

Simulation Experiment Description Markup Language (SED-ML) : Level 1 Version 1 (Release Candidate 1)

March 4, 2011

Editors

Dagmar Waltemath
Frank T. Bergmann
Richard Adams
Nicolas Le Novère

University of Rostock, Germany
University of Washington, Seattle, USA
University of Edinburgh, UK
European Bioinformatics Institute, UK

The latest release of the Level 1 Version 1 specification is available at
<http://biomodels.net/sed-ml#sedmlResources>

To discuss any aspect of the current SED-ML specification as well as language details, please send your messages to the mailing list
sed-ml-discuss@lists.sourceforge.net.

To get subscribed to the mailing list, please write to the same address
sed-ml-discuss@lists.sourceforge.net.

To contact the authors of the SED-ML specification, please write to
sed-ml-editors@lists.sourceforge.net



1	Introduction	3
1.1	Motivation: A sample experiment	3
1.2	Conventions used in this document	6
2	Concepts used in SED-ML	10
2.1	MathML subset	10
2.2	URI Scheme	11
2.3	XPath usage	12
2.4	KiSAO	13
2.5	SED-ML archives	13
2.6	SED-ML resources	14
3	General attributes and classes	15
3.1	The <code>id</code> attribute	15
3.2	The <code>name</code> attribute	15
3.3	The SEDBase class	16
3.4	The SED-ML class	19
3.5	Reference relations	21
3.6	Variable class	24
3.7	Parameter class	25
3.8	ListOf* containers	27
4	SED-ML Components	32
4.1	Model	32
4.2	Change	34
4.3	Simulation	41
4.4	Task	44
4.5	DataGenerator	45
4.6	Output	45
4.7	Curve	48
4.8	Surface	51
4.9	DataSet	53
A	SED-ML UML Overview	56
B	Overview of SED-ML	57
B.1	Conventions	57
C	XML Schema	62
D	Examples	68
D.1	Le Loup Model (CellML)	68
D.2	The IkappaB-NF-kappaB signaling module (SBML)	70

1 Introduction

As Systems Biology transforms into one of the main fields in life sciences, the number of available computational models is growing at an ever increasing pace. At the same time, their size and complexity are also increasing. The need to build on existing studies by reusing models therefore becomes more imperative. It is now generally accepted that one needs to be able to exchange the biochemical and mathematical structure of models. The efforts to standardise the representation of computational models in various areas of biology, such as the *Systems Biology Markup Language* (SBML, [Hucka et al. \[2003\]](#)), *CellML* [Lloyd et al. \[2004\]](#) or *NeuroML* [Goddard et al. \[2001\]](#), result in an increase of the exchange and re-use of models. However, the description of the structure of models is not sufficient to enable the reproduction of simulation results. One also needs to describe the procedures the models are subjected to, as described by the *Minimum Information About a Simulation Experiment (MIASE)* [[Waltemath et al., 2011](#)].

This document presents Level 1 Version 1 of the *Simulation Experiment Description Markup Language* (SED-ML), a format that allows for the encoding of simulation experiments. SED-ML files are encoded in the *eXtensible Markup Language* (XML) [[Bray et al., 2006](#)]. The SED-ML language is defined by an XML Schema [[Fallside et al., 2001](#)].

1.1 Motivation: A sample experiment

To demonstrate how a simulation experiment can be described simply and effectively, we make use of a rather simple, though famous, model that may yet display rich and variable behaviors. The simulation example is taken from [[Waltemath et al., 2011](#)].

The *repressilator* is a synthetic oscillating network of transcription regulators in *Escherichia coli* [[Elowitz and Leibler, 2000](#)]. The network is composed of the three repressor genes Lactose Operon Repressor (*lacI*), Tetracycline Repressor (*tetR*) and Repressor CI (*cI*), which code for proteins binding to the promoter of the other, blocking their transcription. The three inhibitions together in tandem, form a cyclic negative-feedback loop. To describe the interactions of the molecular species involved in the network, the authors built a simple mathematical model of coupled first-order differential equations. All six molecular species included in the network (three mRNAs, three repressor proteins) participated in creation (transcription/translation) and degradation processes. The model was used to determine the influence of the various parameters on the dynamic behavior of the system. In particular, parameter values were sought which induce stable oscillations in the concentrations of the system components. Oscillations in the levels of the three repressor proteins are obtained by numerical integration.

1.1.1 A simple time-course simulation

The first experiment we intend to run on the model is the simulation that will lead to the oscillation shown in Figure 1c of the reference publication [[Elowitz and Leibler, 2000](#)]. The according simulation experiment can be described as:

1. Import the model identified by the Unified Resource Identifier (URI) [[Berners-Lee et al., 2005](#)] `urn:miriam:biomodels.db:BIOMD0000000012`.
2. Select a deterministic method.
3. Run a uniform time course simulation for 1000 min with an output interval of 1 min.
4. Plot the amount of *lacI*, *tetR* and *cI* against time in a 2D Plot.

Following those steps and performing the simulation in the simulation tool COPASI [[Hoops et al., 2006](#)] led to the result shown in Figure 1 on the following page.

1.1.2 Applying pre-processing

The fine-tuning of the model can be shown by adjusting parameters before simulation. When changing the initial values of the parameters *protein copies per promoter* and *leakiness in protein copies per promoter* the system's behavior switches from sustained oscillation to asymptotic steady-state. The adjustments leading to that behavior may be described as:

1. Import the model as above.



Figure 1: Time-course simulation of the repressilator model, imported from BioModels Database and simulated in COPASI. The number of repressor proteins *lacI*, *tetR* and *cI* is shown. (taken from Waltemath et al. [2011])



Figure 2: Time-course simulation of the repressilator model, imported from BioModels Database and simulated in COPASI after modification of the initial values of the protein copies per promoter and the leakiness in protein copies per promoter. The number of repressor proteins *lacI*, *tetR* and *cI* is shown. (taken from Waltemath et al. [2011])

2. Change the value of the parameter `tps_repr` from “0.0005” to “1.3e-05”.
3. Change the value of the parameter `tps_active` from “0.5 “ to “ 0.013”.
4. Select a deterministic method.
5. Run a uniform time course for the duration of 1000 min with an output interval of 1 min.
6. Plot the amount of *lacI*, *tetR* and *cI* against time in a 2D Plot.

Figure 2 shows the result of the simulation.

1.1.3 Applying post-processing

However, the raw numerical output of the simulation steps may be subjected to data post-processing before plotting or reporting. In order to describe the production of a normalized plot of the time-course in the first example (section 1.1.1), depicting the influence of one variable on another (in phase-planes), one could define the following further steps:

(Please note that the description steps 1 - 4 remain as given in section 1.1.2 above.)

5. Collect $PX(t)$ (*lacI*), $PY(t)$ (*tetR*) and $PZ(t)$ (*cI*).
6. Compute the highest value for each of the repressor proteins, $\max(PX(t))$, $\max(PY(t))$, $\max(PZ(t))$.



Figure 3: Time-course simulation of the repressilator model, imported from BioModels Database and simulated in COPASI, showing the normalized temporal evolution of repressor proteins *lacI*, *tetR* and *cI* in phase-plane. (taken from [Waltemath et al. \[2011\]](#))

7. Normalize the data for each of the repressor proteins by dividing each time point by the maximum value, i. e. $PX(t)/\max(PX(t))$, $PY(t)/\max(PY(t))$, and $PZ(t)/\max(PZ(t))$.
8. Plot the normalized **lacI** protein in function of the normalized **cI**, the normalized **cI** in function of the normalized **tetR** protein, and the normalized **tetR** protein against the normalized **lacI** protein in a 2D plot.

Figure 3 illustrates the result of the simulation after post-processing of the output data.

1.2 Conventions used in this document

SED-ML is specified as an XML Schema [W3C, 2004]. We also provide a UML Class diagram representation of that XML Schema (refer to appendix A). UML class diagrams are a subset of the *Unified Markup Language* notation (UML, [OMG, 2009]). Sample experiment descriptions are given as XML snippets that comply with the XML Schema.

1.2.1 UML Classes

A SED-ML UML class (Figure 4) consists of a class name (**ClassName**) and a number of attributes (**attribute**) each of a specific data type (**type**). The SED-ML UML specification does not make use of UML operations.



Figure 4: SED-ML UML Class with class names and attributes

SED-ML class names always begin with upper case letters. If they are composed of different words, the camel case style is used, as in e.g. **DataGenerator**.

1.2.2 UML Relationships

1.2.2.1 UML Relation Types



Figure 5: UML Class connectors

Links between classes specify the connection of objects with each other (Figure 5). The different relation types used in the SED-ML specification include aggregation, composite aggregation, and generalisation. The label on the line is called symbol (**label**) and describes the relation of the objects of both classes.

The **association** (Figure 6) indicates the existence of a connection between the objects of the participating classes. Often associations are directed to show how the label should be read (in which direction). Associations can be uni-directional (one arrowhead), or bidirectional (zero or two arrowheads).



Figure 6: UML Association

The **aggregation** (Figure 7 on the following page, top) indicates that the objects of the participating classes are connected in a way that one class (**Whole**) consists of several parts (**Part**). In an aggregation, the parts may be independent of the whole. For example, a car (**Whole**) has several parts called wheel (**Part**); however, the wheels can exist independently of the car while the car requires the wheels in order to function.



Figure 7: UML Aggregation

The **composite aggregation** (Figure 7, bottom) indicates that the objects of the participating classes are connected in a way that one class (**Whole**) consists of several parts (**Part**). In contrast to the aggregation, the subelements (**Part**) are dependent on the parent class (**Whole**). An example is that a university (**Whole**) consists of a number of departments (**Part**) which have a so-called “lifetime responsibility” with the university, e. g. if the university vanishes, the departments will vanish with it. [Bell, 2003].

The **generalisation** (Figure 8) allows to extend classes (**BaseClass**) by additional properties. The derived class (**DerivedClass**) inherits all properties of the base class and defines additional ones. In the given example, an instance of **DerivedClass** has two attributes **attribute1** and **attribute2**.

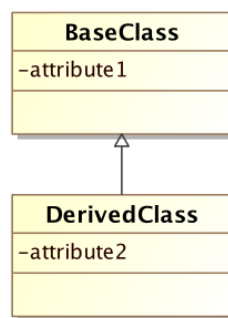


Figure 8: UML Generalisation

1.2.2.2 UML multiplicity

UML multiplicity defines the number of objects in one class that can be related to one object in the other class (also known as **cardinality**). Possible types of multiplicity include values (1), ranges (1..4), intervals (1,3,9), or combinations of ranges and intervals. The standard notation for “many” is the asterix (*).

Multiplicity can be defined for both sides of a relationship between classes. The default relationship is “many to many”. The example in Figure 9 expresses that a class is given by a professor, and a professor might give one to many classes.

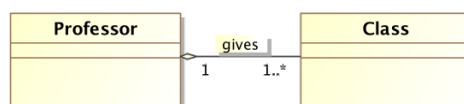


Figure 9: UML Multiplicity in an Aggregation

1.2.3 XML Schema language elements

The main building blocks of an XML Schema specification are

- simple and complex types
- element specifications
- attribute specifications

XML Schema [definitions](#) create new types, [declarations](#) define new elements and attributes. The definition of new (simple and complex) types can be based on a number of already existing, predefined types (string, boolean, float). Simple types are restrictions or extensions of predefined types. Complex types describe how attributes can be assigned to elements and how elements can contain further elements. The current SED-ML XML Schema only makes use of *complex type definitions*. An example for a complex type definition is given in listing 1: It shows the declaration of an element called **computeChange** that is

```

1 <xs:element name="computeChange">
2   <xs:complexType>
3     <xs:sequence>
4       <xs:element ref="listOfVariables" />
5       <xs:element ref="listOfParameters" />
6       <xs:element ref="math:math" />
7     </xs:sequence>
8     <xs:attribute name="target" type="xs:token" />
9   </xs:complexType>
10 </xs:element>

```

Listing 1: Complex Type definition of the SED-ML *computeChange* element

used in SED-ML to change mathematical expressions. The element is defined using an *unnamed* complex type which is build of further elements called **listOfVariables**, **listOfParameters**, and **math**. Additionally, the element **computeChange** has an attribute **target** declared. Please note that the definition of the elements inside the complex type are only referred to and will be found elsewhere in the schema.

The nesting of elements in the schema can be expressed using the **xs:sequence** (a sequence of elements), **xs:choice** (an alternative of elements to choose from), or **xs:all** (a set of elements that can occur in any order) concepts. The current SED-ML XML Schema only uses the *sequence* of elements.

1.2.3.1 Multiplicities

The standard multiplicity for each defined **element** is 1. Explicit multiplicity is to be defined using the **minOccurs** and **maxOccurs** attributes inside the complex type definition, as shown in listing 2.

```

1 <xs:element name="dataGenerator">
2   <xs:complexType>
3     <xs:sequence>
4       <xs:element ref="listOfVariables" minOccurs="0" />
5       <xs:element ref="listOfParameters" minOccurs="0" />
6       <xs:element ref="math:math" />
7     </xs:sequence>
8     [...]
9   </xs:complexType>
10 </xs:element>

```

Listing 2: Multiplicity for complex types in XML Schema

In this example, the **dataGenerator** type is build of a sequence of three elements: The **listOfVariables** element is not necessary for the definition of a valid **dataGenerator** XML structure (it may occur 0 times or once). The same is true for the **listOfParameters** element (it may as well occur 0 times or once). The **math** element, however, uses the implicit standard multiplicity – it must occur exactly 1 time in the **dataGenerator** specification.

1.2.3.2 Type extensions

XML Schema offers mechanisms to restrict and extend previously defined complex types. Extensions add element or attribute declarations to existing types, while restrictions restrict the types by adding further characteristics and requirements (facets) to a type. An example for a type extension is given in listing 3. The **sedML** element is an extension of the previously defined **SEDBase** type. It extends


```

1 <xs:element name="sedML">
2   <xs:complexType>
3     <xs:complexContent>
4       <xs:extension base="SEDBase">
5         <xs:sequence>
6           <xs:element ref="listOfSimulations" />
7           <xs:element ref="listOfModels" />
8           <xs:element ref="listOfTasks" />
9           <xs:element ref="listOfDataGenerators" />
10          <xs:element ref="listOfOutputs" />
11        </xs:sequence>
12        <xs:attribute name="version" type="xs:decimal" use="required" fixed="1" />
13        <xs:attribute name="level" type="xs:decimal" use="required" fixed="1" />
14      </xs:extension>
15    </xs:complexContent>
16  </xs:complexType>
17 </xs:element>

```

Listing 3: Definition of the sedML type through extension of SEDBase in SED-ML

SEDBase by a sequence of five additional elements (`listOfSimulations`, `listOfModels`, `listOfTasks`, `listOfDataGenerators`, and `listOfOutputs`) and a new attribute `version`.

2 Concepts used in SED-ML

2.1 MathML subset

The SED-ML specification allows for the encoding of pre-processing applied to the computational model, as well as for the encoding of post processing applied to the raw simulation data before output. The corresponding mathematical expressions are encoded using MathML 2.0 [Carlisle et al., 2001]. MathML is an international standard for encoding mathematical expressions using XML. It is also used as a representation of mathematical expressions in other formats, such as SBML and CellML, two of the languages supported by SED-ML.

2.1.1 MathML operations

In order to make the SED-ML format easier to adopt, at the beginning we restrict the MathML subset to the following operations:

- *token*: `cn`, `ci`, `csymbol`, `sep`
- *general*: `apply`, `piecewise`, `piece`, `otherwise`, `lambda`
- *relational operators*: `eq`, `neq`, `gt`, `lt`, `geq`, `leq`
- *arithmetic operators*: `plus`, `minus`, `times`, `divide`, `power`, `root`, `abs`, `exp`, `ln`, `log`, `floor`, `ceiling`, `factorial`
- *logical operators*: `and`, `or`, `xor`, `not`
- *qualifiers*: `degree`, `bvar`, `logbase`
- *trigonometric operators*: `sin`, `cos`, `tan`, `sec`, `csc`, `cot`, `sinh`, `cosh`, `tanh`, `sech`, `csch`, `coth`, `arcsin`, `arccos`, `arctan`, `arcsec`, `arccsc`, `arccot`, `arcsinh`, `arccosh`, `arctanh`, `arcsech`, `arccsch`, `arccoth`
- *constants*: `true`, `false`, `notanumber`, `pi`, `infinity`, `exponentiale`
- *MathML annotations*: `semantics`, `annotation`, `annotation-xml`

2.1.2 MathML Symbols

All the operations listed above only operate on *singular* values. However, as one of SED-ML's aims is to provide post processing on the results of simulation experiments, we need to enhance this basic set of operations by some aggregate functions. Therefore a defined set of MathML symbols that represent vector values are supported by SED-ML Level 1 Version 1. To simplify things for SED-ML L1V1 the only symbols to be used are the identifiers of variables defined in the `listOfVariables` of `DataGenerators`. These variables represent the data collected from the simulation experiment with the associated task.

2.1.3 MathML functions

The following aggregate functions are available for use in SED-ML Level 1 Version 1.

- *min*: Where the minimum of a variable represents the smallest value the simulation experiment yielded (listing 4).

```
1 <apply>
2   <csymbol encoding="text" definitionURL="http://www.biomodels.net/sed-ml/#min">
3     min
4   </csymbol>
5   <ci> variableId </ci>
6 </apply>
```

Listing 4: Example for the use of the MathML *min* function.

- *max*: Where the maximum of a variables represents the largest value the simulation experiment yielded (listing 5).
- *sum*: All values of the variable returned by the simulation experiment are added up (listing 6).
- *product*: All values of the variable returned by the simulation experiment are multiplied (listing 7).

```

1 <apply>
2   <csymbol encoding="text" definitionURL="http://www.biomodels.net/sed-ml/#max">
3     max
4   </csymbol>
5   <ci> variableId </ci>
6 </apply>

```

Listing 5: Example for the use of the MathML *max* function.

```

1 <apply>
2   <csymbol encoding="text" definitionURL="http://www.biomodels.net/sed-ml/#sum">
3     sum
4   </csymbol>
5   <ci> variableId </ci>
6 </apply>

```

Listing 6: Example for the use of the MathML *sum* function.

```

1 <apply>
2   <csymbol encoding="text" definitionURL="http://www.biomodels.net/sed-ml/#product">
3     product
4   </csymbol>
5   <ci> variableId </ci>
6 </apply>

```

Listing 7: Example for the use of the MathML *product* function.

These represent the only exceptions. At this point SED-ML Level 1 Version 1 does not define a complete algebra of vector values. For more information see the description of the [DataGenerator](#) class.

2.2 URI Scheme

URIs are needed at different points in SED-ML Level 1 Version 1: Firstly, they are the preferred mechanism to refer to model encodings. Secondly, they are used to specify the language of the referenced model. Thirdly, they enable addressing implicit model variables. Finally, annotations of SED-ML elements should be provided with a standardised annotation scheme.

