, 2001) that can be seen through the large number of individual models curated within the CellML model repository (Lloyd et al., 2004).

Such modelling is collaborative. With the advent of domain specific languages for describing models we are looking at incorporating features from programming language theory to improve collaborative design of physiological models. Modular programming involves describing a system in terms of encapsulated modules and the interactions between them through well-defined interfaces. Such modular program design is commonly used in software engineering when designing large-scale reusable systems, with the key tenets including encapsulation, abstraction, reuse and extensibility (Booch, 2006).

We have been developing our own prototype modelling language, Ode, to investigate modularity and structured software engineering concepts (Somerville, 2010) that may facilitate collaborative design within this domain. Ode is influenced by the mentioned biological DSLs and functional programming languages rooted in the lambda calculus (Barendregt, 1985). Figure 1 presents the Ode grammar.

```
oeStmt ← componentDef | valueDef | odeDef
componentDef ← component f (id, \ldots, id)
| {exprStmt, \ldots, exprStmt, return (E_1, \ldots, E_n)}
valueDef ← valid id, \ldots, id = E
E ← t_1 \& t_2 | t_1 > t_2 | \ldots | t_1 \& \& t_2 | t
\| \{expr\} number | boolean | piecewise | t
piecewise ← piecewise \{E_1 : E_1, \ldots, E_n : E_n, default : E_{\text{def}}\}
odeDef ← ode \{init : id, \{deltaVal : id_{\text{delta}}\} = E\}
```

Figure 1: Abbreviated syntax of the core Ode language using a simplified variant of EBNF that assumes common rules.

The semantics of this core language have been...
described in a previous paper (Gill et al., 2012a). Constant values may be defined as the result of mathematical expressions; the syntax for such expressions is similar to most programming languages. Components are a means to group, abstract and parameterise repeated computations. However they are inline at compile-time in a manner similar to C++ templates. Piecewise terms enable conditional control-flow similar to a switch statement. A deliberate limitation is the lack of looping control-flow constructs that, when coupled with a recursive restriction on components, ensures termination during model simulation.

The basic expression language is extended with several constructs to support the mathematical modelling of (biological) systems over time, including ordinary differential equations (ODEs). ODEs require an initial value \( y(0) \), and a derivative expression \( y' \).

The focus of this research has been to apply software engineering techniques that allow models developed in Ode to be highly reusable and extensible, facilitating rapid composition to investigate certain parameters of interest. To the best of our knowledge, these characteristics have eluded the majority of biological modelling software frameworks to date.

Our paper in outlined as follows. Section 2 motivates modular model design. In Section 3 we provide a precise description of our module system and motivating example. In Section 4 we present results from simulating composite models derived from experimentally-related model modules. Finally in Section 5 we conclude our work and discuss current research. We demonstrate the utility of the module system using examples from a modular form of the Hodgkin-Huxley (HH52) electrophysiological model of the squid giant-axon (Hodgkin and Huxley, 1952).

## 2 MODULAR MODEL DESIGN AND DEVELOPMENT

Modelling biological systems is an integrative, interdisciplinary process, drawing on work from researchers in a global environment from a variety of fields, including physiologists, mathematicians, and computer scientists. Models themselves are comprised from data collected through various studies and are developed iteratively, with newer, generally more realistic, models derived from existing ones (Niederer et al., 2009).

As such a need exists to efficiently facilitate the collaborative development and reuse of models. This would allow for the safe construction of models, guaranteeing the validity of models developed collaboratively and derived from multiple sources, and ensuring that they may be composed successfully. We believe our research provides features that facilitate such collaborative modelling in a structured and safe manner.

We present a summation of the methodologies used to structure and develop large-scale reusable biological mathematical models using Ode, inspired by many techniques utilised in software engineering (Sommerville, 2010; Booch, 2006).

*Encapsulation* is the wrapping up of operations and attributes into a module, so that those attributes may only be manipulated through or accessed via the operations provided by the module. We can view modules as similar to objects, representing a convenient data-centric manner to decompose systems into understandable and manageable units/building blocks. Good encapsulation hides the details of a module’s internal attributes and operations from its users. These techniques are known as information hiding and implementation hiding, and their use is essential for promoting the understandability of code within a reusable domain-specific modelling framework. It allows abstracting away the low-level details and providing a high-level interface to the module functionality. Explicit export definitions within module declarations facilitate implementation hiding and help to achieve this goal. Figure 2 provides an example encapsulating several HH52 cell-level parameters into a module.

