

Efficient Analysis of SBML Models Using Arrays

Leandro Watanabe and Chris Myers

University of Utah

COMBINE

October 14, 2015

Motivation

- Standards are key to the success of systems and synthetic biology since standards give biologists the ability to share models.
- The leading standard representation of biological systems is the *systems biology markup language* (SBML).
- Karr's whole-cell model is difficult to encode in SBML due to large regular structures.
- Similarly, in synthetic biology, population of cells are also large structures that are difficult to encode in SBML.
- The SBML *arrays* package has been proposed to overcome such limitations.

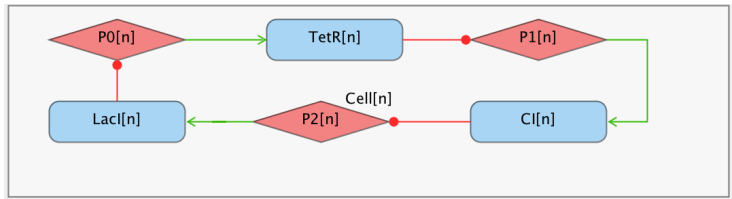
Arrays in iBioSim

- iBioSim provides a user interface to construct SBML models.
- This includes all SBML core constructs: **compartments**, **species**, **reactions**, **parameters**, etc.
- SBML core constructs can be arrayed in iBioSim.
- *Flux balance analysis* and *hierarchical model composition* package constructs can also be arrayed.
- A model can be flattened for simulation.
- There are some overheads in regards to memory.
- This talk focuses on a simulation method that does not flatten models.

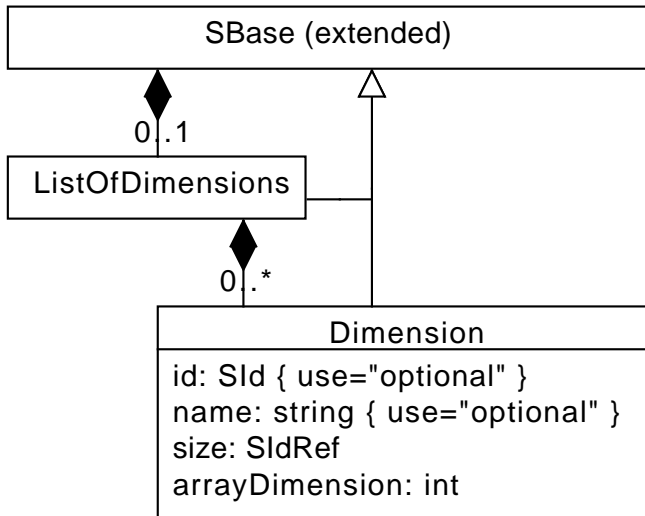
Arrays Package

- The *arrays* package puts dimensions on objects.
- For objects that holds a value (e.g. **species**, **parameters**, **compartments**, etc), a state is created for each element in the array.
- For others, new structures are created.
- Objects that reference arrayed objects need indices to indicate which element in the array is being referenced.

Repressilator Circuit



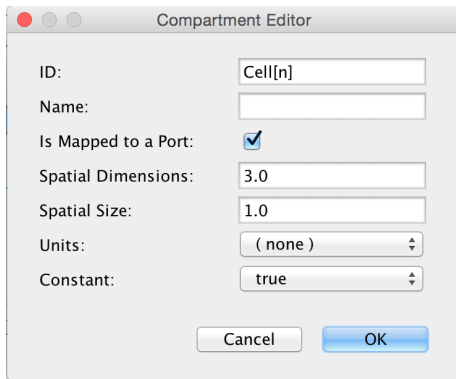
Dimension



SBML Elements That Can Have List Of Dimensions

- In SBML Core, **ListOf** objects are not allowed to have a **ListOfDimensions**.
- All SBML objects defined that inherit from **SBase** are permitted to have a **ListOfDimensions** unless it is explicitly disallowed in the corresponding specification.
- In SBML Core, the elements that can have a **ListOfDimensions** are:
 - **Compartments**
 - **Species**
 - **Parameters**
 - **Initial assignments**
 - **Rules**
 - **Constraints**
 - **Reactions**
 - **Species references**
 - **Events**
 - **Event assignments**

