

Storing SBML in the Berkeley DB XML database for rapid searching, retrieval and construction of biological models

Lafras Uys and Johann M. Rohwer

Triple-J Group for Molecular Cell Physiology
Department of Biochemistry
Stellenbosch University
South Africa

October 2006 / SBML Forum Tokyo

Outline

- 1 Modelling and SBML
- 2 Using Berkeley DB XML
- 3 Would this work?

Modelling

Building models

- Adding/subtracting reactions
- Altering rate equations and parameters

Analysing models

- Parameter substitutions
- Different events

Keep track of changes \Rightarrow Keep track of many files

Modelling

Building models

- Adding/subtracting reactions
- Altering rate equations and parameters

Analysing models

- Parameter substitutions
- Different events

Keep track of changes \Rightarrow Keep track of many files

Why a database?

Database vs. File System

www.sleepycat.com

A database such as BDB XML allows one to:

- Collect/Consolidate,
- Index and,
- Search/Query SBML

Advantages for us:

- BDB XML is open source,
- supported on lots of platforms and,
- has a Python API.

Why a database?

Database vs. File System

www.sleepycat.com

A database such as BDB XML allows one to:

- Collect/Consolidate,
- Index and,
- Search/Query SBML

Advantages for us:

- BDB XML is open source,
- supported on lots of platforms and,
- has a Python API.

Storing SBML

- Consolidate many SBML files/fragments into one db
- SBML is stored as key/value pairs

Grab SBML and dump into DB

```
from dbxml import *
manager = XmlManager()
container = manager.openContainer('sbml.bdbxml',DB_CREATE)
context = manager.createUpdateContext()

files = ['template.xml', 'species.xml', 'reactions.xml']
for f in files:
    path = '/home/lafras/dbxml/'+f
    stream = manager.createLocalFileInputStream(path)
    container.putDocument(f.rstrip('.xml'),stream,context)
container.sync()
```

An SBML fragment

```
<listOfReactions>
<reaction id="1">
  <listOfReactants>
    <speciesReference species="A"/>
  </listOfReactants>
  <listOfProducts>
    <speciesReference species="B"/>
  </listOfProducts>
  <kineticLaw>
    <math xmlns="http://www.w3...
    <apply>
      <times/>
      <ci> Ka </ci>
    ...
```

Retrieving original files

- Documents are available exactly as you deposited them

Retrieving an SBML fragment

```
from dbxml import *

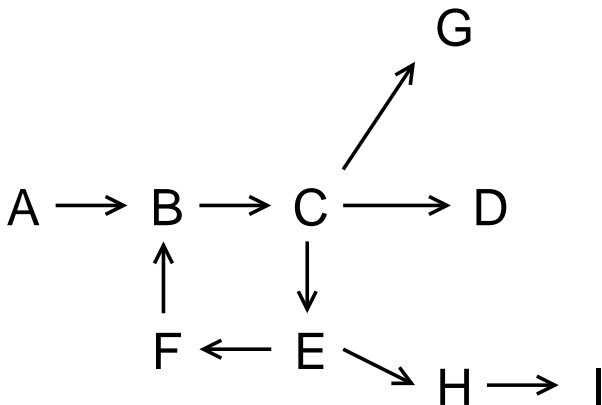
manager = XmlManager()
container = manager.openContainer('sbml.bdbxml')

print container.getDocument("template").getContent()
```

Outputs

```
$ python "getDoc.py"
<?xml version="1.0" encoding="UTF-8"?>
<sbml>
  <model>
    </model>
</sbml>
```

A hypothetical scheme



Assembling SBML fragments

```
from dbxml import *

manager = XmlManager()
container = manager.openContainer('sbml.bdbxml')
context = manager.createQueryContext()

query = """
<sbml><model><listOfSpecies>
{let $available_species := (
  for $reaction in collection('sbml.bdbxml')/listOfReactions/reaction
  where $reaction/listOfReactants/speciesReference/@species='C'
  return $reaction/*/speciesReference/@species
)}
for $specie in collection('sbml.bdbxml')/listOfSpecies/species
  where $specie/@id=$available_species
  return $specie }
</listOfSpecies><listOfReactions>
{for $reaction in collection('sbml.bdbxml')/listOfReactions/reaction
  where $reaction/listOfReactants/speciesReference/@species='C'
  return $reaction }
</listOfReactions></model></sbml>"""

q = manager.prepare(query,context)
results = q.execute(context)
result = results.next()

print result.asString()
```

Other advantages to BDB XML

- Concurrency - distributed processes with shared data
- Data recovery, transaction support etc. . .

Summary

- BDB XML makes it easier to create order when using SBML
- Outlook
 - Start using BDB XML with real modelling projects
 - Use BDB XML as standard data storage for SBML

Acknowledgements

- Prof. Johann M. Rohwer
- Dr. Brett Olivier
- Triple-J group for molecular cell physiology
- National Bioinformatics Network, ZA

So what do you think?