*Generic* modules allow modellers to operate on encapsulated objects in an abstracted manner, enabling polymorphic reuse and substitutability, similar to C++ templates. It is a style of programming in which algorithms are written in terms of to-be-specified-later constructs that are then instantiated when needed by replacing the generic variables with appropriate concrete parameters. In the context of heart modelling, we show that this technique enables module reuse, the creation of alternative implementations, and the mixed usage of ion channel models from a variety of sources.

*Structural subtyping* is implemented at the module level, where we say that module \( B \) is a subtype of module \( A \), the supertype. This occurs when module \( B \) has the same (or a superset of the) interface as \( A \), yet may contain a different implementation. In this case modules of type \( A \) can be substituted by modules of type \( B \). This occurs because methods written to operate on elements of the supertype can also operate on elements of the subtype; in a manner similar to

---

**Figure 2**: Cell and stimulus parameters for the HH52 model grouped and encapsulated into a single parameters module.
Figure 4: Diagram indicates use of containment and aggregation to provide code structure and reuse. We compose the parameters module with the main model, whilst specifying that it is a subtype of HH52Parameters. The HH52DefaultParameters module implements this interface via the aggregation and reuse of two sub-modules containing the required code.

inheritance in object-orientated programming (OOP).

However it is implemented at compile-time without any performance penalty and does not require a direct, nominative relation between the supertype and subtype. Any modules that have a similar ‘shape’ can be considered a subtype, providing low-coupling between modules and allowing rapid development and substitution of alternate implementations. If used properly, subtyping, coupled with generics, can improve the understandability of model code by reducing the conceptual distance between code and the real-world system which the code models. For instance, Figure 3 presents a use of both generics and subtyping to produce a parameterised module of a given supertype.

In software development, code reuse provides many benefits — the code has already been developed, tested and potentially deployed. A similar effect is observed with model code, where equations and definitions are often similar and should be reused if possible. However in both domains, code often requires modifications specific to each particular instance of reuse. Containment-based relationships, such as composition and aggregation, can be implemented by modules to enable lightweight and flexible code reuse.

Modules may be imported and wrapped within containing modules to re-implement and augment existing functionality, delegating to existing code as needed. This provides code reuse through composition and aggregation rather than OO inheritance.

This technique works hand-in-hand with the substitutability property of modules and structural subtyping mechanism to allow modellers to create model implementations which are both malleable and extensible. It enables type-safe substitution whilst avoiding deep inheritance hierarchies often seen in systems implemented using traditional OO-languages (Booch, 2006).

When parameterised modules are utilised with aggregation they allow generic code reuse, abstracting the particular instances until simulation time. This provides compile-time substitutability and a powerful mechanism for rapid model construction and modification. Figure 4 demonstrates aggregation to structure and reuse code with the HH52 parameters module.

3 MODULE SYSTEM – SYNTAX AND SEMANTICS

In this section we describe the Ode module system used to structure models into modules consisting of collections of related value and component definitions that form some logical (and potentially biological) grouping. We describe the syntax and semantics; the grammar for the module language is provided in Figure 5.

```
moduleStmt ← moduleDef | importStmt | applyModule
moduleDef ← module id moduleBody
moduleBody ← {odeStmt1, ..., odeStmtn}
importStmt ← import id1 [as id2]
exportStmt ← export [id1, ..., idn]
applyModule ← id = f(id1, ..., idn)
```

Figure 5: Module extensions to Ode DSL.
We discuss parameterised modules, enabling compile-time substitution of components and generic, interface-driven development that provides further opportunities for abstraction and model reuse.

The module system was inspired by OCaml, where it is effectively a higher-level language that enables the programmable creation of independent modules (Leroy, 2000). The *Ode* module system operates similarly; providing variable definitions, references, function abstraction and application (i.e. parameterised modules), and limited computation, all at the typed module level. This provides extensive power to reconfigure models for particular tasks.

### 3.1 Module Definition

```ml
module Parameters {
  val E_R = -75 { unit : mV }
  val Cm = 1 { unit : mF.cm^-2 }
  val period = 60 { unit : ms }
  val i_Stim = piecewise {
      time >= 10 and time <= 10.5:
        20 { unit : mA.cm^-2 }
      default: 0 { unit : mA.cm^-2 }
  }
}
```

A standard module in *Ode* consists of a collection of expressions, i.e. value and component definitions, that form a grouping. This may be a logical or biological grouping, e.g. an ion channel within a cardiac model. The code fragment above presents the syntax for creating a module encapsulating several HH52 cell-level parameters, as was illustrated in Figure 2.