The use of a standardised URI Scheme ensures long-time availability of particular information that can unambiguously be identified.

2.2.1 Model references

The preferred way for referencing a model from a SED-ML file is adopted from the [MIRIAM URI Scheme](#). MIRIAM enables identification of a data resource (in this case a model resource) by a predefined URN. A data entry inside that resource is identified by an ID. That way each single model in a particular model repository can be unambiguously referenced. To become part of MIRIAM resources, a model repository must ensure permanent and consistent model references, that is stable IDs.

One model repository that is part of MIRIAM resources is the [BioModels Database](#) [Li et al., 2010]. Its data resource name in MIRIAM is `urn:miriam:biomodels.db`. To refer to a particular model, a standardised identifier scheme is defined in [MIRIAM Resources](#)¹. The ID entry maps to a particular model in the model repository. That model is never deleted. A sample BioModels Database ID is `BIOMD0000000048`. Together with the data resource name it becomes unambiguously identifiable by the URN `urn:miriam:biomodels.db:BIOMD0000000048` (in this case referring to the 1999 Kholodenko model on EGFR signaling).

SED-ML recommends to follow the above scheme for model references, if possible. SED-ML does not specify how to resolve the URNs. However, MIRIAM Resources offers web services to do so². For the above example of the `urn:miriam:biomodels.db:BIOMD0000000048` model, the resolved URL may look like:

- <http://biomodels.caltech.edu/BIOMD0000000048> or
- <http://www.ebi.ac.uk/biomodels-main/BIOMD0000000048>

¹<http://www.ebi.ac.uk/miriam/>

²<http://www.ebi.ac.uk/miriam>

depending on the physical location of the resource chosen to resolve the URN.

An alternative means to obtain a model is provided by a MIASE archive. This is a stand-alone, self contained file containing necessary models and a single SED-ML file. Further information is provided in section 2.5. Further information on the [source](#) attribute referencing the model location is provided in section 4.1.2.

2.2.2 Language references

To specify the language a model is encoded in, a set of pre-defined SED-ML URNs can be used. The structure of SED-ML language URNs is `urn:sedml:language:name.version`. SED-ML allows to specify a model representation format very generally as being **XML**, if no standardised representation format has been used to encode the model. On the other hand, one can be as specific as defining a model being in a particular version of a language, as “SBML Level 2, Version 2, Revision 1”.

The list of URNs is available from <http://www.biomodels.net/sed-ml/#sedmlLanguage>. Further information on the [language](#) attribute is provided in section 4.1.1.

2.2.3 Implicit variables

Some variables used in an experiment are not explicitly defined in the model, but may be implicitly contained in it. For example, to plot a variable’s behaviour over time, that variable is defined in an SBML model, while *time* is not explicitly defined.

To overcome this issue and allow SED-ML to refer to such variables in a common way, the notion of *implicit variables* is used. Those variables are called **symbols** in SED-ML. They are defined following the idea of MIRIAM URNs and using the SED-ML URN scheme. The structure of the URNs is `urn:sedml:symbol:implicit variable`. To refer from a SED-ML file to the definition of *time*, for example, the URN is `urn:sedml:symbol:time`.

The list of predefined symbols is available from the SED-ML site on <http://biomodels.net/sed-ml>. From that source, a mapping of SED-ML symbols on possibly existing concepts in the single languages supported by SED-ML is provided.

2.2.4 Annotations

When annotating SED-ML elements with semantic [annotations](#), the [MIRIAM URI Scheme](#) should be used. In addition to providing the data type (e.g. PubMed) and the particular data entry inside that data type (e.g. **10415827**), the relation of the annotation to the annotated element should be described using the standardised [biomodels.net](#) [qualifier](#). The list of qualifiers, as well as further information about their usage, is available from <http://www.biomodels.net/qualifiers/>.

2.3 XPath usage

XPath is a language for finding information in an XML document [Clarke and DeRose \[1999\]](#). Within Level 1 Version 1, XPath version 1 expressions are used to identify nodes and attributes within an XML representation of a biological model in the following ways:

1. Within a [Variable](#) definition, where XPath identifies the model variable required for manipulation in SED-ML.
2. Within a [Change](#) definition, where XPath is used to identify the target XML to which a change should be applied.

For proper application, XPath expressions should contain prefixes that allow their resolution to the correct XML namespace within an XML document. For example, the XPath expression referring to a species *X* in an SBML model:

```
/sbml:sbml/sbml:model/sbml:listOfSpecies/sbml:species[@id='X']  ✓ -CORRECT
```

is preferable to

```
/sbml/model/listOfSpecies/species[@id='X']  ✗ -INCORRECT
```

which will only be interpretable by standard XML software tools if the SBML file declares no namespaces.

2.4 KiSAO

An important aspect of a simulation experiment is the simulation algorithm used to solve the system. But the sole reference of a simulation algorithm through its name in form of a string is error prone and unambiguous. Firstly, typing mistakes or language differences may make the identification of the intended algorithm difficult. Secondly, many algorithms exist with more than one name, having synonyms or various abbreviations that are commonly used.

These problems can be solved by using a controlled vocabulary to refer to a particular simulation algorithm. One attempt to provide such a vocabulary is the *Kinetic Simulation Algorithm Ontology* (KiSAO, <http://www.biomodels.net/kisao/>). KiSAO is a community-driven approach of classifying and structuring simulation approaches by model characteristics and numerical characteristics. Model characteristics include, for instance, the type of variables used for the simulation (such as discrete or continuous variables) and the spatial resolution (spatial or non-spatial descriptions). Numerical characteristics specify whether the system's behavior can be described as deterministic or stochastic, and whether the algorithms use fixed or adaptive time steps. Related algorithms are grouped together, producing classes of algorithms [Courtot et al., 2011]. KiSAO is available from BioPortal at <http://purl.bioontology.org/ontology/KiSAO>.

Although work is still at an early stage, the use of KiSAO is recommended when referring to a simulation algorithm from a SED-ML description. However, the use of KiSAO for the moment is limited. One may look up the algorithm that was used in the simulation experiment (through resolving the KiSAO ID) and then try and use one algorithm that is as similar to the original one as possible. KiSAO will become more supportive for SED-ML as soon as the ontology contains a wider range of relationships between different algorithms, as well as extended descriptions of the algorithm characteristics.

2.5 SED-ML archives

A [SED-ML archive](#) is a self-contained repository of all the resources necessary to run a simulation and display its output. It may be a convenient alternative if a model source URI cannot be resolved, or if an end-user is offline.

An archive is zipped folder containing one SED-ML file and any number of model files. By convention, the name of the archive will be the name of the SED-ML file contained in the archive, with the suffix “.sedx”. Each model file contained in the archive is referred to by a relative URI in the [source](#) attribute of the SED-ML document's [model](#) element.

For example, the contents of an archive, when unzipped, may be as follows:

Name of archive: Mysedml.sedx
Contents:

```
Mysedml.xml
model1.xml
model2.xml
```

Listing 8 shows how a model would be referenced in the SED-ML file *Mysedml.xml* in the above example.

Future versions of SED-ML may expand the permitted contents of an archive to include experimental data files or other resources.

```
1 <listOfModels>
2   <model id="m0001" language="urn:sedml:language:sbml"
3     source="model1.xml">
4     <listOfChanges>
5       <change>
6         [MODEL PRE-PROCESSING]
7       </change>
8     </listOfChanges>
9   </model>
10 </listOfModels>
```

Listing 8: Usage of relative URIs to reference a model in a SED-ML archive

2.6 SED-ML resources

SED-ML is part of the biomodels.net initiative <http://www.biomodels.net>. Information on SED-ML can be found on <http://www.biomodels.net/sed-ml>.

The SED-ML XML Schema, the UML schema and related implementations, libraries, validators and so on can be found on the SED-ML sourceforge project page <http://sed-ml.svn.sourceforge.net/>.

3 General attributes and classes

In this section we introduce attributes and concepts used repeatedly throughout the SED-ML specification.

3.1 The `id` attribute

Most objects in SED-ML carry an `id` attribute. The `id` attribute, if it exists for an object, is always required and identifies SED-ML constituents unambiguously. The data type for `id` is `SId` which is a datatype derived from the basic XML type `string`, but with restrictions about the characters permitted and the sequences in which those characters may appear. The definition is shown in Figure 10.

```
letter ::= 'a'..'z','A'..'Z'
digit  ::= '0'..'9'
idChar ::= letter | digit | '_'
SId    ::= ( letter | '_' ) idChar*
```

Figure 10: The definition of the type `SId`

For a detailed description see also the SBML specification on the “Type `SId`” [Hucka et al., 2010, p. 11].

All `ids` have a global scope, i. e. the `id` must be unambiguous throughout a whole SED-ML document. As such it identifies the constituent it is related to.

An example for a defined `id` is given in listing 9. The defined model carries the ID `m00001`. If the model

```
1 <model id="m00001" language="urn:sedml:language:sbml" source="urn:miriam:biomodels.db:BIOMD0000000012">
2   [MODEL DEFINITION]
3 </model>
```

Listing 9: SED-ML identifier definition, e. g. for a model

is referenced elsewhere in the SED-ML document, it is referred to by that ID.

3.2 The `name` attribute

Besides an `id`, a SED-ML constituent may carry an optional `name`. However, names do not have identifying character; several SED-ML constituents may carry the same name. The purpose of the `name` attribute is to keep a human-readable name of the constituent, e. g. for display to the user. In the XML Schema representation, names are of the data type `String`.

Listing 10 extends the model definition in listing 9 by a model name.

```
1 <model id="m00001" name="Circadian oscillator" language="urn:sedml:language:sbml" source="
   urn:miriam:biomodels.db:BIOMD0000000012">
2   [MODEL DEFINITION]
3 </model>
```

Listing 10: SED-ML name definition, e. g. for a model

3.3 The SEDBase class

SEDBase is the base class of SED-ML Level 1 Version 1. All other classes are derived from it. As such it provides means to attach additional information on all other classes (Figure 11). That information can be specified by human readable **Notes** or custom **Annotations**.

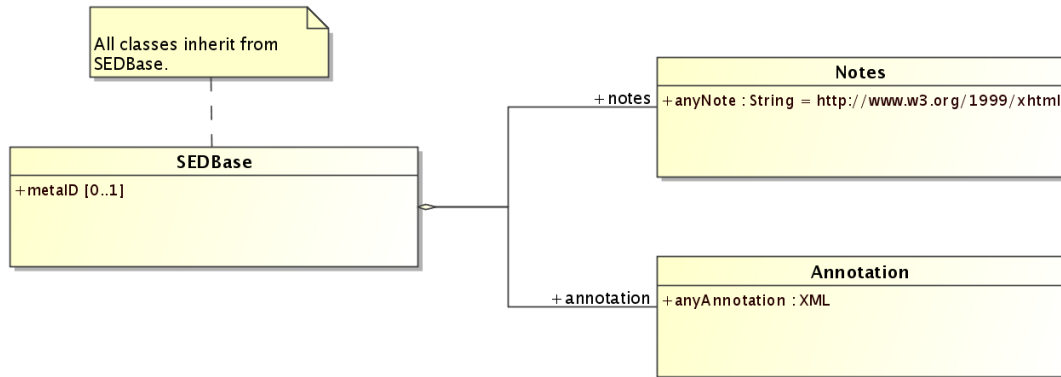


Figure 11: The SEDBase class

Table 3.3 shows all attributes and sub-elements for the **SEDBase** element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaID ^o	page 16
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 1: Attributes and nested elements for **SEDBase**. *xy^o* denotes optional elements and attributes.

3.3.1 metaid Attribute

The main purpose of the **metaid** attribute is to attach semantic annotations in form of the **Annotation** class to SED-ML elements. The type of **metaid** is XML ID and as such the **metaid** attribute is globally unique throughout the whole SED-ML document.

An example showing how to link a semantic annotation to a SED-ML object via the **metaid** is given in the **Annotation** class description.

3.3.2 The Notes class

A **note** is considered a human-readable description of the element it is assigned to. It serves to display information to the user. Instances of the **Notes** class may contain any valid XHTML [Pemberton et al., 2002], ranging from short comments to whole HTML pages for display in a Web browser. The namespace URL for XHTML content inside the **Notes** class is <http://www.w3.org/1999/xhtml>. It may either be declared in the **sedML XML element**, or directly used in top level XHTML elements contained within the **notes** element. For further options of how to set the namespace and detailed examples, please refer to [Hucka et al., 2010, p. 14].

Table 3.3.2 on the next page shows all attributes and sub-elements for the **Notes** element as defined by the SED-ML Level 1 Version 1 XML Schema. **Notes** does not have any further sub-elements defined in SED-ML, nor attributes associated with it.

Listing 11 on the following page shows the use of the **notes** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. In this example, the namespace declaration is inside

attribute	description
xmlns: string "http://www.w3.org/1999/xhtml"	page 19
sub-elements	
<i>well-formed content permitted in XHTML</i>	

Table 2: Attributes and nested elements for *Notes*. xy^o denotes optional elements and attributes.

```

1 <sedML [...]>
2 <notes >
3   <p xmlns="http://www.w3.org/1999/xhtml">The enclosed simulation description shows the oscillating
     behaviour of
4     the Repressilator model using deterministic and stochastic simulators.</p>
5 </notes>
6 </sedML>

```

Listing 11: The *notes* element

the **notes** element and the note is related to the **sedML** root element of the SED-ML file. A note may, however, occur inside *any* SED-ML XML element, except **note** itself and **annotation**.

3.3.3 The Annotation class

An **annotation** is considered a computer-processible piece of information. Annotations may contain any valid XML content. For further guidelines on how to use annotations, we would like to encourage the reading of the corresponding section in the SBML specification [Hucka et al., 2010, pp. 14-16]. The style of annotations in SED-ML is briefly described in section 2.2.4 on page 12.

Table 3.3.3 shows all attributes and sub-elements for the **Annotation** element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
<i>none</i>	
sub-elements	description
<i>none in the SED-ML namespace</i>	

Table 3: Attributes and nested elements for *Annotation*. xy^o denotes optional elements and attributes.

Listing 12 shows the use of the **annotation** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. In that example, a SED-ML **model** element is annotated with a reference to

```

1 <sedML>
2   [...]
3   <model id="model1" metaID="001" language="urn:sedml:language:cellml"
4     source="http://models.cellml.org/workspace/leloup_gonze_goldebeter_1999/@rawfile/
5       d6613d7e1051b3eff2bb1d3d419a445bb8c754ad/leloup_gonze_goldebeter_1999_a.cellml" >
6     <annotation>
7       <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
8         xmlns:bqmodel="http://biomodels.net/model-qualifiers/">
9         <rdf:Description rdf:about="#001">
10          <bqmodel:isDescribedBy>
11            <rdf:Bag>
12              <rdf:li rdf:resource="urn:miriam:pubmed:10415827"/>
13            </rdf:Bag>
14          </bqmodel:isDescribedBy>
15        </rdf:Description>
16      </rdf:RDF>
17    </annotation>
18  </model>
19 </sedML>

```

Listing 12: The *annotation* element

the original publication. The **model** contains an **annotation** that uses the [biomodels.net model-qualifier isDescribedBy](#) to link to the external resource **urn:miriam:pubmed:10415827**. In natural language the annotation content could be interpreted as “The model *is described by* the published article available from *pubmed* under ID *10643740*”. The example annotation follows the proposed [URI Scheme](#) suggested by the MIRIAM reference standard. The MIRIAM URN can be resolved to the PubMed (<http://pubmed.gov>) publication with ID 10415827, namely the article “Alternating oscillations and chaos in a model of two coupled biochemical oscillators driving successive phases of the cell cycle.” published by Romond et al. in 1999.

3.4 The SED-ML class

Each SED-ML Level 1 Version 1 document has a main class called SED-ML which defines the document's structure and content (Figure 12). It consists of several parts; the parts are all connected to the SED-ML class through aggregation: the [Model](#) class (for model specification, see section 4.1), the [Simulation](#) class (for simulation setup specification, see section 4.3), the [Task](#) class (for the linkage of models and simulation setups, see section 4.4), the [DataGenerator](#) class (for the definition of post-processing, see section 4.5), and the [Output](#) class (for the output specification, see section 4.6). All of them are shown in Figure 12 and will be explained in more detail in the relevant sections of this document.

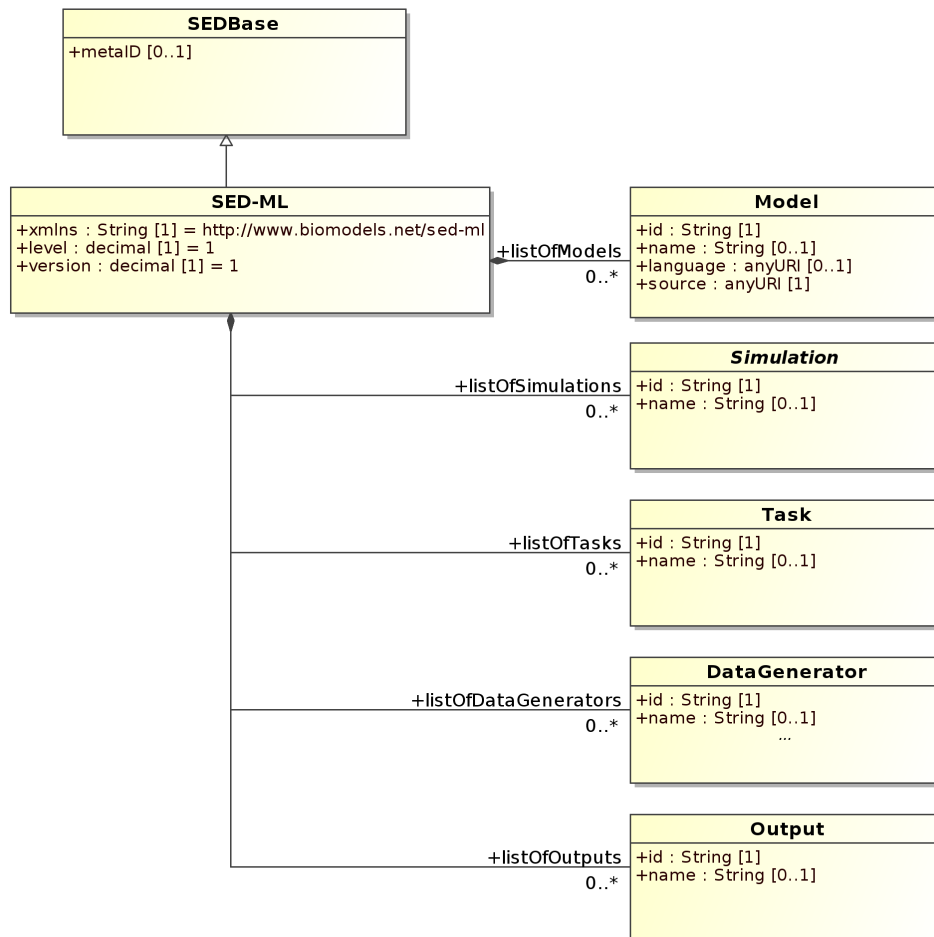


Figure 12: The sub-classes of SED-ML

Table 3.4 on the next page shows all attributes and sub-elements for the [SED-ML](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

A SED-ML document needs to have the SED-ML namespace defined through the mandatory `xmlns` attribute. In addition, the SED-ML `level` and `version` attributes are mandatory.

