Multicomponent Species: a proposal for SBML Level 3

http://www.cds.caltech.edu/~afinney/multi-component-species.pdf

Andrew Finney

University of Hertfordshire, UK

Molecular Sciences Institute, USA

Overview

- Requirements
- Tutorial
- Groups working on Generalized Reaction Systems
- Acknowledgements
- Conclusions

Limitations of SBML Level 2

- **`Species`** represents
  - single state
  - in a specific compartment
- Species states have to be enumerated
- Composition of species from components not represented, thus for example
  - bonds between components in complexes not represented
  - reactions forming complexes not represented

SBML Level 3 Requirements

- "Stateful" species
- Complexes
- Generalized reactions
  - Compact representations of networks that contain a combinatorial explosion of states
- Representing biochemical entities and reactions common to more than one compartment

SpeciesType

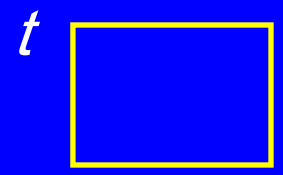
- Introducing **SpeciesType**
  - represents all chemical entities with the same structure and state
    - equivalence in the eye of the modeller
  - Independent of compartment
- **Species have an associated SpeciesType**
- **SpeciesType provides**
  - mechanism for relating Species that represent the same type of entity in different compartment
  - basis for defining entity state and structure
    - more later

SpeciesType defining Entity structure

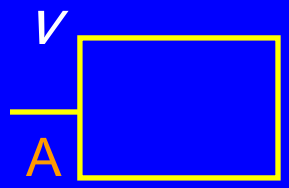
- **SpeciesType** can have a number of binding sites
- **SpeciesType** is graph of **SpeciesTypeInstance** nodes
  - Arcs are pairs of binding sites
  - **SpeciesTypeInstance** nodes refer to other **SpeciesType** structures

Species Types defining Structure: Examples

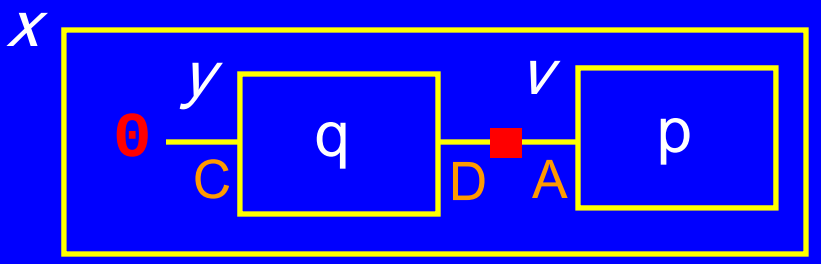
- Empty `SpeciesType` is indivisible



- Binding sites



- Binding

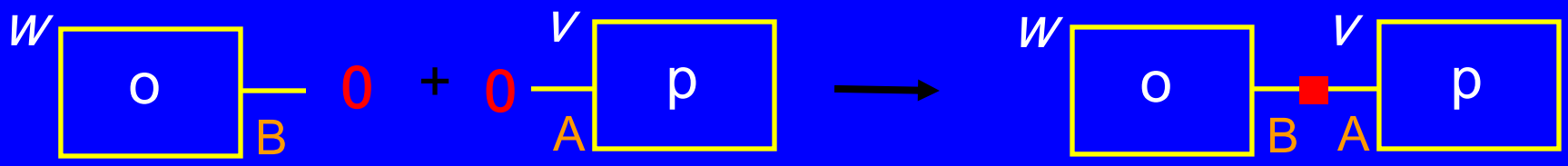


Reactions Generalized Across Species Types

- Reactants, Products and Modifiers are Species Graphs like **`SpeciesType`**
- These graphs contain wildcards which are copied from products to reactants
- Wildcards that do not ‘move’ can be omitted
  - syntactic sugar