Several modules may be defined within a single file, where the directory path in which the files are placed creates a module hierarchy. Modules may be referenced using a dot-notation that corresponds to lookup within the module hierarchy. This structure enables the creation of multiple repositories of modelling components, each with their own hierarchy, that can be shared, reused and uniquely identified.

### 3.2 Importing & Using Modules

```ml
module HH52Model {
  import Cardiac.HH52.Parameters as Pa
  // leakage current
  val E_L = Pa.E_R + 10.613 { unit : mV }
  val i_L = 0.3 { unit : mS.cm^-2 } * (V - E_L)
  // ... channels i_Na & i_K ...
  init V = -75 { unit : mV }
  ode { init : V } =
    -(Pa.i_Stim + i_Na + i_K + i_L) / Pa.Cm
}
```

A module may reference another module’s expressions through an *import* declaration. This is demonstrated in the above code sample, illustrating the composition of the HH52 parameters module with the HH52 base model, as in Figure 2. The declaration uses a dot-notation to reference a filepath within the module hierarchy and a module within the specified file. Thus, in the above example the HH52 base model imports a module *Parameters* from a file *HH52* within the *Cardiac* directory of an available module repository. When processing the import, the module is retrieved and loaded into the global module environment.

The visible definitions of an imported module are again accessed using a dot-notation on the qualified module name, demonstrated by references to `Pa` above.

### 3.3 Module Interface

Modules expose an interface, or signature, comprised from the collection of identifiers and their types visible from outside the module. The uses of a module within a model also leads to the implicit creation of a required interface that any imported modules must implement. When importing a module, interfaces are matched and checked by the module type-system to ensure successful composition.

A modeller may configure the visibility of certain definitions to modify the interface. This may be achieved by structuring the model code such that only exportable definitions are visible, or by explicitly declaring the definitions to export (similar to the *public/private* access modifiers in OO languages) through the use of an *export* command.

In this manner only the vital information is exported in the interface, mimicking biological mechanisms of compartmentalisation and containment.

### 3.4 Parametrised Modules

Grouping related definitions into modules enables modularity and code reuse, however the modules and the abstractions within them are fixed rather than generic. When creating reusable components it is desirable to configure them according to a specific use-case, for instance altering the parameters of a reusable ion channel or creating protocols to compare simulations against experimental data.

In several languages parameterised modules present a flexible and powerful means for accomplishing this. They are implemented as compile-time construct used to create generic components in a similar manner to ML functors (Leroy, 2000) and C++ templates, allowing type-safe substitutability of components with differing implementation details.

Modules may be parameterised by other modules leading to the creation of complex, specialised mod-
ules via a form of aggregation. This flexibility and abstraction enables rapid investigations into parameter variations, including parameter sweeps and sensitivity analysis from a single code-base (O’Hara et al., 2011).

### 3.5 Parameterised Module Definition

```scala
module HH52ModelBase(Pa) {
  // leakage channel
  val E_L = Pa.E_R + 10.613 { unit: mV }
  val i_L = 0.3 { unit: mS.cm^-2 } * (V - E_L)
  // ... channels i_Na & i_K ...
  init V = -75 { unit: mV }
  ode { init: V } =
    -(Pa.i_Stim + i_Na + i_K + i_L) / Pa.Cm
}
```

The above example demonstrates parameterised modules within Ode, where the specific parameters module utilised by the HH52 base model is parameterised and made generic. The syntax extends module definitions to include a list of module parameters, these are generic module arguments that are determined during application. Parameterised modules can be considered analogous to functions over modules, they take a set of input module arguments and return a new module.

Parameterised modules enable a form of interface-based model construction that allows for the specialisation of reusable module components. They facilitate several component-based modelling techniques through the modification of parameters and equations used within a module, for instance the creation of multiple model subcomponents ranging in complexity and accuracy to minimise computational demands.

### 3.6 Parameterised Module Application

```scala
module HH52Model = HH52ModelBase(Parameters)
```

The above code fragment demonstrates applying a concrete version of the HH52 parameters module to the base module, resulting in a complete version of the HH52 model, as was illustrated in Figure 3.

Applying a functor is the process of replacing the generic module parameters with real modules, the effect of which is to instantiate a new module derived from applying the parameters to the functor body. This new, concrete module can then be used elsewhere.

We can dynamically apply functors, enabling the run-time configuration of models that exhibit certain behaviours and investigation of effects using particular parameters. This provides a flexible means for rapidly constructing families of specialised models that can be controlled by an external process at run-time.

## 4 MODULAR CARDIAC SIMULATIONS

In this section we utilise the module system and DSL abstractions previously described to design a component-based modular modelling framework. We hope this will eventually enable the agile development, reuse and investigation/experimentation of cardiac electrophysiological models in a collaborative fashion (Noble and Rudy, 2001).