The basic XML structure of a SED-ML file is shown in listing 13 on the following page. The root element of each SED-ML XML file is the `sedML` element, encoding `version` and `level` of the file, and setting the necessary namespaces. Nested inside the `sedML` element are the five lists serving as containers for the encoded data (`listOfModels` for all models, `listOfSimulations` for all simulations, `listOfTasks` for all tasks, `listOfDataGenerators` for all post-processing definitions, and `listOfOutputs` for all output definitions).

3.4.1 The `xmlns` attribute

The `xmlns` attribute declares the namespace for the SED-ML document. The pre-defined namespace for SED-ML documents is <http://www.biomodels.net/sed-ml>.

attribute	description
metaID ^o	page 16
xmlns	page 19
level	page 20
version	page 20
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
listOfModels ^o	page 28
listOfSimulations ^o	page 28
listOfTasks ^o	page 29
listOfDataGenerators ^o	page 29
listOfOutputs ^o	page 30

Table 4: Attributes and nested elements for *SED-ML*. *xy^o* denotes optional elements and attributes.

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <sedML xmlns:math="http://www.w3.org/1998/Math/MathML"
3   xmlns="http://www.biomodels.net/sed-ml" level="1" version="1">
4   <listOfModels />
5   [MODEL REFERENCES AND APPLIED CHANGES]
6   <listOfSimulations />
7   [SIMULATION SETUPS]
8   <listOfTasks />
9   [MODELS LINKED TO SIMULATIONS]
10  <listOfDataGenerators />
11  [DEFINITION OF POST-PROCESSING]
12  <listOfOutputs />
13  [DEFINITION OF OUTPUT]
14 </sedML>

```

Listing 13: The *SED-ML* root element

In addition, SED-ML makes use of the *MathML* namespace <http://www.w3.org/1998/Math/MathML> to enable the encoding of mathematical expressions in MathML 2.0. SED-ML uses a subset of MathML as described in section 2.1 on page 10.

SED-ML *notes* use the XHTML namespace <http://www.w3.org/1999/xhtml>. The *Notes* class is described in section 3.3.2 on page 16.

Additional external namespaces might be used in *annotations*.

3.4.2 The level attribute

The current SED-ML *level* is “level 1”. Major revisions containing substantial changes will lead to the definition of forthcoming levels.

The level attribute is **required** and its value is a **fixed** decimal. For SED-ML Level 1 Version 1 the value is set to 1, as shown in the example in listing 13.

3.4.3 The version attribute

The current SED-ML *version* is “version 1”. Minor revisions containing corrections and refinements of SED-ML elements will lead to the definition of forthcoming versions.

The version attribute is **required** and its value is a **fixed** decimal. For SED-ML Level 1 Version 1 the value is set to 1, as shown in the example in listing 13.

3.5 Reference relations

The [reference](#) concept is used to refer to a particular element inside the SED-ML document. It may occur in five different ways in the SED-ML document:

1. as an association between a [Variable](#) and a [Model](#) ([modelReference](#)),
2. as an association between a [Variable](#) and a [Task](#) ([taskReference](#)),
3. as an association between a [Task](#) and the associated [Model](#) ([modelReference](#)) or
4. as an association between a [Task](#) and the [Simulation](#) ([simulationReference](#)).
5. as an association between an [Output](#) and a [DataGenerator](#) ([dataReference](#)),

The definition of a [Task](#) object demands a reference to a particular Model object ([modelReference](#), see section 3.5.1 on page 21); furthermore, the Task object must be associated with a particular Simulation object ([simulationReference](#), see section 3.5.3 on page 22).

Depending on the use of the [reference](#) relation in connection with a [Variable](#) object, it may take different roles:

- a. The [reference](#) association might occur between a Variable object and a Model object, if the variable is to define a [Change](#). In that case the **variable** element contains a [modelReference](#) to refer to the particular model that contains the variable used to define the change (see section 3.5.1 on page 21).
- b. If the [reference](#) is used as an association between a Variable object and a Task object inside the [dataGenerator](#) class, then the **variable** element contains a [taskReference](#) to unambiguously refer to an observable in a given task (see section 3.5.2 on page 22).

Four different types of [data references](#) exist in SED-ML Level 1 Version 1. They are used depending on the *type* of output for the simulation. A 2d plot has an [xDataReference](#) and a [yDataReference](#) assigned. A 3D plot has in addition a [zDataReference](#) assigned. To define a report, each data column has a [dataReference](#) assigned.

3.5.1 modelReference

The [modelReference](#) either represents a relation between a [Variable](#) object and a [Model](#) object, or a relation between a [Task](#) object and a [Model](#) object.

If pre-processing needs to be applied to a model before simulation, then the model update can be specified by creating a [Change](#) object. In the particular case that a change must be calculated with a mathematical function, variables need to be defined. To refer to an existing entity in a defined [Model](#), the [modelReference](#) is used.

The **modelReference** attribute of the **variable** element contains the **id** of a model that is defined in the document. Listing 14 shows the use of the **modelReference** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. In the example, a change is applied on model **m0001**. In

```
1 <model id="m0001" [...]>
2   <listOfChanges>
3     <computeChange>
4       <listOfVariables>
5         <variable id="v1" modelReference="cellML" target="/cellml:model/cellml:component[@cmeta:id='MP']/
          cellml:variable[@name='vsP']/@initial_value" />
6       [...]
7     </listOfVariables>
8     <listOfParameters [...] />
9     <math>
10      [CALCULATION OF CHANGE]
11    </math>
12  </computeChange>
13 </listOfChanges>
14 [...]
```

Listing 14: SED-ML *modelReference* attribute inside a variable definition of a *computeChange* element

the `computeChange` element a list of variables is defined. One of those variable is `v1` which is defined in another model (`cellML`). The XPath expression given in the `target` attribute identifies the variable in the model which carries the ID `cellML`.

The `modelReference` is also used to indicate that a `Model` object is used in a particular `Task`. Listing 15 shows how this can be done for a sample SED-ML document. The example defines two different tasks;

```
1 <listOfTasks>
2 <task id="t1" name="Baseline" modelReference="model1" simulationReference="simulation1" />
3 <task id="t2" name="Modified" modelReference="model2" simulationReference="simulation1" />
4 </listOfTasks>
```

Listing 15: SED-ML `modelReference` definition inside a `task` element

the first one applies the simulation settings of `simulation1` on `model1`, the second one applies the same simulation settings on `model2`.

3.5.2 `taskReference`

`DataGenerator` objects are created to apply post-processing to the simulation results before simulation output.

For certain types of post-processing `Variable` objects need to be created. These link to a defined `Task` from which the model that contains the variable of interest can be inferred. A `taskReference` association is used to realise that link from a `Variable` object inside a `DataGenerator` to a `Task` object. Listing 16 gives an example. The example shows the definition of a variable `v1` in a `dataGenerator` element. The

```
1 <listOfDataGenerators>
2 <dataGenerator id="tim3" name="tim mRNA (difference v1-v2+20)">
3 <listOfVariables>
4 <variable id="v1" taskReference="t1" [...] />
5 </listOfVariables>
6 <math />
7 </dataGenerator>
8 </listOfDataGenerators>
```

Listing 16: SED-ML `taskReference` definition inside a `dataGenerator` element

variable appears in the model that is used in task `t1`. The task definition of `t1` might look as shown in listing 17. Task `t1` references the model `model1`. Therefore we can conclude that the variable `v1`

```
1 <listOfTasks>
2 <task id="t1" name="task definition" modelReference="model1" simulationReference="simulation1" />
3 </listOfTasks>
```

Listing 17: Use of the reference relations in a task definition

defined in listing 16 targets an element of the model with ID `model1`. The targeting process itself will be explained in section 3.6.1 on page 24.

3.5.3 `simulationReference`

The `simulationReference` is used to refer to a particular `Simulation` in a `Task`. Listing 15 shows the reference to a defined simulation for a sample SED-ML document. In the example, both tasks `t1` and `t2` use the simulation settings defined in `simulation1` to run the experiment.

3.5.4 `dataReference`

The `dataReference` is used to refer to a particular `DataGenerator` instance from an `Output` instance. Listing 18 shows the reference to a defined data set for a sample SED-ML document. In the example, the output type is a 2D plot, which defines one curve with id `c1`. A curve must refer to two different

```

1 <listOfOutputs>
2   <plot2D id="p1" [...] >
3     <curve id="c1" xDataReference="dg1" yDataReference="dg2" />
4     [...]
5   </plot>
6 </listOfOutputs>

```

Listing 18: *Example for the use of data references in a curve definition*

data generators which describe how to procure the data that is to be plotted on the x-axis and y-axis respectively.

3.6 Variable class

Variables are references to already existing entities, either existing in one of the defined **models** or externally defined **symbols** (Figure 13). If the variable is defined through a reference to a model constituent,

Variable
+id : String [1] +name : String [0..1] +target : XPath [0..1] +symbol : String [0..1]

Figure 13: The Variable class

such as an SBML species, then the reference is specified using the **target** attribute. If the variable is defined through a reference to an external entity, then the **symbol** attribute is used. It holds a SED-ML **URI**. A **variable** is always placed inside a **listOfVariables**. **Symbol** and **target** must not be used together in a single instance of Variable.

Table 3.6 shows all attributes and sub-elements for the **Variable** element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
symbol	page 25
taskReference	page 22
modelReference	page 21
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 5: Attributes and nested elements for **Variable**. xy^o denotes optional elements and attributes.

A **variable** element must contain a **taskReference** if it occurs inside a **listOfVariables** inside a **dataGenerator** element. A **variable** element must contain a **modelReference** if it occurs inside a **listOfVariables** inside a **computeChange** element.

Listing 19 on the following page shows the use of the **variable** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

Listing 19 defines a variable **v1** (line 7) to compute a change on a model constituent (referenced by the **target** attribute on **computeChange** in line 5). The value of **v1** corresponds with the value of the targeted model constituent references by the **target** attribute in line 8. The second variable, **v2** (line 21), is used inside a **dataGenerator**. As the variable is **time** as used in **task1**, the **symbol** attribute is used to refer to the SED-ML URI for time (line 21).

3.6.1 The target attribute

An instance of **Variable** refers to a model constituent inside a particular **model** through an **XPath** expression stored in the required **target** attribute. XPath unambiguously identifies an element or attribute in an XML file.

Listing 20 on the next page shows the use of the **target** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. It should be noted that the identifier and names inside the SED-ML document do not have to comply with the identifiers and names that the model and its constituents carry in the model definition. In listing 20 on the following page, the variable with ID **v1** is defined. It is described as the **TetR protein**. The reference points to a species in the referenced SBML


```

1 <sedML>
2 <listOfModels>
3 <model [...]>
4 <listOfChanges>
5 <computeChange target="TARGET ELEMENT OR ATTRIBUTE">
6 <listOfVariables>
7 <variable id="v1" name="maximum velocity"
8 target="XPath TO A MODEL ELEMENT OR ATTRIBUTE IN ANY SPECIFIED MODEL" />
9 [FURTHER VARIABLE DEFINITIONS]
10 </listOfVariables>
11 [...]
12 </computeChange>
13 </listOfChanges>
14 [...]
15 </model>
16 [...]
17 </listOfModels>
18 <listOfDataGenerators>
19 <dataGenerator [...]>
20 <listOfVariables>
21 <variable id="v2" name="time" task="task1" symbol="urn:sedml:symbol:time" />
22 [FURTHER VARIABLE DEFINITIONS]
23 </listOfVariables>
24 </dataGenerator>
25 </listOfDataGenerators>
26 [...]
27 </sedML>

```

Listing 19: SED-ML variable definitions inside the `computeChange` element and inside the `dataGenerator` element

```

1 <listOfVariables>
2 <variable id="v1" name="TetR protein"
3 target="/sbml:sbml/sbml:listOfSpecies/sbml:species[@id='PY']" />
4 </listOfVariables>

```

Listing 20: SED-ML target definition

model. The particular species can be identified through its ID in the SBML model, namely `PY`. However, SED-ML does not forbid to use identical identifiers and names as in the referenced models neither. The following listing 21 is another valid example for the specification of a variable, but uses the same naming in the variable definition as in the original model (as opposed to listing 20):

```

1 <listOfVariables>
2 <variable id="PY" name="TetR protein"
3 target="/sbml:sbml/sbml:listOfSpecies/sbml:species[@id='PY']" />
4 </listOfVariables>

```

Listing 21: SED-ML variable definition using the original model identifier and name in SED-ML

The XPath expression used in the `target` attribute unambiguously leads to the particular place in the XML SBML model – the species is to be found in the `sbml` element, and there inside the `listOfSpecies` (listing 22 on the following page).

3.6.2 The `symbol` attribute

Symbols are predefined, implicit variables that can be called in a SED-ML file by referring to the defined URNs representing that variable's concept. The notion of implicit variables is explained in section 2.2.3 on page 12.

Listing 23 on the next page shows the use of the `symbol` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The example encodes a computed change of model `m001`. To specify that change, a symbol is defined (i.e. the SED-ML symbol for `time` is assigned to the variable `t1`). How to compute the change itself is explained in section 4.2.7.

3.7 Parameter class

The SED-ML **Parameter** class creates instances with a constant value (Figure 14 on page 27). SED-

```

1 <sbml [...]>
2 <listOfSpecies>
3 <species metaid="PY" id="PY" name="TetR protein" [...]>
4 [...]
5 </species>
6 </listOfSpecies>
7 [...]
8 </sbml>

```

Listing 22: Species definition in the referenced model (extracted from [urn:miriam:biomodels.db:BIOMD0000000012](https://miriam.org/uri/urn:miriam:biomodels.db:BIOMD0000000012))

```

1 <listOfVariables>
2 <variable id="t1" name="time" symbol="urn:sedml:symbol:time" />
3 </listOfVariables>

```

Listing 23: SED-ML symbol definition

ML uses parameters in two ways: Firstly, parameters may be defined in the [ComputeChange](#) class for describing the mathematical computation of a change of a model’s observable. Secondly, parameters may be part of a [DataGenerator](#) specification. In both cases the parameter definitions are local to the particular class defining them.

Table 3.7 shows all attributes and sub-elements for the [parameter](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaID ^o	page 16
id	page 15
name ^o	page 15
value	page 26
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 6: Attributes and nested elements for [parameter](#). xy^o denotes optional elements and attributes.

A parameter can unambiguously be identified through its given [id](#). It may additionally carry an optional [name](#). Each parameter has one associated [value](#).

Listing 24 shows the use of the [parameter](#) element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The listing shows the definition of a parameter [p1](#) with the [value](#)="40" assigned.

```

1 <listOfParameters>
2 <parameter id="p1" name="KM" value="40" />
3 </listOfParameters>

```

Listing 24: The definition of a parameter in SED-ML

3.7.1 The value attribute

Each [parameter](#) has exactly one fixed [value](#). The [value](#) attribute of XML data type [Double](#) is required for each [parameter](#) element.

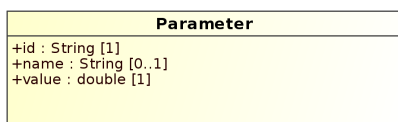


Figure 14: The *Parameter* class

3.8 ListOf* containers

SED-ML [listOf*](#) elements serve as containers for a collection of objects of the same type. For example, the [listOfModels](#) contains all [Model](#) objects of a SED-ML document. Lists do not carry any further semantics nor do they add additional attributes to the language. They might, however, be annotated with [Notes](#) and [Annotations](#) as they are derived from [SBase](#). All [listOf*](#) elements are optional in a SED-ML document.

3.8.1 listOfVariables: The variable definition container

SED-ML uses the [variable](#) concept to refer to existing entities inside a model. The container for all variables is [listOfVariables](#) (Figure 15). It includes all variables that need to be defined to either describe a change in the model by means of mathematical equations ([ComputeChange](#)) or to set up a [dataGenerator](#).

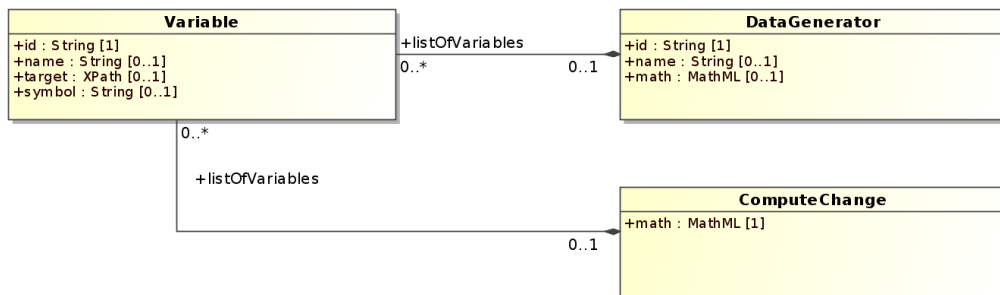


Figure 15: The SED-ML *listOfVariables* container

Listing 25 shows the use of the [listOfVariables](#) element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The [listOfVariables](#) is optional and may contain zero to many variables.