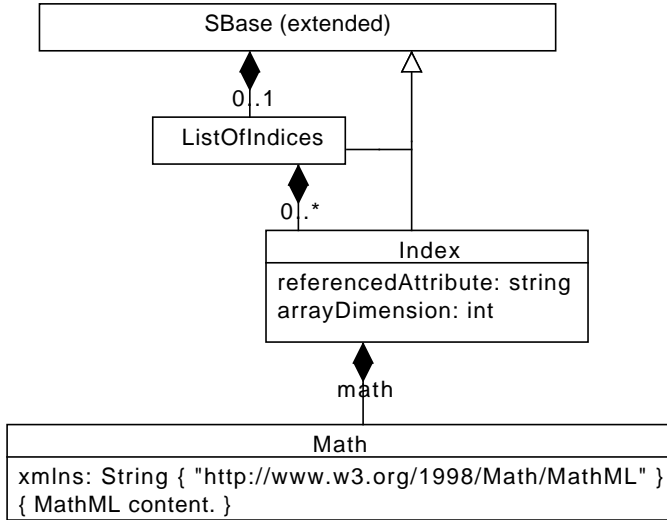
Creating Array of Compartments



The image shows a 'Compartment Editor' dialog box with the following fields and controls:

- ID:** A text field containing 'Cell[n]'.
- Name:** An empty text field.
- Is Mapped to a Port:** A checked checkbox.
- Spatial Dimensions:** A text field containing '3.0'.
- Spatial Size:** A text field containing '1.0'.
- Units:** A dropdown menu showing '(none)'.
- Constant:** A dropdown menu showing 'true'.
- Buttons:** 'Cancel' and 'OK' buttons at the bottom right.

Index



SBML Core Elements That Can Have List of Indices

- Only SBML objects that include defined attributes of type SIdRef are permitted to have a **ListOfIndices**.
- The following SBML Core objects can have a **ListOfIndices**:
 - **Model** - conversionFactor
 - **Species** - compartment, conversionFactor
 - **Initial assignments** - symbol
 - **Rules** - variable
 - **Species references** - species
 - **Events assignments** - variable

Indexing Compartment Attribute

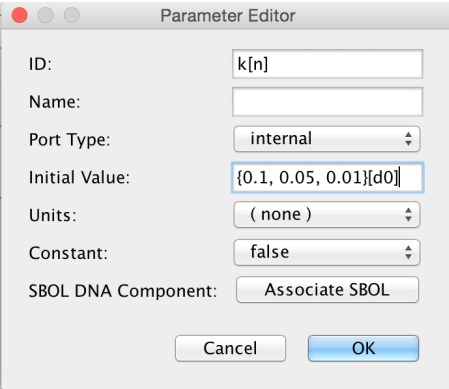
Species Editor

ID	LacI[n]	
Name		
Port Type	internal	
Compartment	Cell	
Compartment Indices	[d0]	
Initial Amount/Concentration	0.0	
Units	(none)	
Conversion Factor	(none)	
Conversion Factor Indices		
<input type="checkbox"/> Boundary Condition	<input type="checkbox"/> Constant	<input checked="" type="checkbox"/> Has Only Substance Units
<input type="checkbox"/> Constitutive	<input checked="" type="checkbox"/> Degrades	<input type="checkbox"/> Diffusible
Open complex production rate (ko)	default	0.05
Stoichiometry of production (np)	default	10.0
Degradation rate (kd)	default	0.0075
Complex formation equilibrium (Kc)	default	0.05/1
Membrane diffusion rate (fd/rv) (kmdiff)	default	1.0/0.01
SBOL DNA Component	Associate SBOL	