Our use case is influenced by research in (Niederer et al., 2009) that investigated the common history between cardiac and ion channel models, but we conduct similar simulations through programmable modular composition rather than manual model construction. Our sample framework is used to conduct several simulations, generating action potential (AP) membrane voltage curves for cardiac models created from the dynamic substitution of ion channel modules. This is intended to show the ease at which a well-designed framework lends itself to reuse, extensibility and flexibility for modellers to investigate particular phenomena. Furthermore it demonstrates the use of the type system to ensure subsequent model validity.

### 4.1 Methodology

We investigate and construct a modelling architecture suitable for creating models in a flexible and extensible manner comprised from individual modules. This framework is used to develop a repository of cardiac models to help qualitatively determine the DSL’s effectiveness in the large-scale structure and rapid, collaborative development, modification and reuse of models. Several cardiac models are safely composed in a scripted manner from modules in the repository.

#### 4.1.1 Cardiac Models

The cardiac models are listed in Table 1, and are chosen to demonstrate the iterative development process and reuse of parameters and equations.

#### 4.1.2 Model Structure & Development

We separate the cell models into reusable and substitutable ion channel objects that each model the flow of a particular charged ion across the cell membrane, resulting in the generation of an AP. We abstract and modularise at this point as most cardiac cell experiments, and subsequent research data, occurs at the ion channel level. As more information is derived regarding their function we hope that their differing representations within cell models may be used interchangeably and gradually unify into single canonical
Table 1: Cardiac ventricular electrophysiological models used within study, where the base model is denoted with a *.

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
<th>Species</th>
<th>Ion Channels</th>
<th>State Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR77 * (Beeler and Reuter, 1977)</td>
<td>1977</td>
<td>Mammalian</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>LR91 (Luo and Rudy, 1991)</td>
<td>1991</td>
<td>Guinea Pig</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>LRd94 (Luo and Rudy, 1994)</td>
<td>1994</td>
<td>Guinea Pig</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

Figure 6: Interface for reusable ion channel models, consists of component, `getCurrent`, that computes $I_{\text{ion}}$ for calculating the membrane voltage. All ion channel models implement this interface, as `NaChannel` and `KChannel` do.

A standardised interface for ion channels models was determined that enables their use and replacement within cell models. This interface is illustrated in Figure 6, and consists of a single component, `getCurrent`. The input parameters represent the membrane potential and equilibrium potential respectively, and the output is the current generated by ionic flow.

A cardiac model structurally depends upon and contains the ion channel modules, requiring that they expose the `IonChannel` signature to be compatible. This signature is checked at compile-time by the type system to ensure that only valid ion channel objects are used. The structure used to associate a cardiac model and its related ion channels is illustrated in Figure 7.

We have used encapsulation to separate definitions for ion channels into modules that expose a specified interface. We now demonstrate how aggregation, encapsulation and subtyping may be used to abstract out common functionality and enable code reuse. From a cardiac modelling perspective, we can define common functionality for ion channels. These are then encapsulated and extended in later, more complex channel model subtypes through aggregation (mirroring the real life development of such models), as was demonstrated in Figure 2. We can thus delegate to existing code, and override and specialise as needed within the newer encapsulated module, whilst ensuring it exposes the same interface to enable substitutability.

We use generics to specify the input ion channels to each model via parameterised modules, allowing simulation-time configuration and instantiation of concrete modules. Generics provide a standardised manner to alter models and parameters, providing substitutability without incurring a performance penalty. For example, in Figure 8 the model, `BR77Model`, contains the generic module reference `NaChannel` that exposes the `IonChannel` interface. We do not need to specify during model development which specific `NaChannel` implementation we are referring to.

We utilise encapsulation, aggregation, subtyping and generics to abstract behaviour and provide substitutability of reusable model components. This benefits modellers by enabling model reuse, component-driven development, and type-checked model composition. These techniques orthogonally help model development by increasing cohesion and decreasing coupling.

