Introduction to SBML

SBML: Rationale

- The amount of experimental data on biological systems is growing at an impressive pace.
- To properly exploit/understand such vast information base, computational models has become mandatory.
- Many different modeling approaches can be applied, as well as many different SW tools.
- In order to be automatically analyzed, simulated, etc., models of any kind must be:
  - Formal
  - Machine-readable
- Model representations must address these properties; moreover must not be biased towards any particular approach, so to make a model usable across different approaches/tools.
- Developed solution: SBML, actually an XML grammar

What Models?

- SBML has been designed to formally describe systems where biological entities take part in *processes* (with temporal evolution).
- Typical example: networks of biochemical reactions.
- SBML is suitable for representing:
  - cell signaling pathways
  - metabolic pathways
  - biochemical reactions
  - gene regulation networks
  - …

SBML Benefits

1. Enabling the use of multiple tools
   - No need to rewrite models for each tool
2. Enabling models to be effectively shared among researchers that use different SW environments and tools
3. Ensuring the survival of models beyond the lifetime of the SW used to create them
4. SBML is based on XML, so it is neutral with respect to programming languages
   - XML-related libraries can be used over SBML
5. SBML is formal and unambiguous.

Reactions and Models

$Fru_{ex} \rightarrow Fru$
$GLC_{ex} \rightarrow Glc$
$ATP + Glc \rightarrow ADP + HexP$
$ATP + Fru \rightarrow ADP + HexP$
$2\ HexP \rightarrow Suc6P + UDP$
$Suc6P \rightarrow phos + Suc$
$Fru + HexP \rightarrow Suc + UDP$
$Suc \rightarrow Fru + Glc$
$HexP \rightarrow glycolysis$
$Suc \rightarrow Suc_{vac}$

Basic question:

- *Can you predict what a set of reactions like this will do when you start the system with different initial quantities?*

Beyond the enumeration of reactions, additional information is required.

Hints for an SBML Model

$$S_1 \xrightarrow{k_1[S_1]/([S_1]+k_2)} S_2$$
$$S_2 \xrightarrow{k_3[S_2]} S_3 + S_4$$

- Roughly speaking, SBML can encode models consisting of **entities** *linked* and *modified* by **processes**

- In the simple model above, "explicit" components are present:
  - reactant species and product species
  - reactions
  - reaction rates, and parameters in the rate expressions
- To analyze/simulate this system, other "implicit" components must be made explicit:
  - e.g., compartments where the species are located, and units on the various quantities.

Structure of an SBML Model

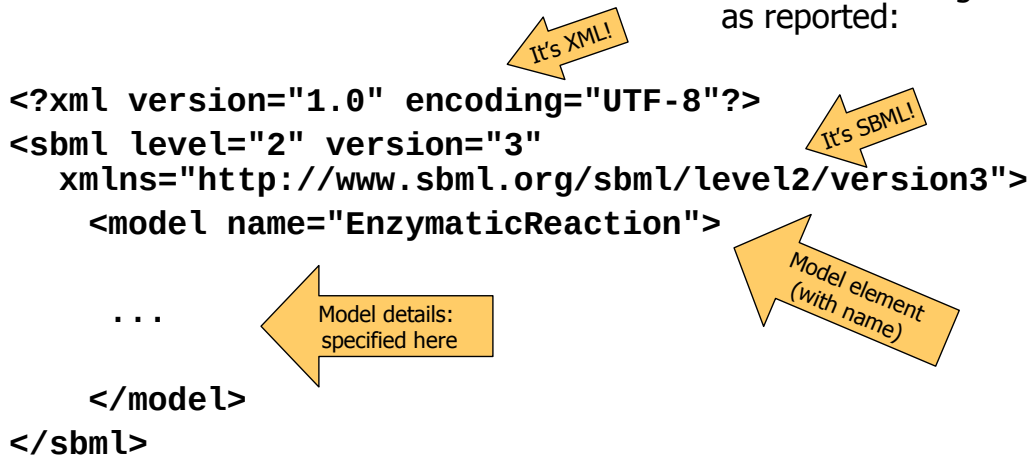
```
beginning of model definition

    list of function defs (opt.)
    list of unit defs (opt.)
    list of compartment types (opt.)
    list of species types (opt.)
    list of compartments (opt.)
    list of species (opt.)
    list of parameters (opt.)
    list of initial assignments (opt.)
    list of rules (opt.)
    list of constraints (opt.)
    list of reactions (opt.)
    list of events (opt.)

end of model definition
```

- A model is an XML element containing a number of optional elements
- The "types" elements are useful to develop well organized, flexible models
- Each model element has its own internal structure; e.g. parameters can be declared both at a global and at a local level

Example (I)

$$E + S \underset{k_{OFF}}{\overset{k_{ON}}{\rightleftarrows}} ES \xrightarrow{k_{CAT}} E + P$$

- Given the enzymatic reaction reported here, the corresponding SBML model will be organized as reported:




```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="3"
   xmlns="http://www.sbml.org/sbml/level2/version3">
     <model name="EnzymaticReaction">

       ...

     </model>
</sbml>
```

Example (II)

$$E + S \underset{k_{OFF}}{\overset{k_{ON}}{\rightleftharpoons}} ES \xrightarrow{k_{CAT}} E + P$$