Cancel Ok

- *Constructors:*
 - **vector**
- *Element referenced operator:*
 - **selector**

Initial Assignment Example



A screenshot of a 'Parameter Editor' dialog box. The dialog has a title bar with three window control buttons (red, yellow, green) on the left. The main area contains several labeled fields: 'ID:' with a text box containing 'k[n]'; 'Name:' with an empty text box; 'Port Type:' with a dropdown menu showing 'internal'; 'Initial Value:' with a text box containing '{0.1, 0.05, 0.01}[d0]' which is highlighted with a blue border; 'Units:' with a dropdown menu showing '(none)'; 'Constant:' with a dropdown menu showing 'false'; and 'SBOL DNA Component:' with a button labeled 'Associate SBOL'. At the bottom right, there are two buttons: 'Cancel' and 'OK'.

Parameter Editor

ID:

Name:

Port Type:

Initial Value:

Units:

Constant:

SBOL DNA Component:

Analysis of Models with Arrays

- Models with arrays can be flatten for analysis.
- *Arrays* package is syntatic sugar for SBML models.
- Eases the integration of the arrays package into existing analysis tools.
- A model using arrays can be converted to a new one with objects inlined.
- However, this approach has some limitations:
 - Restricts arrays objects to be statically computable (i.e. constant sizes).
 - Can cause model blow up.

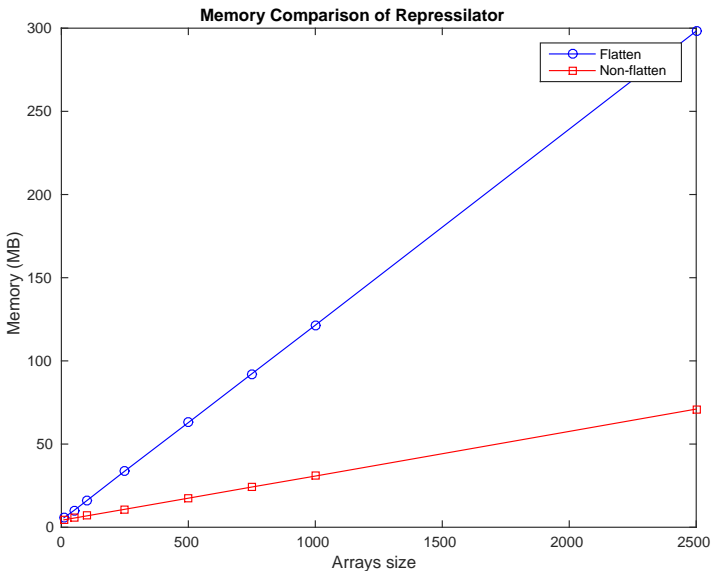
Model Blowup

Population	# of Species	Arrays Memory (MB)	Flattened Memory (MB)
10 ckts	60	3.568	4.0521
100 ckts	600	3.569	9.093
250 ckts	1500	3.568	19.803
1000 ckts	6000	3.568	69.2344
2500 ckts	15000	3.567	168.27
5000 ckts	30000	3.568	333.206