4.2 Results

```cpp
to each model via parameterised modules, allowing simulation-time configuration and instantiation of concrete modules. Generics provide a standardised manner to alter models and parameters, providing substitutability without incurring a performance penalty. For example, in Figure 8 the model, `BR77Model`, contains the generic module reference `NaChannel` that exposes the `IonChannel` interface. We do not need to specify during model development which specific `NaChannel` implementation we are referring to.

We utilise encapsulation, aggregation, subtyping and generics to abstract behaviour and provide substitutability of reusable model components. This benefits modellers by enabling model reuse, component-driven development, and type-checked model composition. These techniques orthogonally help model development by increasing cohesion and decreasing coupling.

4.2 Results

```cpp
module BR77NaChannel {
    export (getCurrent)
    component getCurrent(V, E_R) {
        // ... channel constants & computations ...
        // calculate current
        val i_Na = (g_Na * pow(m, 3.0) * h * j + g_Nac) * (V - E_Na)
        return (i_Na)
    }
}
```

The above listing contains a segment from the BR77 $I_{Na}$ model (Beeler and Reuter, 1977), this segment includes the top-level component that comprises the module interface and implements the type-signature required for reusable channels within our framework.

```cpp
module BR77Model(NaChannel) {
    // membrane voltage
    init V = -84.624 { unit : mV }
    // ... model constants and computations ...
    val i_Na = NaChannel.getCurrent(V, E_R)
    // Membrane Voltage ODE
    ode { initVal : V, deltaVal : dV } =
        (i_Stim - (i_Na + i_s + i_x1 + i_K1))/C
    }
```

Similarly, the listing above contains a segment from a cardiac model within the framework. It shows a BR77 cell model parametrised by a module representing $I_{Na}$, this may be provided by any ion channel model that implements the correct module signature.

We ran a series of simulations that demonstrate the substitution and impact of the `NaChannel` within the
models in a fully scripted manner. This was inspired by the work in (Niederer et al., 2009) and the modular framework was structured to safely enable such rapid-modifications and experimentation. We initially composed each cardiac base model with its corresponding sodium channel and simulated the resulting complete model to generate the reference AP curves depicted in Figure 9. The sodium channels from the alternate models were then applied to the BR77 cell model to generate a new family of models whose AP curves are presented in Figure 10. This plot shows that the LR91 and LRd94 $I_{Na}$ models slightly shortens the AP duration of the BR77 cell model. $I_{Zero}$, implementing the same module signature but returning a zero current, fails to trigger an AP. This is expected as sodium ions are responsible for AP initialisation.

We performed the inverse operation, applying the $I_{Na}$ BR77 model into the remaining cellular models to investigate its influence. The AP curves from simulating these models are presented in Figure 11, where replacing the $I_{Na}$ from the LR91 and LRd94 models increases the AP duration and alters the peak AP. We do not consider these results particularly important in terms of their biological meaning, instead they serve to demonstrate the application and benefits of modular programming and architecture to developing reusable, extensible and dynamic models. The results were obtained from a high-performance DSL implementation that we intend to discuss in future publications.

Figure 7: Modular structure of a cardiac model in the repository. The base model represents the cell membrane and calculates the membrane voltage through references to many ionic channels that implement the IonChannel interface.

Figure 8: Using generics within the model framework — either BR77NaChannel or LR91NaChannel can be applied as module argument to the base BR77Model cardiac model to represent the generic NaChannel.

Figure 9: AP curves from applying original $I_{Na}$ channels into models using modules.

Figure 10: AP curves from applying alternate $I_{Na}$ channels into BR77 cardiac cell model using modules.

Figure 11: AP curves from applying BR77 $I_{Na}$ channel into alternate models using modules.
5 CONCLUSIONS

Our use-case driven design process has yielded a modular framework which allows a family of cardiac electrophysiological models to be extended both intuitively and easily. This extensibility was achieved by utilising features such as generics and subtyping which exploit the substitutability property of module hierarchies along with encapsulation for code structure and aggregation for reuse. This is a common approach used to structure large-scale software projects. New cardiac models may utilise existing ion channel models, enabling the DSL to naturally capture the reuse of models and experimental data. Thereby increasing robustness which is important if these models are ever used in a predictive pharmaceutical or clinical setting.

We can create hybrid models that include ion channel representations from several models. This can be undertaken automatically simply by utilising ion channel modules from different models when instantiating a generic cell model. Generic modules may be used to perform sensitivity analysis of the model to parameter fluctuations. They can be used to alter equations, for example to model ion channel changes and resulting effect on the AP caused by mutation or drug block (O’Hara et al., 2011) in an abstracted manner.

Models created in the fashion that we have illustrated do not depend on each other explicitly. They do not communicate with one another either. Parameterisation only requires the type signature of the parameter object to be known. Consequently they demonstrate low coupling and high cohesion, an aspect of model design that we feel is important for reusability and extensibility (McKeever et al., 2013).

We are currently developing a repository to reproduce existing models. Our notation is textual but as we have shown the modularity constructs trialled in Ode have equivalent UML-style graphical representations. Therefore visual physiological modelling environments could also be developed to support collaborative efforts to construct future models.

REFERENCES