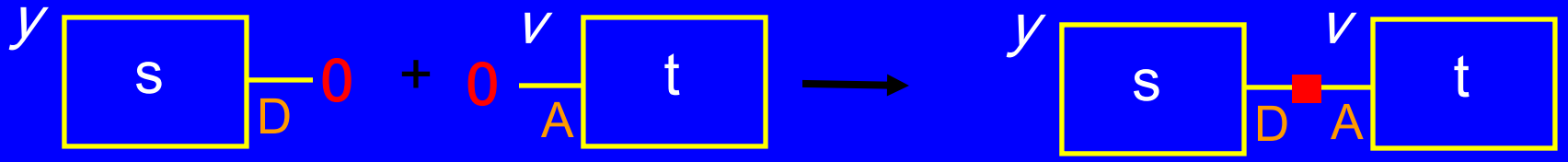
Reactions Generalized Across Species Types: Example

- Basic Reaction



- Generalized Reaction

- Generalized Reaction simplified



Groups tackling Generalization across Species Types

- **Larry Lok**
  - Molecular Sciences Institute
  - Moleculizer
  - Alpha project
    - Yeast Pheromone Pathway
  - Stochastic simulation generates ODE model
  - http://www.molsci.org/
- **T10 Group**
  - Los Alamos National Laboratory
  - BioNetGen
  - http://cellsignaling.lanl.gov/
  - Deterministic generation of ODE model
- **Brent Foy**
  - Wright State University
  - Stochastic Simulation

Groups tackling Generalization across Species Types - continued

- **Cell Systems Initiative (CSI)**
  - University of Washington
  - BALSA
  - http://www.csi.washington.edu/
  - Deterministic generation of Stochastic model

- **Carl Firth, Tom Shimizu, Nicolas Le Novère**
  - Cambridge University
  - Stochsim
  - http://info.anat.cam.ac.uk/groups/comp-cell/StochSim.html
  - Stochastic simulation of individual chemical entities
  - Entities only have state
    - complexes not represented in reaction system

Further Work

- Reaching high level of confidence
- Multiplicity - Blinov
  - 'array' of binding sites
- State Variables – Blinov, Le Novère
  - Decoupled from binding site
  - Examples
    - folding state
    - phosphorylation simplified
- Reactions having alternative kinetics when reactants in a given state - Lok
  - Normally reactions run in parallel
  - Group reactions where reactions within group run exclusively to each other
  - Example
    - enzyme catalyses faster in one specific state

Further Work (continued)

- Correct terminology
- Clarify Semantics
- Sort out compartments
  - Start from Blinov's ideas
  - How do we represent species spanning compartments?
  - What are the correct semantics?
  - Where is the variable?
- Simplification of Hierarchy?
  - Inheritance?
- Observables - Blinov
  - Single variable is the sum of a generalized set of species types
- More extensive set of examples in proposal

Vision

- This scheme can represent a wide range of biochemical systems that have a combinatorial explosion of states beyond signaling pathways
  - Examples:
    - Promoter Complexes
    - Metabolic Channeling
- This scheme can perhaps provide convenient representations of generic operations on very large molecules
  - Example: splicing DNA
  - Simulation in these cases is highly problematic
    - Topic for research in Systems Biology?

Acknowledgements

- This proposal has benefited from discussions with
  - Larry Lok
  - Nicolas Le Novère
  - Michael Blinov
  - Maria Schilstra
- Support from
  - Mike Hucka
  - Roger Brent

Conclusions

- Proposal introduces features which overcome limitations of SBML Level 2
- Proposal supports generalized reactions which eliminates combinatorial explosion
- Tools that deal with generalized reaction systems are emerging

Discussion

- Are you modeling systems that require a multicomponent species representation scheme?
- Would you be able to represent such a system in the proposed schemes?
  - Is something missing?
  - Could we do better?

- Do we require at least one implementation of these proposals before going forward?
- What minimal capability is required of this implementation?

- An informal working group is forming to develop proposal(s) for multicomponent species
  - If you're interested in this let us know