- No function definition is inserted
- Units and compartments: implicit information

```
<listOfUnitDefinitions>
   <unitDefinition id="per_second">
       <listOfUnits>
           <unit kind="second" exponent="-1"/>
       </listOfUnits>
   </unitDefinition>
   <unitDefinition id="litre_per_mole_per_second">
       <listOfUnits>
           <unit kind="mole"   exponent="-1"/>
           <unit kind="litre"  exponent="1"/>
           <unit kind="second" exponent="-1"/>
       </listOfUnits>
   </unitDefinition>
</listOfUnitDefinitions>

<listOfCompartments>
   <compartment id="cytosol" size="1e-14"/>
</listOfCompartments>
         ...
```

Unit definition

Unit definition

Compartment definition

Example (III)

$$E + S \underset{k_{OFF}}{\overset{k_{ON}}{\rightleftharpoons}} ES \xrightarrow{k_{CAT}} E + P$$

- For each species in the model, location, id, initial amount (and name) are specified

```
   ...

<listOfSpecies>
  <species compartment="cytosol" id="ES" initialAmount="0"     name="ES"/>
  <species compartment="cytosol" id="P"  initialAmount="0"     name="P"/>
  <species compartment="cytosol" id="S"  initialAmount="1e-20" name="S"/>
  <species compartment="cytosol" id="E"  initialAmount="5e-21" name="E"/>
</listOfSpecies>

    ...
```

Example (IV)

$$E + S \underset{k_{OFF}}{\overset{k_{ON}}{\rightleftharpoons}} ES \xrightarrow{k_{CAT}} E + P$$

- Reactions must be described…

```
<listOfReactions>
   <reaction id="veq">
      <listOfReactants>
         <speciesReference species="E"/>
         <speciesReference species="S"/>
      </listOfReactants>
      <listOfProducts>
         <speciesReference species="ES"/>
      </listOfProducts>
      <kineticLaw>
         [redacted]
      </kineticLaw>
   </reaction>
   <reaction id="vcat" reversible="false">
         ...
   </reaction>
</listOfReactions>
```

Use of MathML

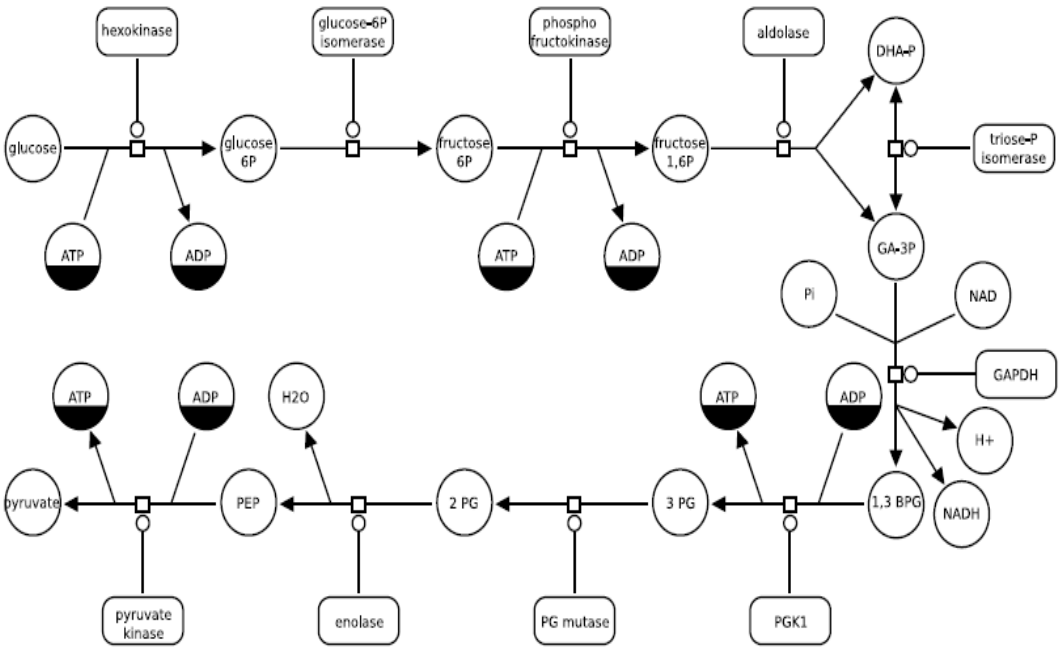
```
<math xmlns="http://www.w3.org/1998/Math/MathML">
   <apply>
      <times/>
      <ci>cytosol</ci>
         <apply>
            <minus/>
               <apply>
                  <times/>
                  <ci>kon</ci> <ci>E</ci> <ci>S</ci>
               </apply>
               <apply>
                  <times/>
                  <ci>koff</ci> <ci>ES</ci>
               </apply>
         </apply>
   </apply>
</math>
<listOfParameters>
   <parameter id="kon"  value="1000000"
     units="litre_per_mole_per_second"/>
   <parameter id="koff" value="0.2"
     units="per_second"/>
</listOfParameters>
```

"Local" Params

SBML and Python

Going Graphical: SBGN

An Example: Glycolysis



Compartments: IGF Signaling