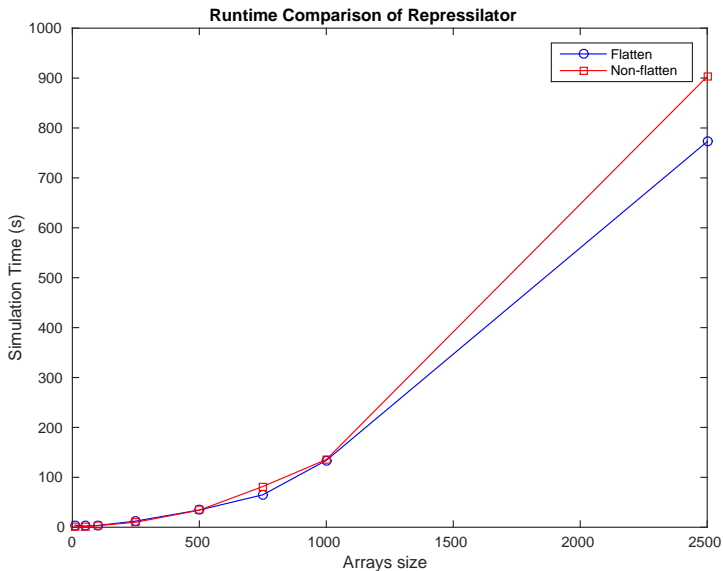
Model Analysis Technique

- Valuable information is lost with flattening.
- Another way is to simulate the original model as it is.
- In this approach, only one copy of each construct is necessary with the exception of variables.
- Variables, such as **species**, **parameters**, and **compartments**, among others, need a record of the state of each member of the array.
- Other constructs need a record of the size of the array.
- When performing arrayed **reactions**, **events**, **rules**, and others constructs that change the state of the simulation, the simulator iterates through each of the components of the array and performs the necessary updates.