```

1 <listOfVariables>
2 <variable id="v1" name="maximum velocity" target="/cellml:model/cellml:component[@cmeta:id='MP']/
  cellml:variable[@name='vsP']/@initial_value" />
3 <variable id="v2" symbol="urn:sedml:symbol:time" />
4 </listOfVariables>

```

Listing 25: SED-ML *listOfVariables* element

3.8.2 listOfParameters: The parameter definition container

All parameters needed throughout the simulation experiment, either to compute a change on a model prior to simulation ([ComputeChange](#)) or to set up a [DataGenerator](#), are defined inside a [listOfParameters](#) (Figure 16 on the next page).

Listing 26 on the following page shows the use of the [listOfParameters](#) element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The element is optional and may contain zero to many parameters.

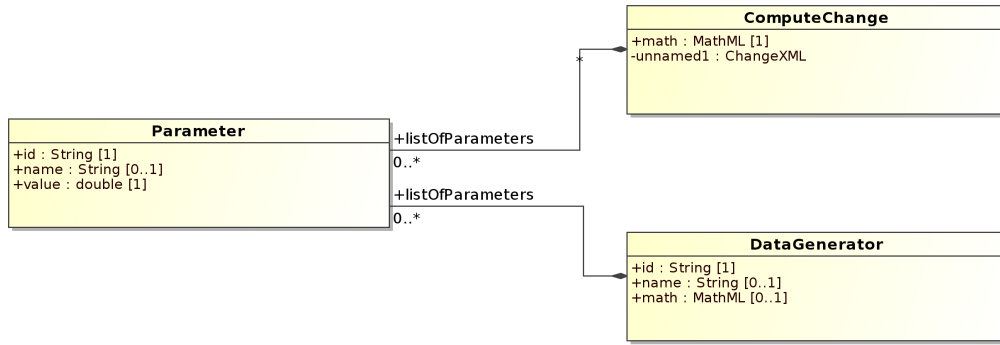


Figure 16: The SED-ML *listOfParameters* container

```

1 <listOfParameters>
2   <parameter id="p1" value="1" />
3   <parameter id="p2" name="Kadp_2" value="0.23" />
4 </listOfParameters>

```

Listing 26: SED-ML *listOfParameters* element

3.8.3 *listOfModels*: The model description container

In order to specify a simulation experiment, the participating models have to be defined. SED-ML uses the *listOfModels* container for all necessary models (Figure 17).

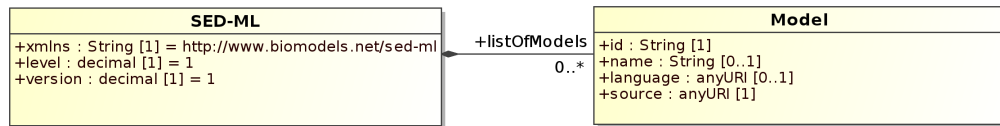


Figure 17: The SED-ML *listOfModels* container

Listing 27 shows the use of the *listOfModels* element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The *listOfModels* is optional and may contain zero to many models. However, if the Level 1 Version 1 document contains one or more **Task** elements, at least one **Model** element must be defined to which the **Task** element refers - see section 3.5.1 on page 21.

```

1 <listOfModels>
2   <model id="m0001" language="urn:sedml:language:sbml"
3     source="urn:miriam:biomodels.db:BIOMD0000000012" />
4   <model id="m0002" language="urn:sedml:language:cellml"
5     source="http://www.cellml.org/models/leloup_gonze_goldbeter_1999_version02" />
6 </listOfModels>

```

Listing 27: SED-ML *listOfModels* element

3.8.4 *listOfChanges*: The change definition container

The *listOfChanges* contains the defined changes to be applied to a particular *model* (Figure 18 on the next page). It always occurs as an optional subelement of the *model* element.

Listing 28 on the following page shows the use of the *listOfChanges* element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The *listOfChanges* is nested inside the *model* element.

3.8.5 *listOfSimulations*: The simulation description container

The *listOfSimulations* element is the container for *simulation* descriptions (Figure 19 on page 30).

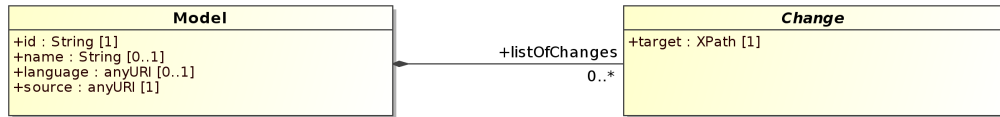


Figure 18: *The SED-ML listOfChanges container*

```

1 <model id="m0001" [...]>
2   <listOfChanges>
3     [CHANGE DEFINITION]
4   </listOfChanges>
5 </model>

```

Listing 28: *The SED-ML listOfChanges element, defining a change on a model*

Listing 29 shows the use of the `listOfSimulations` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. For all SED-ML Level 1 Version 1 documents, the encoded simulation

```

1 <listOfSimulations>
2   <simulation id="s1" [...]>
3     [UNIFORM TIMECOURSE DEFINITION]
4   </simulation>
5   <simulation id="s2" [...]>
6     [UNIFORM TIMECOURSE DEFINITION]
7   </simulation>
8 </listOfSimulations>

```

Listing 29: *The SED-ML listOfSimulations element, containing two simulation setups*

definitions are instances of the [Uniform Timecourse](#) class. The `listOfSimulations` is optional and may contain zero to many simulations. However, if the Level 1 Version 1 document contains one or more **Task** elements, at least one **Simulation** element must be defined to which the **Task** element refers - see section 3.5.3 on page 22.

3.8.6 *listOfTasks: The task specification container*

The `listOfTasks` element contains the defined tasks for the simulation experiment (Figure 20 on the next page).

Listing 30 shows the use of the `listOfTasks` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The `listOfTasks` is optional and may contain zero to many tasks. However, if

```

1 <listOfTasks>
2   <task id="t1" name="simulating v1" modelReference="m1" simulationReference="s1">
3     [FURTHER TASK DEFINITIONS]
4   </listOfTasks>

```

Listing 30: *The SED-ML listOfTasks element, defining one task*

the Level 1 Version 1 document contains a **DataGenerator** element with at least one **Variable** element, at least one **Task** must be defined to which variable(s) in the **DataGenerator** element refers - see section 3.5.2 on page 22.

3.8.7 *listOfDataGenerators: The post-processing container*

In SED-ML, all variable- and parameter values that shall be used in the [Output](#) class need to be defined as a [dataGenerator](#) beforehand. The container for those data generators is the `listOfDataGenerators` (Figure 21 on page 31).

Listing 31 on the following page shows the use of the `listOfDataGenerators` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

The `listOfDataGenerators` is optional and in general may contain zero to many **DataGenerators**. How-

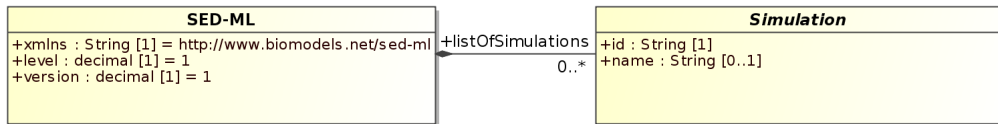


Figure 19: The *listOfSimulations* container

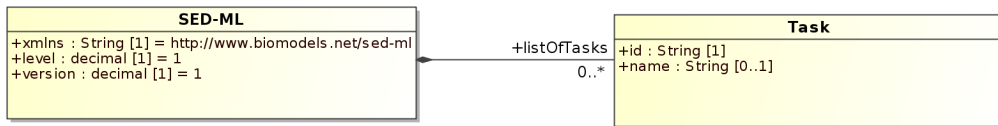


Figure 20: The SED-ML *listOfTasks* container

```

1 <listOfDataGenerators>
2 <dataGenerator id="d1" name="time">
3   [DATA GENERATOR DEFINITION FOLLOWING]
4 </dataGenerator>
5 <dataGenerator id="LaCI" name="LaCI repressor">
6   [DATA GENERATOR DEFINITION FOLLOWING]
7 </dataGenerator>
8 </listOfDataGenerators>

```

Listing 31: The *listOfDataGenerators* element, defining two data generators time and LaCI repressor

ever, if the Level 1 Version 1 document contains an **Output** element, at least one **DataGenerator** must be defined to which the **Output** element refers - see section 3.5.4 on page 22.

3.8.8 *listOfOutputs*: The output specification container

The *listOfOutputs* container holds the output specifications for a simulation experiment.

The output can be defined as either a *report*, a *plot2D* or as a *plot3D*.



Figure 21: *The SED-ML listOfDataGenerators container*

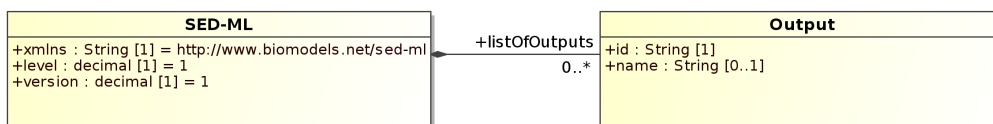


Figure 22: *The SED-ML listOfOutputs container*

Listing 32 shows the use of the `listOfOutputs` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The `listOfOutputs` is optional and may contain zero to many outputs.

```

1 <listOfOutputs>
2   <report id="report1">
3     [REPORT DEFINITION FOLLOWING]
4   </report>
5   <plot2D id="plot1">
6     [2D PLOT DEFINITION FOLLOWING]
7   </plot2D>
8 </listOfOutputs>
  
```

Listing 32: *The listOfOutput element*

4 SED-ML Components

In this section we describe the major components of SED-ML. We use the UML notation presented in section 1.2.1, and we show the use of SED-ML with XML examples. In addition, we provide a detailed BNMP diagram with explanation of the SED-ML workflow in Appendix B and an XML Schema in appendix C.

4.1 Model

The **Model** class defines the models to be used in the simulation experiment (Figure 23).



Figure 23: The SED-ML Model class

Each instance of the Model class has an unambiguous and mandatory **id**. An additional, optional **name** may be given to the model.

The **language** may be specified, defining the format the model is encoded in, if such a format exists. Example formats are SBML or CellML.

The **Model** class refers to the particular model of interest through the **source** attribute. The restrictions on the model reference are

- The model must be encoded in an XML format.
- To refer to the model encoding language, a reference to a valid definition of that XML format must be given (**language** attribute).
- To refer to a particular model in an external resource, an unambiguous reference must be given (**source** attribute).

A model might need to undergo pre-processings before simulation. Those pre-processings are specified in the SED-ML **Change** class.

Table 4.1 shows all attributes and sub-elements for the **model** element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
language ^o	page 33
source	page 33
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
change ^o	page 34

Table 7: Attributes and nested elements for **model**. *xy^o* denotes optional elements and attributes.

Listing 33 on the following page shows the use of the **model** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

The above **listOfModels** contains three models: The first model **m0001** is the Repressilator model taken from BioModels Database. The original model is available from [urn:miriam:biomodels.db:m0001](http://miriam.org/uri/urn:miriam:biomodels.db:m0001).


```

1 <listOfModels>
2 <model id="m0001" language="urn:sedml:language:sbml"
3   source="urn:miriam:biomodels.db:BIOMD0000000012">
4   <listOfChanges>
5     <change>
6       [MODEL PRE-PROCESSING]
7     </change>
8   </listOfChanges>
9 </model>
10 <model id="m0002" language="urn:sedml:language:sbml" source="m0001">
11   <listOfChanges>
12     [MODEL PRE-PROCESSING]
13   </listOfChange>
14 </model>
15 <model id="m0003" language="urn:sedml:language:cellml" source="http://www.cellml.org/models/
16   leloup_gonze_goldbeter_1999_version02">
17   [MODEL PRE-PROCESSING]
18 </model>
19 </listOfModels>

```

Listing 33: *SED-ML model element*

[BIOMD0000000012](#). For the SED-ML simulation, the model might undergo pre-processings, described in the `change` element (lines 5-7). Based on the description of the first model `m0001`, the second model is built. It refers to the model `m0001` in the `source` attribute, that is the modified version of the Repressilator model. `m0002` might then have even further changes applied to it on top of the changes defined in the pre-processing of `m0001`. The third model in the code example above (lines 13-15) is a different model in CellML representation. `m0003` is the model available from the given URL in the `source` attribute. Again, it might have additional pre-processing applied to it before used in the simulation.

4.1.1 The language attribute

The evaluation of a SED-ML document is required in order for software to decide whether or not it can be used for a particular simulation environment. One crucial criterion is the particular model representation language used to encode the model. A simulation software usually only supports a small subset of the representation formats available to model biological systems computationally.

To help software decide whether or not it supports a SED-ML description file, the information on the model encoding for each referenced model can be provided through the `language` attribute, as the description of a language name and version through an unrestricted `String` is error-prone. A prerequisite for a language to be fully supported by SED-ML is that a formalised language definition, e.g. an XML Schema, is provided online. SED-ML also defines a set of standard URIs to refer to particular language definitions. The list of URNs for languages so far associated with SED-ML is available from the SED-ML web site on <http://biomodels.net/sed-ml> (see again section 2.2.2 on page 12). To specify language and version, following the idea of MIRIAM URNs, the SED-ML URN scheme `urn:sedml:language:language name` is used. A model's language being "SBML Level 2 Version 2" can be referred to, for example, through the URN `urn:sedml:language:sbml.level-2.version-2`.

The `language` attribute is optional in the XML representation of a SED-ML file. If it is not explicitly defined in the SED-ML file, the default value for the `language` attribute is `urn:sedml:language:xml`, referring to any XML based model representation.

However, the use of the `language` attribute is strongly encouraged for two reasons. Firstly, it helps a user decide whether or not he is able to run the simulation, that is to parse the model referenced in the SED-ML file. Secondly, the language attribute is also needed to decide how to handle the implicit variables in the `Variable` class, as the interpretation of implicit variables depends on the language of the representation format. The concept of implicit variables has been introduced in section 2.2.3 on page 12.

4.1.2 The source attribute

To make a model available for the execution of a SED-ML file, the model `source` must be specified through a URI. The URI should preferably point to a public, consistent location that provides the model description file and follows the proposed [URI Scheme](#). References to curated, open model bases are recommended, such as the BioModels Database. However, any resource registered with MIRIAM resources³ can easily be referenced. Even without a MIRIAM URN, SED-ML can be used (see again

³<http://www.ebi.ac.uk/miriam/main/>

section 2.2.1 on page 11).

An example for the definition of a model, and using the [URI scheme](#) is given in listing 34. The example

```
1 <model id="m1" name="repressilator" language="urn:sedml:language:sbml"
2   source="urn:miriam:biomodels.db:BIOMD0000000012">
3   <listOfChanges>
4     [MODEL PRE-PROCESSING]
5   </listOfChanges>
6 </model>
```

Listing 34: The SED-ML source element, using the URI scheme

defines one model `m1`. `urn:miriam:biomodels.db:BIOMD0000000012` defines the source of the model code. The MIRIAM URN can be resolved into the SBML model stored in BioModels Database under ID `BIOMD0000000012` using the MIRIAM web service. The resulting URL is <http://www.ebi.ac.uk/biomodels-main/BIOMD0000000012>.

An example for the definition of a model and using a URL is given in listing 35. In the example one

```
1 <model id="m1" name="repressilator" language="urn:sedml:language:cellml"
2   source="http://models.cellml.org/exposure/bba4e39f2c7ba8af51fd045463e7bdd3/aguda_b_1999.cellml">
3   <listOfChanges />
4 </model>
```

Listing 35: The SED-ML source element, using a URL

model is defined. The `language` of the model is `CellML`. As the CellML model repository currently does not provide a MIRIAM URI for model reference, the `URL` pointing to the model code is used to refer to the model. The URL is given in the `source` attribute.

4.2 The Change class

SED-ML not only allows to use the sole model for simulation, but also enables the description of [changes](#) to be made on the model before simulation (Figure 24 on the next page). Changes can be of three distinct types:

1. Changes on attributes of the model's XML representation ([ChangeAttribute](#))
2. Changes on any XML snippet of the model's XML representation ([AddXML](#), [ChangeXML](#), [RemoveXML](#))
3. Changes based on mathematical calculations ([ComputeChange](#))

The [Change](#) class is abstract and serves as the container for different types of changes. Therefore, a SED-ML document will only contain the derived classes, i.e. [ChangeAttribute](#), [AddXML](#), [ChangeXML](#), [RemoveXML](#), or [ComputeChange](#).

Table 4.2 on the following page shows all attributes and sub-elements for the [change](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

Each Change has a [target](#) attribute that holds a valid XPath expression pointing to the XML element or XML attribute that is to undergo the defined changes.

4.2.1 The NewXML class

The `newXML` element provides a piece of XML code (Figure 24 on the next page). `NewXML` must hold a valid piece of XML which after insertion into the original model must lead to a valid model file, according to the model language specification (as given by the [language](#) attribute).

Table 4.2.1 on the following page shows all attributes and sub-elements for the `newXML` element as defined by the SED-ML Level 1 Version 1 XML Schema.

The `newXML` element is used at two different places inside SED-ML Level 1 Version 1:

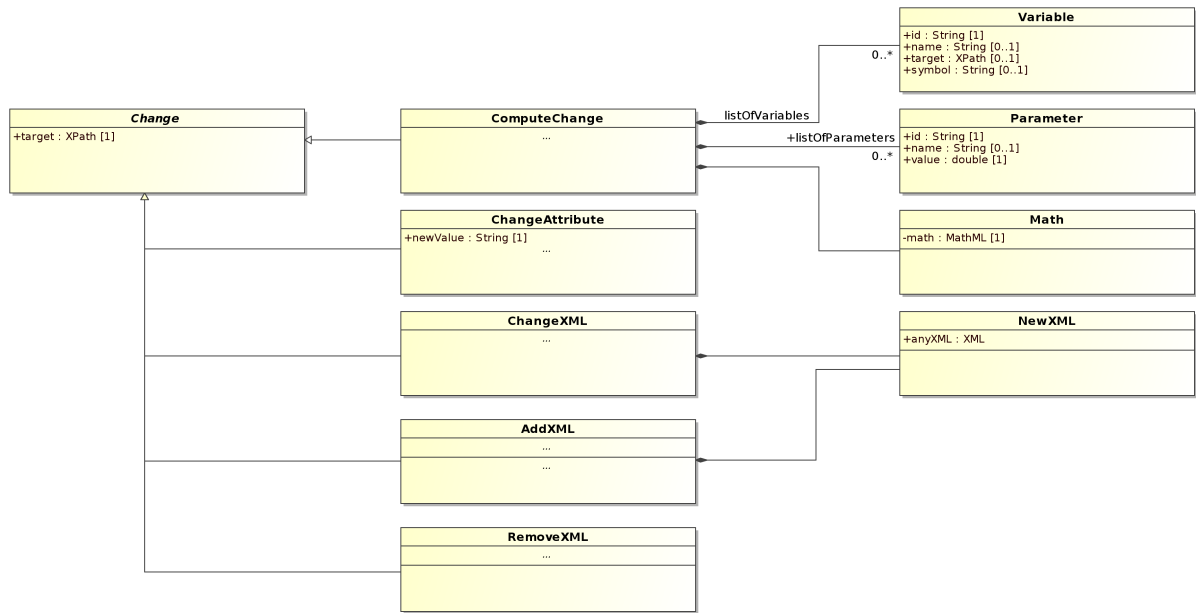


Figure 24: The SED-ML Change class

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
addXML ^o	page 35
changeXML ^o	page 36
removeXML ^o	page 37
changeAttribute ^o	page 38
computeChange ^o	page 39

Table 8: Attributes and nested elements for [change](#). xy^o denotes optional elements and attributes.

attribute	description
none	
sub-elements	description
anyXML	

Table 9: Attributes and nested elements for [newXML](#). xy^o denotes optional elements and attributes.

1. If it is used as a sub-element of the [addXML](#) element, then the XML it contains it is to be *inserted as a child* of the XML element addressed by the XPath.
2. If it is used as a sub-element of the [changeXML](#) element, then the XML it contains is to *replace* the XML element addressed by the XPath.

Examples are given in the relevant change class definitions.

4.2.2 The AddXML class

The [AddXML](#) class specifies a snippet of XML that is to be added as a child of the specified XPath [target](#) attribute (Figure 25 on the next page). The new piece of XML code is provided by the [NewXML](#) class.

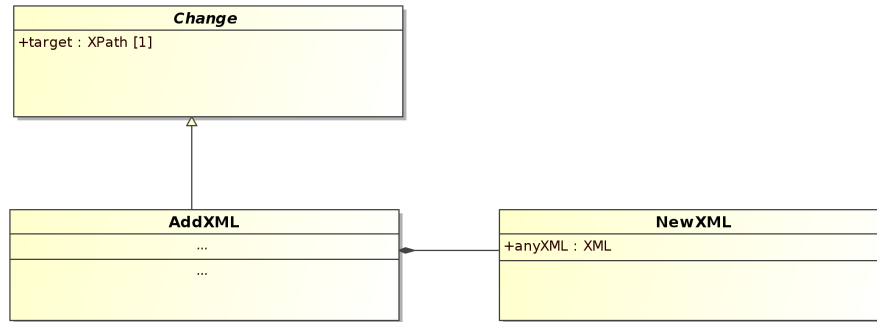


Figure 25: The SED-ML *AddXML* class

Table 4.2.2 shows all attributes and sub-elements for the `addXml` element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
newXML	page 34

Table 10: Attributes and nested elements for *addXML*. *xy^o* denotes optional elements and attributes.

An example for a change that adds an additional parameter to a model is given in listing 36.

```

1 <model language="urn:sedml:language:sbml" [..]>
2   <listOfChanges>
3     <addXML target="/sbml:sbml/sbml:model/sbml:listOfParameters" >
4       <newXML>
5         <parameter metaid="metaid_0000010" id="V_mT" value="0.7" />
6       </newXML>
7     </addXML>
8   </listOfChanges>
9 </model>

```

Listing 36: The *addXML* element with its *newXML* sub-element

The code of the model is changed so that a parameter with ID `V_mT` is added to its list of parameters. The `newXML` element adds an additional XML element to the original model. The element's name is `parameter` and it is added to the existing parent element `listOfParameters` that is addressed by the XPath expression in the `target` attribute.

4.2.3 The *ChangeXML* class

The `ChangeXML` class defines changes of any XML element in the model that can be addressed by a valid XPath expression (Figure 26 on the next page). The XPath is specified in the required `target` attribute (see again section 3.6.1 on page 24). The change of XML is specified in the `NewXML` class.

Table 4.2.3 on the following page shows all attributes and sub-elements for the `changeXml` element as defined by the SED-ML Level 1 Version 1 XML Schema.

An example for a change that adds an additional parameter to a model is given in listing 37. The code of

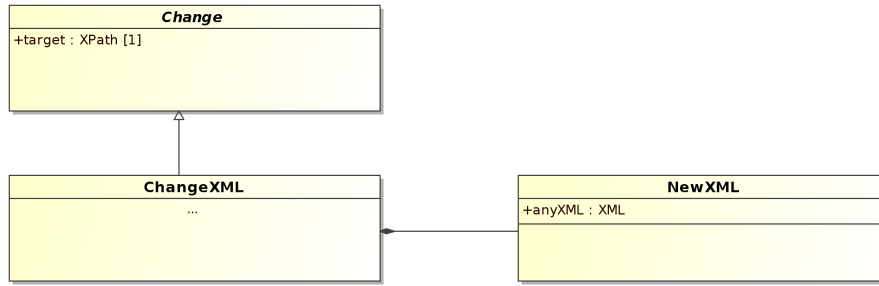


Figure 26: *The ChangeXML class*

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
newXML	page 34

Table 11: *Attributes and nested elements for [changeXML](#). xy^o denotes optional elements and attributes.*

```

1 <model [...]>
2   <listOfChanges>
3     <changeXML target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='V_mT']" >
4       <newXML>
5         <parameter metaid="metaid_0000010" id="V_mT_1" value="0.7" />
6         <parameter metaid="metaid_0000050" id="V_mT_2" value="4.6" />
7       </newXML>
8     </changeXML>
9   </listOfChanges>
10 </model>

```

Listing 37: *The changeXML element*

the model is changed in the way that its parameter with ID **V_{mT}** is substituted by two other parameters **V_{mT_1}** and **V_{mT_2}**. The **target** attribute defines that the parameter with ID **V_{mT}** is to be changed. The **newXML** element then specifies the XML that is to be exchanged for that parameter.

4.2.4 The RemoveXML class

The [RemoveXML](#) class can be used to delete the XML element of the model that is addressed by the XPath expression (Figure 27 on the next page).

The XPath is specified in the required **target** attribute.

Table 4.2.4 on the following page shows all attributes and sub-elements for the [removeXml](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

An example for the removal of an XML element from a model is given in listing 38.

```

1 <model [...]>
2   <listOfChanges>
3     <removeXML target="/sbml:sbml/sbml:model/sbml:listOfReactions/sbml:reaction[@id='J1']" />
4   </listOfChanges>
5 </model>

```

Listing 38: *The removeXML element*

The code of the model is changed by deleting the reaction with ID **V_{mT}** from the model's list of reactions.

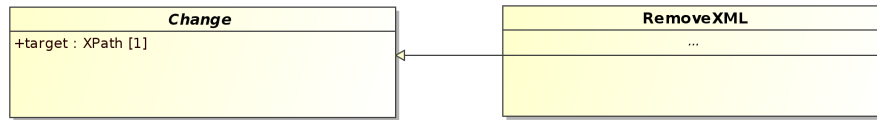


Figure 27: *The RemoveXML class*

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 12: *Attributes and nested elements for [removeXML](#). xy^o denotes optional elements and attributes.*

4.2.5 The ChangeAttribute class

The [ChangeAttribute](#) class allows to define updates on the XML attribute values of the corresponding model (Figure 28).

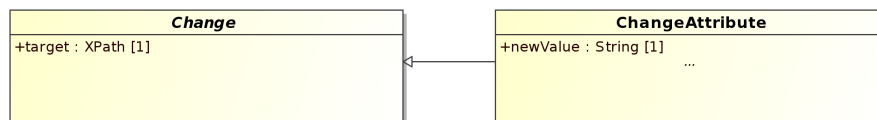


Figure 28: *The ChangeAttribute class*

The [ChangeXML](#) class covers the possibilities provided by the [ChangeAttribute](#) class. That is, everything that can be expressed by a [ChangeAttribute](#) construct can also be expressed by a [ChangeXML](#). However, both concepts exist to allow for being very specific in defining changes. It is recommended to use the [ChangeAttribute](#) for any changes of an XML attribute, and to use the more general [ChangeXML](#) for all other cases.

[ChangeAttribute](#) requires to specify the [target](#) of change, i.e. the location of the addressed XML attribute, and also the [new value](#) of that attribute.

Table 4.2.5 shows all attributes and sub-elements for the [changeAttribute](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
newValue	page 39
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 13: *Attributes and nested elements for [ChangeAttribute](#). xy^o denotes optional elements and attributes.*

4.2.6 The `newValue` attribute

The mandatory `newValue` attribute assigns a new value to the targeted XML attribute.

The example in listing 39 shows the update of the initial concentration of two parameters inside an SBML model.

```

1 <model id="model1" name="Circadian Chaos" language="urn:sedml:language:sbml"
2   source="urn:miriam:biomodels.db:BIOMD0000000021">
3   <listOfChanges>
4     <changeAttribute target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='V_mT']/@value"
5       newValue="0.28"/>
6     <changeAttribute target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='V_dT']/@value"
7       newValue="4.8"/>
8   </listOfChanges>
9 </model>

```

Listing 39: The `changeAttribute` element and its `newValue` attribute

4.2.7 The `ComputeChange` class

The `ComputeChange` class permits changes on any element of the XML file addressable by an XPath expression (Figure 29). The changes are described by mathematical expressions using a subset of MathML

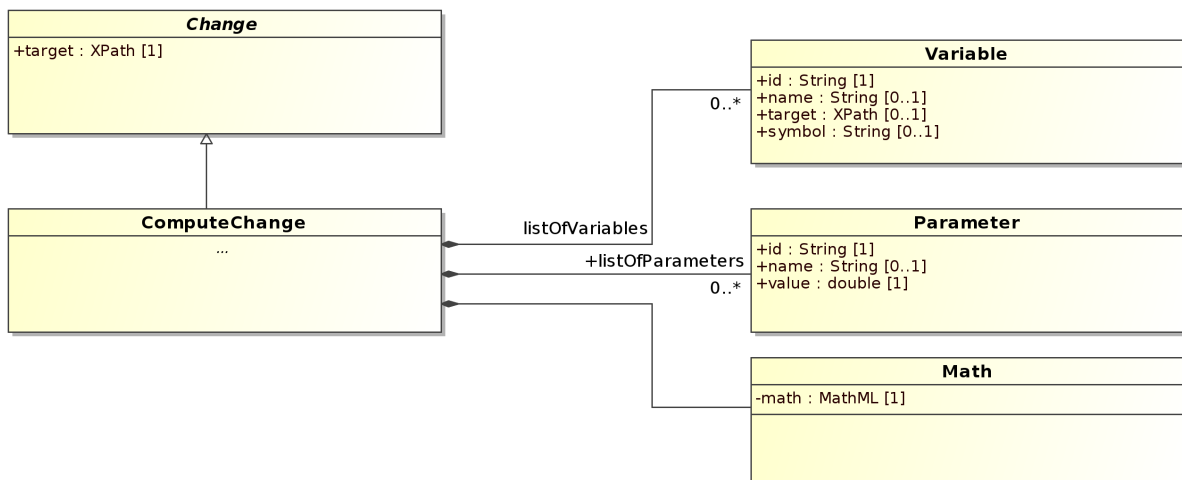


Figure 29: The `ComputeChange` class

(see again section 2.1 on page 10).

To define a computable change it is often necessary to introduce additional parameters for the mathematics. Therefore, the `parameters` first need to be defined. They can then be referenced through their ID. To use model variables for the definition of a mathematical expression, those `variables` need to be defined first, and can then be addressed by their ID. Finally, the change itself is specified using an instance of the `Math` class.

Table 4.2.7 on the following page shows all attributes and sub-elements for the `computeChange` element as defined by the SED-ML Level 1 Version 1 XML Schema.

4.2.7.1 The `math` element

The `math` element encodes mathematical functions. If used as an element of the `ComputeChange` class, it computes the change of the element or attribute addressed by the `target` attribute. Level 1 Version 1 supports the subset of MathML 2.0 shown in section 2.1.

Listing 40 on the next page shows the use of the `computeChange` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
target	page 24
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
listOfVariables ^o	page 27
listOfParameters ^o	page 27
math	page 39

Table 14: Attributes and nested elements for *computeChange*. *xy^o* denotes optional elements and attributes.

```

1 <model [...]>
2   <computeChange target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='w']">
3     <listOfVariables>
4       <variable id="camkii" name="active calcium/calmoduline kinase II"
5         target="/sbml:sbml/sbml:model[@id='calcium']/sbml:listOfSpecies/sbml:species[@id='KII']"
6         />
7       <variable id="w" name="synaptic weight"
8         target="/sbml:sbml/sbml:model[@id='synapse']/sbml:listOfParameters/sbml:parameter[@id='
9         w']" />
10     </listOfVariables>
11     <listOfParameters>
12       <parameter id="w0" name="synaptic weight change" value="1">
13       <parameter id="n" name="utrasensitivity to calcium" value="2">
14       <parameter id="K" name="sensitivity to calcium" value="1e-6">
15     </listOfParameters>
16     <math>
17       <apply>
18         <plus />
19         <ci>w</ci>
20         <apply>
21           <times />
22           <ci>w0</ci>
23           <apply>
24             <divide />
25             <apply>
26               <power />
27               <ci>camkii</ci>
28               <ci>n</ci>
29             </apply>
30             <plus />
31             <apply>
32               <power />
33               <ci>K</ci>
34               <ci>n</ci>
35             </apply>
36             <apply>
37               <power />
38               <ci> camkii </ci>
39               <ci>n</ci>
40             </apply>
41           </apply>
42         </apply>
43       </math>
44     </computeChange>
45   </listOfChanges>
46 </model>

```

Listing 40: The *computeChange* element

The example in listing 40 computes a change on the model parameter w . To do so, it defines two variables. `camkii` corresponds to the value of the species `KII` defined in the model with ID `calcium`. `w` corresponds to the value of the parameter `w` defined in the model with ID `synapse`. In addition, three further parameters are defined: $w_0=1$, $n=2$, and $K=1e-6$. The mathematical expression in the mathML then computes the change of w using the encoded equation:

$$w + w_0 \times \frac{camkii^n}{camkii^n + K^n} .$$

4.3 The Simulation Class

A simulation is the execution of some defined algorithm(s). Simulations are described differently depending on the type of simulation experiment to be performed (Figure 30). **Simulation** is an abstract class

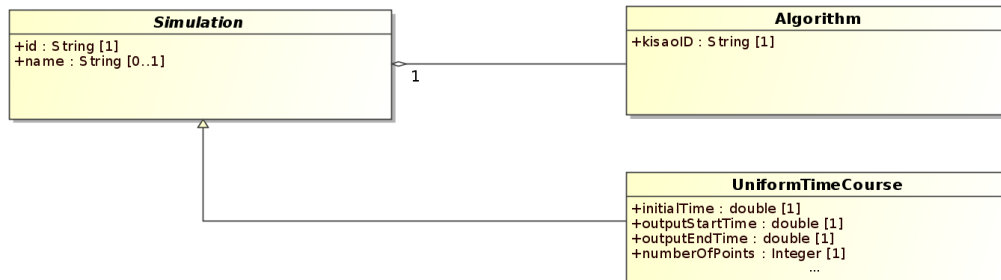


Figure 30: The SED-ML Simulation class

and serves as the container for the different types of simulation experiments. SED-ML Level 1 Version 1 offers the predefined simulation class **UniformTimeCourse**. Further simulation classes are planned for future versions of SED-ML, including simulation classes for bifurcation analysis and parameter scans. Simulation algorithms used for the execution of a simulation setup are defined in the **Algorithm** class.

Table 4.3 shows all attributes and sub-elements for the **simulation** element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
algorithm	page 42

Table 15: Attributes and nested elements for **simulation**. *xy^o* denotes optional elements and attributes.

Listing 41 on the following page shows the use of the **simulation** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. Two timcourses with uniform range are defined.

4.3.1 The UniformTimeCourse class

SED-ML Level 1 Version 1 so far only supports the encoding of uniform time course experiments.

Table 4.3.1 on page 43 shows all attributes and sub-elements for the **uniformTimeCourse** element as defined by the SED-ML Level 1 Version 1 XML Schema.

Listing 42 on the following page shows the use of the **uniformTimeCourse** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

4.3.2 The initialTime attribute

The attribute **initialTime** of type **double** represents the time from which to start the simulation. Usually this will be **0**. For an example, see listing 42.

4.3.3 The outputStartTime attribute

Sometimes a researcher is not interested in simulation results at the start of the simulation (i.e. the initial time). To accommodate this in SED-ML the **uniformTimeCourse** class uses the attribute **outputStartTime**

```

1 <listOfSimulations>
2   <uniformTimeCourse [...]>
3     [SIMULATION SPECIFICATION]
4   </uniformTimeCourse>
5   <uniformTimeCourse [...]>
6     [SIMULATION SPECIFICATION]
7   </uniformTimeCourse>
8 </listOfSimulations>

```

Listing 41: The SED-ML *listOfSimulations* element, defining two different simulations

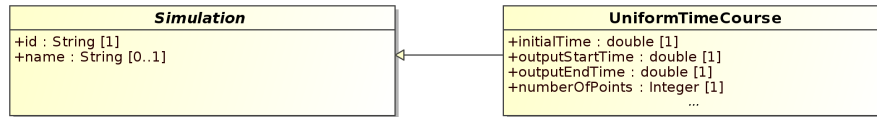


Figure 31: The *UniformTimeCourse* class

```

1 <listOfSimulations>
2   <uniformTimeCourse id="s1" name="time course simulation of variable v1 over 100 minutes"
3     initialTime="0" outputStartTime="0" outputEndTime="2500" numberOfPoints="1000">
4     <algorithm [...] />
5   </uniformTimeCourse>
6 </listOfSimulations>

```

Listing 42: The SED-ML *uniformTimeCourse* element, defining a uniform time course simulation over 2500 time units with 1000 simulation points.

of type `double`. To be valid the `outputStartTime` cannot be before `initialTime`. For an example, see listing 42.

4.3.4 The *outputEndTime* attribute

The attribute `outputEndTime` of type `double` marks the end time of the simulation. See listing 42 for an example.

4.3.5 The *numberOfPoints* attribute

When executed, the `uniformTimeCourse` simulation produces output on a regular grid starting with `outputStartTime` and ending with `outputEndTime`. The attribute `numberOfPoints` of type `integer` describes the number of points expected in the result. Software interpreting the `uniformTimeCourse` is expected to produce a first outputPoint at time `outputStartTime` with the initial values of the model to be simulated, and then `numberOfPoints` output points with the results of the simulation. Thus a total of `numberOfPoints + 1` output points will be produced.

Just because the output points lie on the regular grid described above, this does not mean that the simulation algorithm has to work with the same step size. Usually the step size the simulator chooses will be adaptive and much smaller than the required output step size. On the other hand a stochastic simulator might not have any new events occurring between two grid points. Nevertheless the simulator has to produce data on this regular grid. For an example, see listing 42.

4.3.6 The *Algorithm* class

SED-ML makes use of the [KiSAO ontology](#) (see again section 2.4 on page 13) to refer to a term in the controlled vocabulary identifying the particular simulation algorithm to be used in the simulation.

Each instance of the [Simulation](#) class must contain one reference to a simulation algorithm (Figure 32 on the next page).

Each instance of the [Algorithm](#) class must contain a [KiSAO](#) reference to a simulation algorithm. The reference should define the simulation algorithm to be used in the simulation as precisely as possible, and should be defined in the correct syntax, as defined by the regular expression `KISA0:[0-9]{7}`.

Table 4.3.6 on page 44 shows all attributes and sub-elements for the [Algorithm](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

The example given in code snippet 41, completed by algorithm definitions looks as in listing 43. For

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
initialTime	page 41
outputStartTime	page 41
outputEndTime	page 42
numberOfPoints	page 42
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
algorithm	page 42

Table 16: Attributes and nested elements for *uniformTimeCourse*. xy^o denotes optional elements and attributes.

Algorithm
+kisaoID : String [1]

Figure 32: The *Algorithm* class