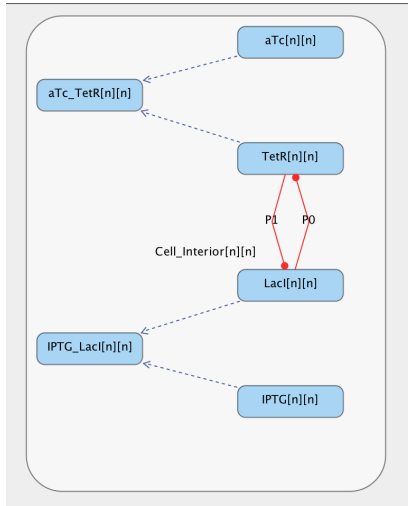
Results: Memory



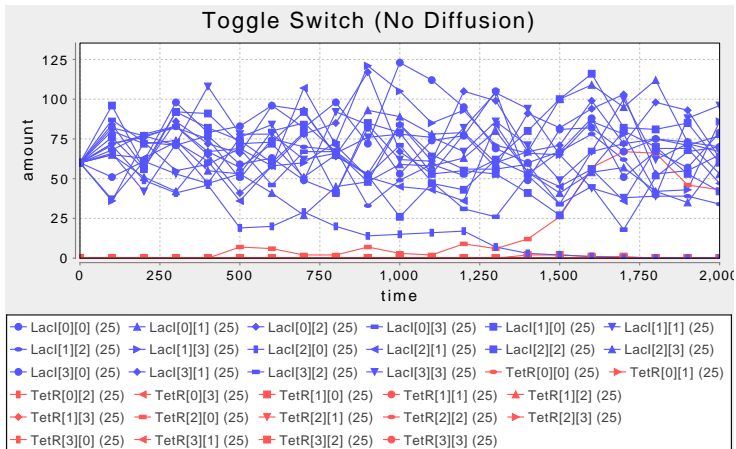
Results: Runtime



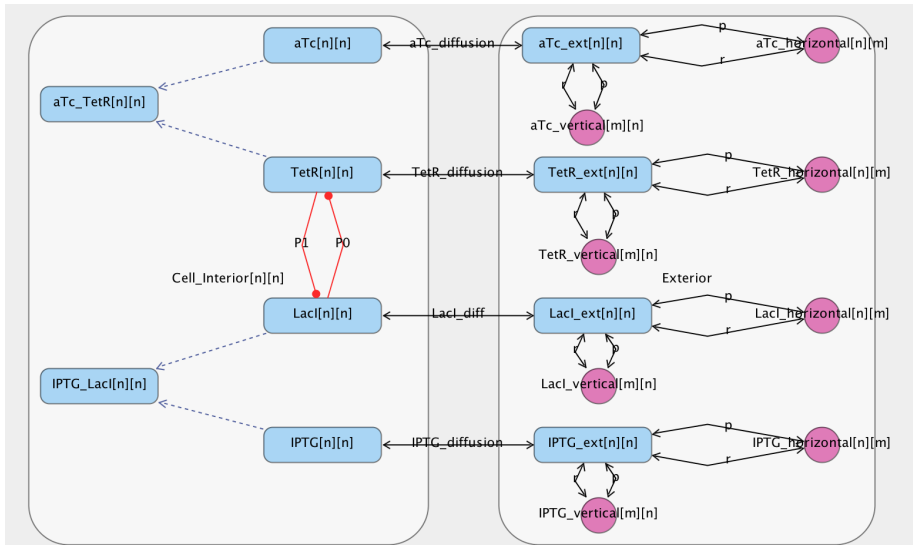
Genetic Toggle Switch



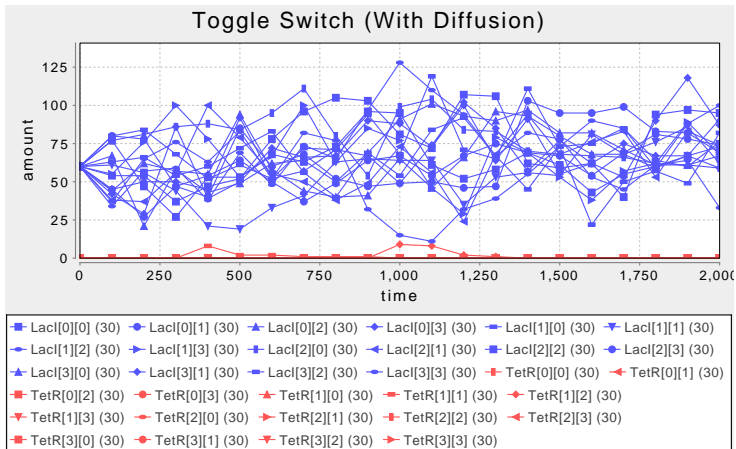
Results: Plot



Genetic Toggle Switch With Diffusion



Results: Plot



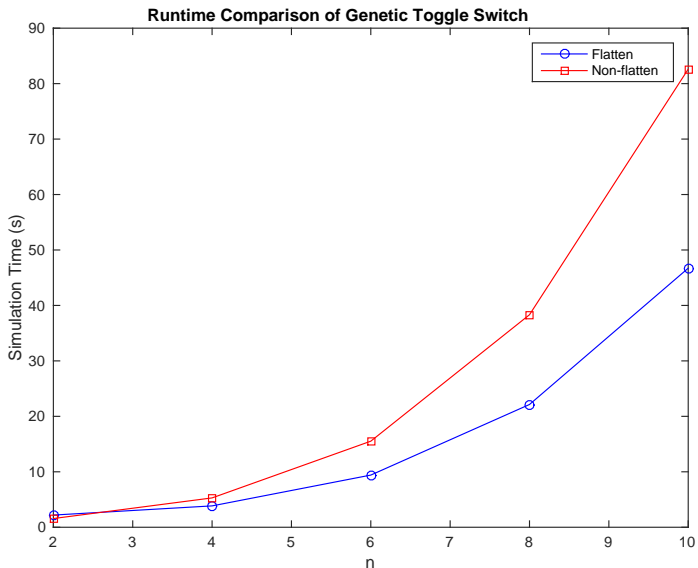
Results: Probabilities

Model	Number of Cells	Number of Failures	Probability
Without Diffusion	18750	219	$\sim 1.2 \%$
With Diffusion	18750	90	~ 0.5

Results: Memory



Results: Runtime



Discussion

- The *arrays* package enables efficient representation of regular constructs.
- Without arrays, the same population of genetic circuits would have to be instantiated explicitly multiple times.
- Flattening of arrays constructs eases package integration.
- Efficient simulation methods should exploit array structure.
- While the runtime is a bit higher, the method discussed in this talk reduces the memory usage substantially.

Future Work

- Re-encode the whole-cell model in SBML using arrays.
- Develop efficient ways to encode population models.
- Implement support for sparse array.

Acknowledgments



This work is supported by the National Science Foundation under Grant No. CCF-1218095.