```

1 <listOfSimulations>
2   <uniformTimeCourse id="s1" name="time course simulation over 100 minutes" [...]>
3     <algorithm kisaoID="KISAO:0000030" />
4   </uniformTimeCourse>
5   <uniformTimeCourse id="s2" name="time course definition for concentration of p" [...]>
6     <algorithm kisaoID="KISAO:0000021" />
7   </uniformTimeCourse>
8 </listOfSimulations>

```

Listing 43: The SED-ML *algorithm* element for two different time course simulations, defining two different algorithms. KISAO:0000030 refers to the Euler forward method ; KISAO:0000021 refers to the StochSim nearest neighbor algorithm.

both simulations, one algorithm is defined. In the first simulation **s1** a deterministic approach has been chosen (Euler forward method), in the second simulation **s2** a stochastic approach is used (Stochsim nearest neighbor).

attribute	description
metaid ^o	page 16
kisaoID	page 13
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 17: Attributes and nested elements for [algorithm](#). ^oxy^o denotes optional elements and attributes.

4.4 The Task class

A task in SED-ML links a [model](#) to a certain [simulation](#) description via their respective identifiers (Figure 33), using the [modelReference](#) and the [simulationReference](#). In SED-ML Level 1 Version 1 it

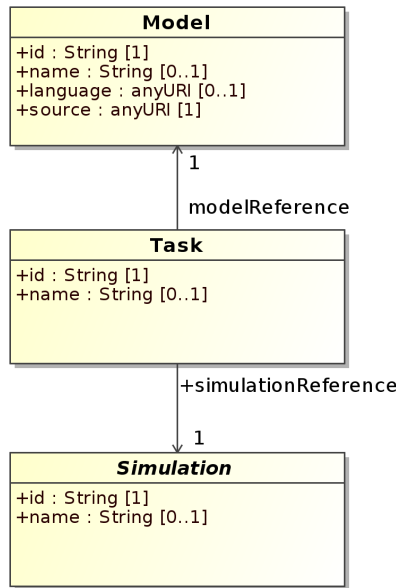


Figure 33: The SED-ML Task class

is only possible to link one simulation description to one model at a time. However, one can define as many tasks as needed within one experiment description. Please note that the tasks may be executed in any order, as XML does not have an ordering concept.

Table 4.4 on the following page shows all attributes and sub-elements for the [task](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

Listing 44 shows the use of the **task** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

```

1 <listOfTasks>
2   <task id="t1" name="task definition" modelReference="model1"
3     simulationReference="simulation 1" />
4   <task id="t2" name="another task definition" modelReference="model2"
5     simulationReference="simulation 1" />
6 </listOfTasks>

```

Listing 44: The task element

In the example, a simulation setting *simulation1* is applied first to *model1* and then is applied to *model2*.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
modelReference	page 21
simulationReference	page 22
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 18: Attributes and nested elements for [task](#). *xy^o* denotes optional elements and attributes.

4.5 The DataGenerator class

The [DataGenerator](#) class prepares the raw simulation results for later output (Figure 34). It encodes the post-processing to be applied to the simulation data. The post-processing steps could be anything, from simple normalisations of data to mathematical calculations. Each instance of the [DataGenerator](#)

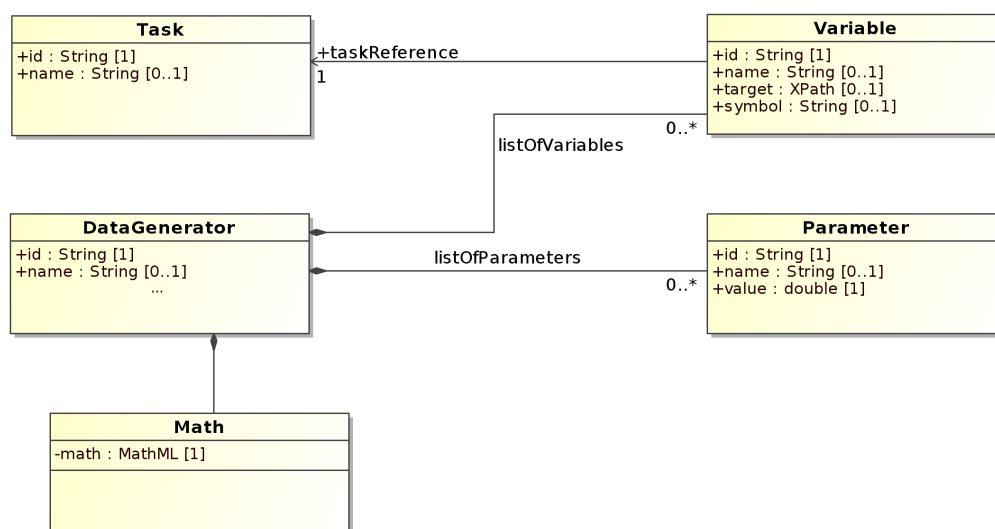


Figure 34: The SED-ML DataGenerator class

class is identifiable within the experiment by its unambiguous [id](#). It can be further characterised by an optional [name](#). The related [Math](#) class contains a `mathML` expression for the calculation of the data generator. Mathematical functions available for the specification of data generators are given in section 2.1 on [page 10](#). [Variable](#) and [Parameter](#) instances can be used to encode the mathematical expression.

Table 4.5 on the next page shows all attributes and sub-elements for the [dataGenerator](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

Listing 45 on the following page shows the use of the `dataGenerator` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The `listOfDataGenerator` contains two `dataGenerator` elements. The first one, `d1`, refers to the task definition `t1` (which itself refers to a particular model), and from the corresponding model it reuses the symbol `time`. The second one, `d2`, references a particular species defined in the same model (and referred to via the `taskReference="t1"`). The model species with ID `PX` is reused for the data generator `d2` without further post-processing.

4.6 The Output class

The [Output](#) class describes how the results of a simulation should be presented to the user (Figure 35 on [page 47](#)). It does not contain the data itself, but the type of output and the [data generators](#) used to produce a particular output.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15

sub-elements	description
math	page 39
notes ^o	page 16
annotation ^o	page 17
variable ^o	page 24
parameter ^o	page 25

Table 19: Attributes and nested elements for *dataGenerator*. *xy^o* denotes optional elements and attributes.

```

1 <listOfDataGenerators>
2   <dataGenerator id="d1" name="time">
3     <listOfVariables>
4       <variable id="time" taskReference="task1" symbol="urn:sedml:symbol:time" />
5     </listOfVariables>
6     <listOfParameters />
7     <math xmlns="http://www.w3.org/1998/Math/MathML">
8       <ci> time </ci>
9     </math>
10   </dataGenerator>
11   <dataGenerator id="d2" name="LaCI repressor">
12     <listOfVariables>
13       <variable id="v1" taskReference="task1"
14         target="/sbml:sbml/sbml:model/sbml:listOfSpecies/
15           sbml:species[@id='PX']" />
16     </listOfVariables>
17     <math:math>
18       <math:ci>v1</math:ci>
19     </math:math>
20   </dataGenerator>
21 </listOfDataGenerators>

```

Listing 45: Definition of two *dataGenerator* elements, time and LaCI repressor

The types of output pre-defined in SED-ML Level 1 Version 1 are plots and [reports](#). The output can be defined as a [2D plot](#) or alternatively as a [3D plot](#).

Note that even though the terms “2D plot” and “3D plot” are used, the exact type of plot is not specified. In other words, whether the 3D plot represents a surface plot, or three dimensional lines in space, cannot be distinguished by SED-ML alone. It is expected that applications use [annotations](#) for this purpose.

Table 4.6 shows all attributes and sub-elements for the [output](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15

sub-elements	description
notes ^o	page 16
annotation ^o	page 17
plot2D ^o	page 46
plot3D ^o	page 47
report ^o	page 48

Table 20: Attributes and nested elements for *output*. *xy^o* denotes optional elements and attributes.

4.6.1 The Plot2D class

A [2 dimensional plot](#) (Figure 36 on the next page) contains a number of [curve](#) definitions.

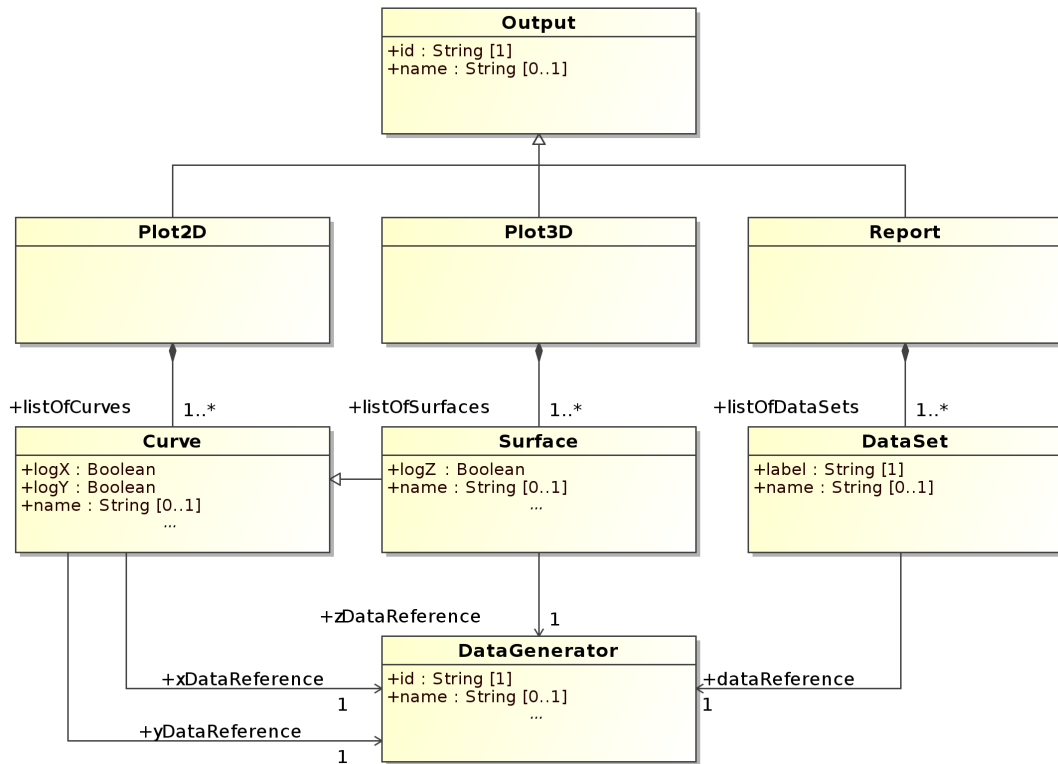


Figure 35: The SED-ML Output class



Figure 36: The SED-ML Plot2D class

Table 4.6.1 shows all attributes and sub-elements for the `plot2D` element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
notes ^o	page 16
annotation ^o	page 17
curve	page 48

Table 21: Attributes and nested elements for `plot2D`. xy^o denotes optional elements and attributes.

Listing 46 on the next page shows the use of the `listOfCurves` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The listing shows the definition of a 2 dimensional plot containing one `curve` element inside the `listOfCurves`. The curve definition follows in section 4.7 on page 48.

4.6.2 The Plot3D class

A 3 dimensional plot (Figure 37 on the following page) contains a number of `surface` definitions.

Table 4.6.2 on page 49 shows all attributes and sub-elements for the `plot3D` element as defined by the SED-ML Level 1 Version 1 XML Schema.

Listing 47 on the following page shows the use of the `plot3D` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. The example defines a `surface` for the 3 dimensional plot.

```

1 <plot2D>
2 <listOfCurves>
3 <curve>
4   [CURVE DEFINITION]
5 </curve>
6 [FURTHER CURVE DEFINITIONS]
7 </listOfCurves>
8 </plot2D>

```

Listing 46: The `plot2D` element with the nested `listOfCurves` element



Figure 37: The SED-ML `Plot3D` class

```

1 <plot3D>
2 <listOfSurfaces>
3 <surface>
4   [SURFACE DEFINITION]
5 </surface>
6 [FURTHER SURFACE DEFINITIONS]
7 </listOfSurfaces>
8 </plot3D>

```

Listing 47: The `plot3D` element with the nested `listOfSurfaces` element

The surface definition follows in section 4.8 on page 51.

4.6.3 The Report class

The `Report` class defines a data table consisting of several single instances of the `DataSet` class (Figure 38 on the next page). Its output returns the simulation result in actual *numbers*. The particular columns of the report table are defined by creating an instance of the `DataSet` class for each column.

Table 4.6.3 on page 50 shows all attributes and sub-elements for the `report` element as defined by the SED-ML Level 1 Version 1 XML Schema.

Listing 48 shows the use of the `listOfDataSets` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

```

1 <report>
2 <listOfDataSets>
3 <dataSet>
4   [DATA REFERENCE]
5 </dataSet>
6 </listOfDataSets>
7 </report>

```

Listing 48: The `report` element with the nested `listOfDataSets` element

The simulation result itself, i.e. concrete result numbers, are not stored in SED-ML, but the directive how to *calculate* them from the output of the simulator is provided through the `dataGenerator`.

The encoding of simulation results is outside the scope of SED-ML, but other efforts exist, for example the *Systems Biology Result Markup Language* (SBRML, [Dada et al., 2010]).

4.7 The Curve class

One or more instances of the `Curve` class define a 2D plot. A `curve` needs a data generator reference to refer to the data that will be plotted on the x-axis, using the `xDataReference`. A second data generator reference is needed to refer to the data that will be plotted on the y-axis, using the `yDataReference`.

Table 4.7 on page 51 shows all attributes and sub-elements for the `curve` element as defined by the

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
surface	page 51

Table 22: Attributes and nested elements for [plot3D](#). ^oxy denotes optional elements and attributes.



Figure 38: The SED-ML Report class

SED-ML Level 1 Version 1 XML Schema.

Listing 49 shows the use of the **curve** element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema. Here, only one curve is created, results shown on the x-axis are generated by

```

1 <listOfCurves>
2   <curve id="c1" name="v1 / time" xDataReference="dg1" yDataReference="dg2" logX="true" logY="false" />
3 </listOfCurves>

```

Listing 49: The SED-ML **curve** element, defining the output curve showing the result of simulation for the referenced dataGenerators

the data generator **dg1**, results shown on the y-axis are generated by the data generator **dg2**. Both **dg1** and **dg2** need to be already defined in the [listOfDataGenerators](#). The x-axis is plotted logarithmically.

4.7.1 The **logX** attribute

logX is a required attribute of the [Curve](#) class and defines whether or not the data output on the x-axis is logarithmic. The data type of **logX** is **boolean**. To make the output on the x-axis of a plot logarithmic, **logX** must be set to “true”, as shown in the sample listing 49.

logX is also used in the definition of a [Surface](#) output.

4.7.2 The **logY** attribute

logY is a required attribute of the [Curve](#) class and defines whether or not the data output on the y-axis is logarithmic. The data type of **logY** is **boolean**. To make the output on the y-axis of a plot logarithmic, **logY** must be set to “true”, as shown in the sample listing 49.

logY is also used in the definition of a [Surface](#) output.

4.7.3 The **xDataReference** attribute

The **xDataReference** is a mandatory attribute of the [Curve](#) object. Its content refers to a dataGenerator ID which denotes the [DataGenerator](#) object that is used to generate the output on the x-axis of a [Curve](#) in a [2D Plot](#). The **xDataReference** data type is **string**. However, the valid values for the **xDataReference** are restricted to the IDs of already defined [DataGenerator](#) objects.

An example for the definition of a curve is given in listing 49. **xDataReference** is also used in the definition of the x-axis of a [Surface](#) in a [3D Plot](#).

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
sub-elements	description
notes ^o	page 16
annotation ^o	page 17
dataSet	page 53

Table 23: Attributes and nested elements for [report](#). xy^o denotes optional elements and attributes.

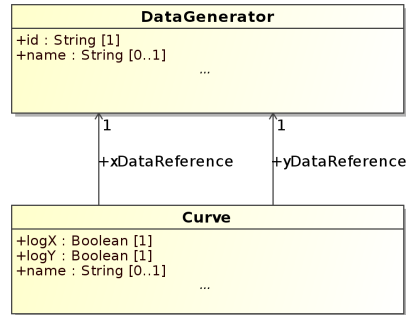


Figure 39: The SED-ML Curve class

4.7.4 The yDataReference attribute

The [yDataReference](#) is a mandatory attribute of the [Curve](#) object. Its content refers to a [dataGenerator](#) ID which denotes the [DataGenerator](#) object that is used to generate the output on the y-axis of a [Curve](#) in a [2D Plot](#). The [yDataReference](#) data type is **string**. However, the number of valid values for the [yDataReference](#) is restricted to the IDs of already defined [DataGenerator](#) objects.

An example for the definition of a curve is given in listing 49. [yDataReference](#) is also used in the definition of the y-axis of a [Surface](#) in a [3D Plot](#).

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
logX	page 49
xDataReference	page 49
logY	page 49
yDataReference	page 50
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 24: Attributes and nested elements for [curve](#). ^o denotes optional elements and attributes.

4.8 The Surface class

A [surface](#) is a three-dimensional figure representing a simulation result (Figure 40).

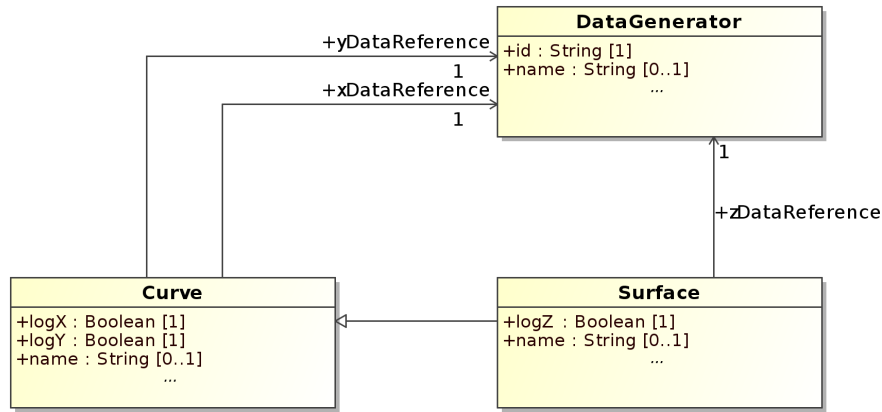


Figure 40: The SED-ML Surface class

Creating an instance of the [Surface](#) class demands the definition of three different axes, that is which data to plot on which axis and in which way. The aforementioned `xDataReference` and `yDataReference` attributes define the according [data generators](#) for both the x- and y-axis of a surface. In addition, the `zDataReference` attribute defines the output for the z-axis. All axes might be logarithmic or not. This can be specified through the `logX`, `logY`, and the `logZ` attributes in the according `dataReference` elements.

Table 4.8 on the next page shows all attributes and sub-elements for the [surface](#) element as defined by the SED-ML Level 1 Version 1 XML Schema. Listing 50 shows the use of the `surface` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

```

1 <listOfSurfaces>
2   <surface id="s1" name="surface" xDataReference="dg1" yDataReference="dg2" zDataReference="dg3"
3     logX="true" logY="false" logZ="false" />
4   [FURTHER SURFACE DEFINITIONS]
5 </listOfSurfaces>

```

Listing 50: The SED-ML surface element, defining the output showing the result of the referenced task

Here, only one surface is created, results shown on the x-axis are generated by the data generator `dg1`, results shown on the y-axis are generated by the data generator `dg2`, and results shown on the z-axis are generated by the data generator `dg3`. All `dg1`, `dg2` and `dg3` need to be already defined in the [listOfDataGenerators](#).

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
logX	page 49
xDataReference	page 49
logY	page 49
yDataReference	page 50
logZ	page 52
zDataReference	page 52
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 25: Attributes and nested elements for [surface](#). *xy^o* denotes optional elements and attributes.

4.8.1 The logZ attribute

[logZ](#) is a required attribute of the [Surface](#) class and defines whether or not the data output on the z-axis is logarithmic. The data type of [logZ](#) is **boolean**. To make the output on the z-axis of a surface plot logarithmic, [logZ](#) must be set to “true”, as shown in the sample listing [50](#).

4.8.2 The zDataReference attribute

The [zDataReference](#) is a mandatory attribute of the [Surface](#) object. Its content refers to a dataGenerator ID which denotes the [DataGenerator](#) object that is used to generate the output on the z-axis of a [3D Plot](#). The [zDataReference](#) data type is **string**. However, the valid values for the [zDataReference](#) are restricted to the IDs of already defined [DataGenerator](#) objects.

An example using the [zDataReference](#) is given in listing [50](#) on [page 51](#).

4.9 The DataSet class

The [DataSet](#) class holds definitions of data to be used in the [Report](#) class (Figure 41). Data sets are

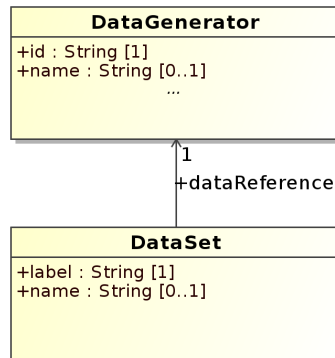


Figure 41: *The SED-ML DataSet class*

labeled references to instances of the [DataGenerator](#) class.

Table 4.9 shows all attributes and sub-elements for the [dataSet](#) element as defined by the SED-ML Level 1 Version 1 XML Schema.

attribute	description
metaid ^o	page 16
id	page 15
name ^o	page 15
dataReference	page 53
label	page 53
sub-elements	description
notes ^o	page 16
annotation ^o	page 17

Table 26: *Attributes and nested elements for [dataSet](#). xy^o denotes optional elements and attributes.*

4.9.1 The label attribute

Each data set in a [Report](#) does have to carry an unambiguous [label](#). A label is a human readable descriptor of a data set for use in a [report](#). For example, for a tabular data set of time series results, the label could be the column heading.

4.9.2 The dataReference attribute

The [dataReference](#) attribute contains the ID of a [dataGenerator](#) element and as such represents a link to it. The data produced by that particular data generator fills the according data set in the [report](#).

Listing 51 shows the use of the `dataSet` element in a SED-ML file as defined by the SED-ML Level 1 Version 1 XML Schema.

```
1 <listOfDataSets>
2   <dataSet id="d1" name="v1 over time" dataReference="dg1" label="_1">
3 </listOfDataSets>
```

Listing 51: *The SED-ML `dataSet` element, defining a data set containing the result of the referenced task*

Acknowledgements

The SED-ML specification has been developed with the input of many people. Main contributors of the current specification include Richard Adams, Frank Bergmann, Stefan Hoops, Nicolas Le Novère, Ion Moraru, Sven Sahle, Henning Schmidt and Dagmar Waltemath.

Moreover, we would like to thank all the participants of the meetings where SED-ML has been discussed as well as the subscribers of the sed-ml-discuss mailing list.

A SED-ML UML Overview

Figure 42 shows the complete UML diagram of the SED-ML. It gives the full picture of all implemented classes (see the XML Schema definition in 62).

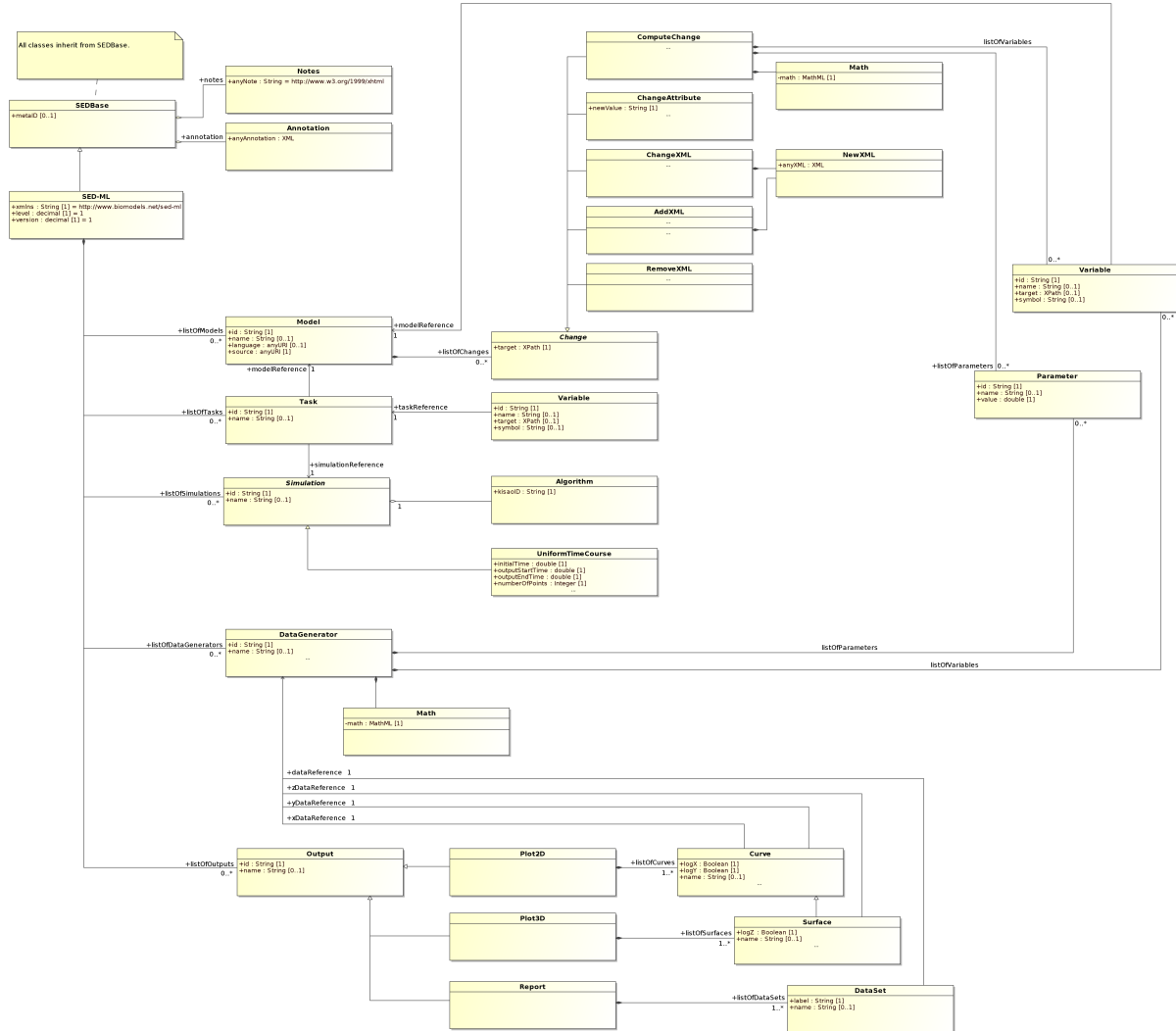


Figure 42: The SED-ML UML class diagram

B Overview of SED-ML

The *Simulation Experiment Description Markup Language* (SED-ML) is an XML-based format for the description of simulation experiments. It serves to store information about the simulation experiment performed on one or more models with a given set of outputs. Support for SED-ML compliant simulation descriptions will enable the exchange of simulation experiments across tools.

B.1 Conventions

The Business Process Modeling Notation Version 1.2 (BPMN) was initially intended to describe internal business procedures (processes) in a graphical way. However, we will use BPMN to graphically describe the steps and processes of setting up a simulation experiment description. The major parts of BPMN that are used to specify SED-ML are activities, gateways, events, data, and documentation.

An *activity* is “work that is performed on a [...] process”, for example “Specify the simulation settings”. Activities may be atomic or non-atomic. SED-ML in particular makes use of the *task* activities, i.e., specific work units that need to be performed. Non-atomic tasks might be collapsed or expanded in the graphical representation (see Figure 43). Each collapsed subprocess has a corresponding expanded subprocess definition.

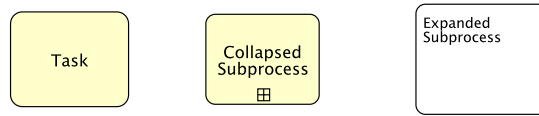


Figure 43: BPMN activities: task, collapsed process, expanded subprocess

Gateways serve as means to control the flow of sequence in the diagram. As the term already implies, a gateway needs some “mechanism that either allows or disallows passage through” [White et al., 2004]. The result of a gateway pass-through can be that processes are merged or splitted. Graphically, a gateway is represented as a diamond.

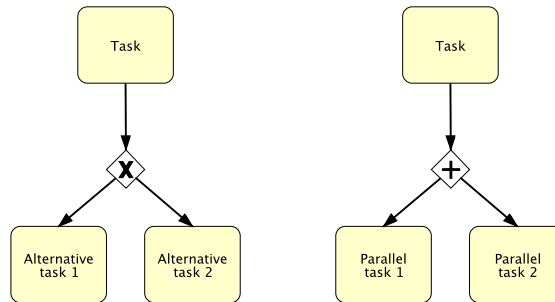


Figure 44: BPML gateway types: Exclusive (left), parallel (right)

While there exist a number of different gateway types (see [White et al., 2004], pp. 93), the SED-ML specification only uses the parallel and the exclusive gates (see Figure 44).

Exclusive gateways – also denoted as decisions – allow the sequence flow to take two or more alternative paths (Figure 44, left hand side). However, *only one* of the paths may be chosen (not more). Sometimes two alternative branches need to be merged together again, in which case the exclusive gate must be used as well: The sequence flow continues as soon as *one* of the incoming processes send a signal. An exclusive gateways is marked by an **X** in the graphical notation.

Parallel gateways, “provide a mechanism to synchronize parallel flow and to create parallel flow” [White et al., 2004] (Figure 44, right hand side). They are used to show parallel paths in the workflow; even if

sometimes not required they might help in understanding the process. Synchronisation allows to start two processes in parallel at the same time in the sequence flow: The sequence flow will continue with *all* processes leaving the parallel gateway. Joining two processes with a parallel gateway is also possible: the process flow will only continue after a signal has arrived from *all* processes coming in the parallel gateway. A parallel gateway is marked by a + in the graphical notation.

Events mark everything happening during the execution of the sequence flow, usually they interrupt the business process, having some cause or impact on the execution. From the broad range of events that BPMN offers, SED-ML only uses a small subset, namely the start event and the end event (Figure 45).

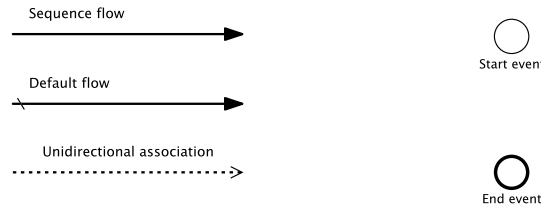


Figure 45: BPMN connectors (left) and events (right).

All events are graphically drawn as small circles. A *start event* is drawn with a single thin line and mark the start of a process, it can not have any incoming sequence flow. Start events may be triggered by different mechanisms, for the case of SED-ML the untyped start event (no marker inside the circle) is used. The trigger to start the process is “Create new simulation experiment”. The *end event* is marked with a thick line. It indicates the end of a process. SED-ML specification makes use of the untyped end event (no marker inside the circle). The end event is used to show the end of sub-processes as well as processes. If the end of a sub-process is reached, the sequence flow returns to the according parent process.

Connectors are used to combine different BPMN objects with each other ([White et al., 2004] page 30 shows the full list of valid connections). SED-ML uses only a subset of available connectors, namely sequence flow, default flow, and unidirectional associations (Figure 45). *Sequence flow* defines the execution order of activities. *Default flow* marks the default branch to be chosen if other conditions leave various possibilities for further execution of the sequence flow. A *unidirectional association* is used to indicate that a data object is modified, i. e. read and written during the execution of an activity [Business Process Technology group, 2009].

The rough SED-ML workflow is shown in Figure 46. The process of defining a SED-ML simulation

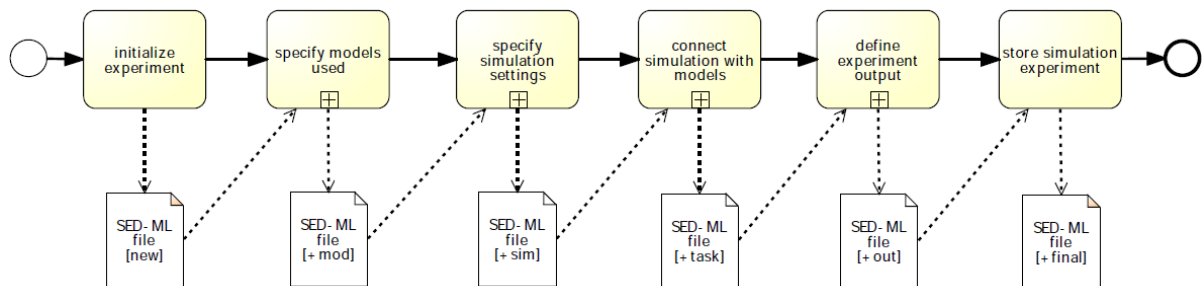


Figure 46: The process of defining a simulation experiment in SED-ML (overview)

experiment starts by initialising the experiment and creating a new sed-ml file. Afterwards, the [models](#) needed for the simulation are specified and stored into the existing sed-ml file (see section B.1.1). In a third step, the simulation experiment [setups](#) are defined and stored into the same file (see section B.1.2). To assign a setup to a number of models used in the experiment, these connections have to be defined and recorded (see section B.1.3), called [task](#) in SED-ML. After simulation, the [output](#) should defined,

based on the specified tasks and performed simulation experiment. The information is added to the existing SED-ML file (see section B.1.4). In the end, the whole experiment is stored in the final SED-ML file. All collapsed processes are described in the following. Examples in XML are provided in the more technical description.

B.1.1 Models

To define a simulation experiment, first of all a new SED-ML file is created. The models to be used in the experiment (zero or many) are referenced, using a link to a model description in some open, curated model base (such as Biomed Models Database [Li et al. \[2010\]](#), CellML Repository [Beard et al. \[2009\]](#), or alike). Changes that are necessary to simulate the model correctly are defined, e.g. assigning new parameter values or updating the mathematics of the model (Figure 47). The procedure is repeated until all models

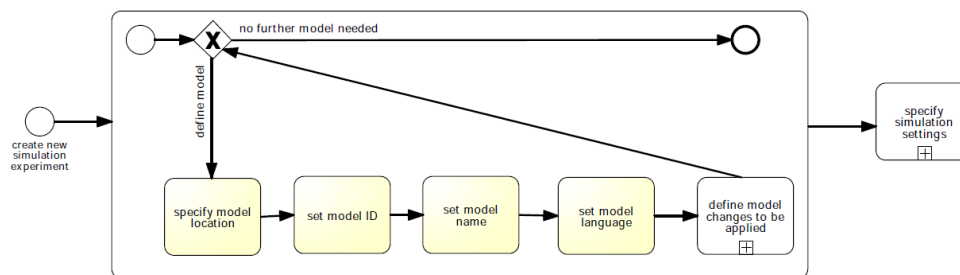


Figure 47: The process of defining model(s) in SED-ML

participating in the experiment have been described. Each model used gets an internal SED-ML ID and an optional name.

B.1.2 Simulation setup

Secondly, the simulation setups (zero or many) used throughout the simulation experiment are described (Figure 48). Those may stem from various different types of simulation, e.g. steady state analysis or

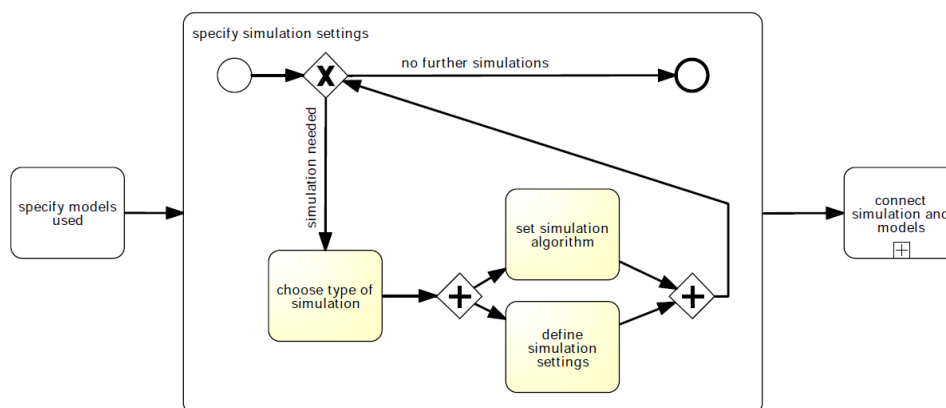


Figure 48: The process of defining simulation(s) in SED-ML

bifurcation. Depending on the specific type of experiment, the information encoded for the simulation setup might differ. Thus, the definition of simulation settings is specific to the simulation experiment.

In a simple case the experiment consists of one simulation, but it can get far more complex. For example, one might define a nested sequence of simulations, in which case every simulation has to be defined separately. Each simulation setup gets its own internal ID and an optional name. For each of the setups, the simulation algorithm to be used for that simulation is defined through a reference to a well-defined algorithm name, e.g. an ontology or controlled vocabulary. One approach to define such a controlled

vocabulary of simulation algorithms is the *Kinetic Simulation Algorithm Ontology* (KiSAO, [Courtot et al. \[2011\]](#)). The setup definition is repeated until all different simulations have been described.

B.1.3 Task

SED-ML allows to apply one defined simulation setting to one defined model at a time. However, any number of [tasks](#) may be defined inside a simulation experiment description (Figure 49). To do so, each

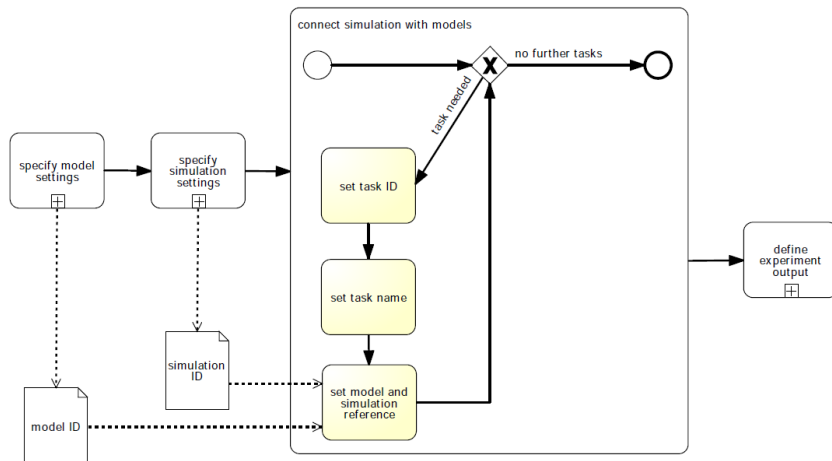


Figure 49: The process of defining simulation task(s) in SED-ML

task refers to one of the formerly specified models and to one of the formerly specified simulation setups. Each task has its own ID and an optional name. The process of task definition is repeated until all tasks have been defined.

The current SED-ML does not allow to nest or order tasks. However, these features are evaluated for future versions of SED-ML.

B.1.4 Output

The SED-ML finally consists of output definitions that describe what kind of output the experiment uses to present the simulation result to the user, i.e. a plot or a data table (Figure 50), and also which data is part of the output. Therefore, SED-ML first defines a set of [data generators](#) (Figure 51), which are

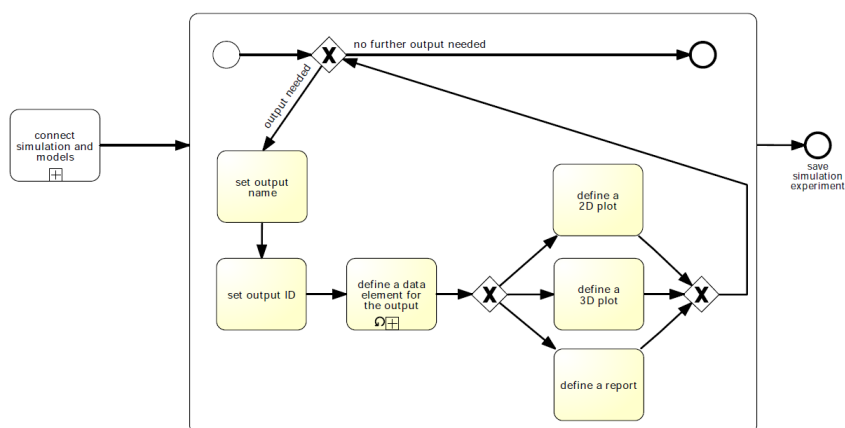


Figure 50: The process of defining output(s) in SED-ML

then used to specify a particular result, i.e. output (see section [B.1.5](#)).

The SED-ML specification comes with three pre-defined types of outputs: 2D- and 3D plots, and reports. All use the aforementioned data generators to specify the information to be plotted on the different axes, or in the table columns respectively.

B.1.5 Data Generator

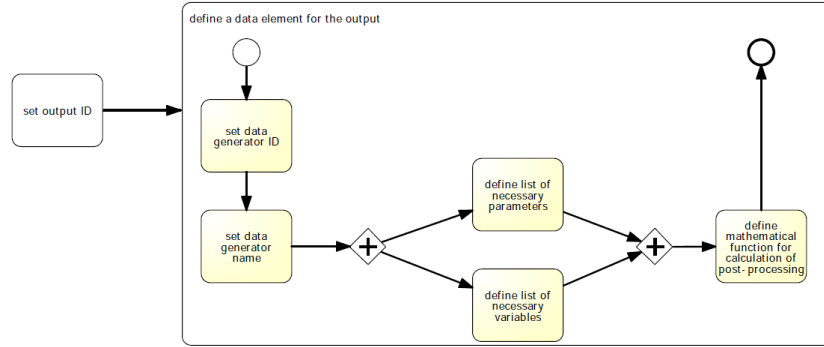


Figure 51: *The process of defining data generator(s) in SED-ML*

A data generator may use data elements, e. g. variables or parameters, that either (1) have been taken directly from the model, or (2) have been generated in a post-processing step. If post-processing needs to be applied, variables and parameters from the various, previously defined models may be used, but also existing global parameters, such as *time*. If the variables are taken from existing models, a reference to the model and the particular variable needs to be given. If post-processing is necessary, a reference to an existing variable or parameter, including other data generators, has to be provided. Additional mathematical rules to be applied on the referred variable or parameter must then be specified. In a SED-ML file, any number of data generators can be created for later re-use in the output definition.

C XML Schema

Listing 52 shows the full SED-ML XML Schema. The code is commented inline.

```
1 <xs:schema targetNamespace="http://www.biomodels.net/sed-ml"
2   xmlns="http://www.biomodels.net/sed-ml" xmlns:xs="http://www.w3.org/2001/XMLSchema"
3   xmlns:math="http://www.w3.org/1998/Math/MathML">
4   <xs:import namespace="http://www.w3.org/1998/Math/MathML"
5     schemaLocation="sbml-mathml.xsd" />
6
7
8   <xs:simpleType name="SId">
9     <xs:annotation>
10       <xs:documentation>
11         The type SId is used throughout SED-ML as the
12         type of the 'id' attributes on model elements.
13       </xs:documentation>
14     </xs:annotation>
15     <xs:restriction base="xs:string">
16       <xs:pattern value="(_|[a-z]|[A-Z])(_|[a-z]|[A-Z]|[0-9])*" />
17     </xs:restriction>
18   </xs:simpleType>
19
20   <!-- attribute group for elements with ID/name att -->
21   <xs:attributeGroup name="idGroup">
22     <xs:attribute name="id" use="required" type="SId"></xs:attribute>
23     <xs:attribute name="name" use="optional" type="xs:string"></xs:attribute>
24   </xs:attributeGroup>
25
26
27   <!-- SED Base class -->
28   <xs:complexType name="SEDBase">
29     <xs:annotation>
30       <xs:documentation xml:lang="en">
31         The SEDBase type is the
32         base type of all main types in SED-ML. It
33         serves as a container for
34         the annotation of any part of the
35         experiment description.
36       </xs:documentation>
37     </xs:annotation>
38     <xs:sequence>
39       <xs:element ref="notes" minOccurs="0" />
40       <xs:element ref="annotation" minOccurs="0" />
41     </xs:sequence>
42   </xs:complexType>
43   <!--
44     This must be a variable-type identifier ,i.e., (Letter | '_' )
45     (NCNameChar)* that is unique in the document.
46   -->
47   <xs:attribute name="metaid" type="xs:ID" use="optional"></xs:attribute>
48 </xs:complexType>
49 <xs:element name="sedML">
50   <xs:complexType>
51     <xs:complexContent>
52       <xs:extension base="SEDBase">
53         <xs:sequence>
54           <xs:element ref="listOfSimulations" minOccurs="0" />
55           <xs:element ref="listOfModels" minOccurs="0" />
56           <xs:element ref="listOfTasks" minOccurs="0" />
57           <xs:element ref="listOfDataGenerators" minOccurs="0" />
58           <xs:element ref="listOfOutputs" minOccurs="0" />
59         </xs:sequence>
60         <xs:attribute name="level" type="xs:decimal" use="required"
61           fixed="1" />
62         <xs:attribute name="version" type="xs:decimal" use="required"
63           fixed="1" />
64       </xs:extension>
65     </xs:complexContent>
66   </xs:complexType>
67 </xs:element>
68 <!-- notes and annotations -->
69 <xs:element name="notes">
70   <xs:complexType>
71     <xs:sequence>
72       <xs:any namespace="http://www.w3.org/1999/xhtml"
73         processContents="skip" minOccurs="0" maxOccurs="unbounded" />
74     </xs:sequence>
75   </xs:complexType>
76 </xs:element>
77 <xs:element name="annotation">
78   <xs:complexType>
79     <xs:sequence>
80       <xs:any processContents="skip" minOccurs="0" maxOccurs="unbounded" />
81     </xs:sequence>
82   </xs:complexType>
83 </xs:element>
84 <!-- KISA ID type -->
```

```

85 <xs:simpleType name="KisaoType">
86   <xs:restriction base="xs:string">
87     <xs:pattern value="KISAO:[0-9][0-9][0-9][0-9][0-9][0-9]" />
88   </xs:restriction>
89 </xs:simpleType>
90
91 <!-- global element declarations -->
92 <xs:element name="variable">
93   <xs:complexType>
94     <xs:complexContent>
95       <xs:extension base="SEDBase">
96         <!-- at least one of taskReference or modelReference must be set -->
97         <xs:attribute name="taskReference" type="SId"
98           use="optional" />
99         <xs:attribute name="modelReference" type="SId"
100           use="optional" />
101
102         <!--
103           either target or symbol have to be used in the variable definition
104         -->
105         <xs:attribute name="target" type="xs:token" use="optional" />
106         <xs:attribute name="symbol" type="xs:string" use="optional" />
107         <xs:attributeGroup ref="idGroup" />
108       </xs:extension>
109     </xs:complexContent>
110   </xs:complexType>
111 </xs:element>
112 <xs:element name="parameter">
113   <xs:complexType>
114     <xs:complexContent>
115       <xs:extension base="SEDBase">
116         <xs:attributeGroup ref="idGroup" />
117         <xs:attribute name="value" type="xs:double" use="required" />
118       </xs:extension>
119     </xs:complexContent>
120   </xs:complexType>
121 </xs:element>
122 <xs:element name="algorithm">
123   <xs:complexType>
124     <xs:complexContent>
125       <xs:extension base="SEDBase">
126         <xs:attribute name="kisaoID" type="KisaoType" use="required" />
127       </xs:extension>
128     </xs:complexContent>
129   </xs:complexType>
130 </xs:element>
131 <xs:element name="uniformTimeCourse">
132   <xs:complexType>
133     <xs:complexContent>
134       <xs:extension base="SEDBase">
135         <xs:sequence>
136           <xs:element ref="algorithm" />
137         </xs:sequence>
138         <xs:attributeGroup ref="idGroup" />
139         <xs:attribute name="outputStartTime" type="xs:double"
140           use="required" />
141         <xs:attribute name="outputEndTime" type="xs:double"
142           use="required" />
143         <xs:attribute name="numberOfPoints" type="xs:integer"
144           use="required" />
145         <xs:attribute name="initialTime" type="xs:double" use="required" />
146       </xs:extension>
147     </xs:complexContent>
148   </xs:complexType>
149 </xs:element>
150 <xs:element name="task">
151   <xs:complexType>
152     <xs:complexContent>
153       <xs:extension base="SEDBase">
154         <xs:attribute name="simulationReference" type="SId"
155           use="required" />
156
157         <xs:attribute name="modelReference" type="SId"
158           use="required" />
159         <xs:attributeGroup ref="idGroup" />
160       </xs:extension>
161     </xs:complexContent>
162   </xs:complexType>
163 </xs:element>
164 <xs:element name="plot2D">
165   <xs:complexType>
166     <xs:complexContent>
167       <xs:extension base="SEDBase">
168         <xs:sequence>
169           <xs:element ref="listOfCurves" minOccurs="0" />
170         </xs:sequence>
171         <xs:attributeGroup ref="idGroup" />
172       </xs:extension>
173     </xs:complexContent>

```

```

174     </xs:complexType>
175 </xs:element>
176 <xs:element name="plot3D">
177   <xs:complexType>
178     <xs:complexContent>
179       <xs:extension base="SEDBase">
180         <xs:sequence>
181           <xs:element ref="listOfSurfaces" minOccurs="0" />
182         </xs:sequence>
183         <xs:attributeGroup ref="idGroup" />
184       </xs:extension>
185     </xs:complexContent>
186   </xs:complexType>
187 </xs:element>
188 <xs:element name="report">
189   <xs:complexType>
190     <xs:complexContent>
191       <xs:extension base="SEDBase">
192         <xs:sequence>
193           <xs:element ref="listOfDataSets" minOccurs="0" />
194         </xs:sequence>
195         <xs:attributeGroup ref="idGroup" />
196       </xs:extension>
197     </xs:complexContent>
198   </xs:complexType>
199 </xs:element>
200 <xs:element name="model">
201   <xs:complexType>
202     <xs:complexContent>
203       <xs:extension base="SEDBase">
204         <xs:sequence>
205           <xs:element ref="listOfChanges" minOccurs="0" />
206         </xs:sequence>
207         <xs:attribute name="language" type="xs:anyURI" use="optional"
208           default="urn:sedml:language:xml" />
209         <xs:attribute name="source" type="xs:anyURI" use="required" />
210         <xs:attributeGroup ref="idGroup" />
211       </xs:extension>
212     </xs:complexContent>
213   </xs:complexType>
214 </xs:element>
215 <!-- math element, does not inherit from SEDBase -->
216 <xs:element name="math" type="math:Math" />
217 <!-- listOf elements -->
218 <xs:element name="listOfVariables">
219   <xs:complexType>
220     <xs:complexContent>
221       <xs:extension base="SEDBase">
222         <xs:sequence>
223           <xs:element ref="variable" minOccurs="0" maxOccurs="unbounded" />
224         </xs:sequence>
225       </xs:extension>
226     </xs:complexContent>
227   </xs:complexType>
228 </xs:element>
229 <xs:element name="listOfParameters">
230   <xs:complexType>
231     <xs:complexContent>
232       <xs:extension base="SEDBase">
233         <xs:sequence>
234           <xs:element ref="parameter" minOccurs="0" maxOccurs="unbounded" />
235         </xs:sequence>
236       </xs:extension>
237     </xs:complexContent>
238   </xs:complexType>
239 </xs:element>
240 <xs:element name="listOfTasks">
241   <xs:complexType>
242     <xs:complexContent>
243       <xs:extension base="SEDBase">
244         <xs:sequence>
245           <xs:element ref="task" minOccurs="0" maxOccurs="unbounded" />
246         </xs:sequence>
247       </xs:extension>
248     </xs:complexContent>
249   </xs:complexType>
250 </xs:element>
251 <xs:element name="listOfSimulations">
252   <xs:complexType>
253     <xs:complexContent>
254       <xs:extension base="SEDBase">
255         <xs:sequence>
256           <xs:element ref="uniformTimeCourse" minOccurs="0"
257             maxOccurs="unbounded" />
258         </xs:sequence>
259       </xs:extension>
260     </xs:complexContent>
261   </xs:complexType>
262 </xs:element>

```



```

263 <xs:element name="listOfOutputs">
264   <xs:complexType>
265     <xs:complexContent>
266       <xs:extension base="SEDBase">
267         <xs:sequence minOccurs="0">
268           <xs:element ref="plot2D" minOccurs="0" maxOccurs="unbounded" />
269           <xs:element ref="plot3D" minOccurs="0" maxOccurs="unbounded" />
270           <xs:element ref="report" minOccurs="0" maxOccurs="unbounded" />
271         </xs:sequence>
272       </xs:extension>
273     </xs:complexContent>
274   </xs:complexType>
275 </xs:element>
276 <xs:element name="listOfModels">
277   <xs:complexType>
278     <xs:complexContent>
279       <xs:extension base="SEDBase">
280         <xs:sequence>
281           <xs:element ref="model" minOccurs="0" maxOccurs="unbounded" />
282         </xs:sequence>
283       </xs:extension>
284     </xs:complexContent>
285   </xs:complexType>
286 </xs:element>
287 <xs:element name="listOfDataGenerators">
288   <xs:complexType>
289     <xs:complexContent>
290       <xs:extension base="SEDBase">
291         <xs:sequence>
292           <xs:element ref="dataGenerator" minOccurs="0"
293             maxOccurs="unbounded" />
294         </xs:sequence>
295       </xs:extension>
296     </xs:complexContent>
297   </xs:complexType>
298 </xs:element>
299 <xs:element name="listOfCurves">
300   <xs:complexType>
301     <xs:complexContent>
302       <xs:extension base="SEDBase">
303         <xs:sequence>
304           <xs:element ref="curve" maxOccurs="unbounded" />
305         </xs:sequence>
306       </xs:extension>
307     </xs:complexContent>
308   </xs:complexType>
309 </xs:element>
310 <xs:element name="listOfSurfaces">
311   <xs:complexType>
312     <xs:complexContent>
313       <xs:extension base="SEDBase">
314         <xs:sequence>
315           <xs:element ref="surface" maxOccurs="unbounded" />
316         </xs:sequence>
317       </xs:extension>
318     </xs:complexContent>
319   </xs:complexType>
320 </xs:element>
321 <xs:element name="listOfDataSets">
322   <xs:complexType>
323     <xs:complexContent>
324       <xs:extension base="SEDBase">
325         <xs:sequence>
326           <xs:element ref="dataSet" maxOccurs="unbounded" />
327         </xs:sequence>
328       </xs:extension>
329     </xs:complexContent>
330   </xs:complexType>
331 </xs:element>
332 <!-- change -->
333 <xs:element name="listOfChanges">
334   <xs:complexType>
335     <xs:complexContent>
336       <xs:extension base="SEDBase">
337         <xs:sequence>
338           <xs:element ref="changeAttribute" minOccurs="0"
339             maxOccurs="unbounded" />
340           <xs:element ref="changeXML" minOccurs="0" maxOccurs="unbounded" />
341           <xs:element ref="addXML" minOccurs="0" maxOccurs="unbounded" />
342           <xs:element ref="removeXML" minOccurs="0" maxOccurs="unbounded" />
343           <xs:element ref="computeChange" minOccurs="0"
344             maxOccurs="unbounded" />
345         </xs:sequence>
346       </xs:extension>
347     </xs:complexContent>
348   </xs:complexType>
349 </xs:element>
350 <xs:element name="newXML">
351   <xs:complexType>

```

```

352         <xs:sequence>
353             <xs:any processContents="skip" minOccurs="1" maxOccurs="unbounded" />
354         </xs:sequence>
355     </xs:complexType>
356 </xs:element>
357 <xs:element name="changeAttribute">
358     <xs:complexType>
359         <xs:complexContent>
360             <xs:extension base="SEDBase">
361                 <xs:attribute name="target" use="required" type="xs:token" />
362                 <xs:attribute name="newValue" type="xs:string" use="required" />
363             </xs:extension>
364         </xs:complexContent>
365     </xs:complexType>
366 </xs:element>
367 <xs:element name="changeXML">
368     <xs:complexType>
369         <xs:complexContent>
370             <xs:extension base="SEDBase">
371                 <xs:sequence>
372                     <xs:element ref="newXML" />
373                 </xs:sequence>
374                 <xs:attribute name="target" use="required" type="xs:token" />
375             </xs:extension>
376         </xs:complexContent>
377     </xs:complexType>
378 </xs:element>
379 <xs:element name="addXML">
380     <xs:complexType>
381         <xs:complexContent>
382             <xs:extension base="SEDBase">
383                 <xs:sequence>
384                     <xs:element ref="newXML" />
385                 </xs:sequence>
386                 <xs:attribute name="target" use="required" type="xs:token" />
387             </xs:extension>
388         </xs:complexContent>
389     </xs:complexType>
390 </xs:element>
391 <xs:element name="removeXML">
392     <xs:complexType>
393         <xs:complexContent>
394             <xs:extension base="SEDBase">
395                 <xs:attribute name="target" use="required" type="xs:token" />
396             </xs:extension>
397         </xs:complexContent>
398     </xs:complexType>
399 </xs:element>
400 <xs:element name="computeChange">
401     <xs:complexType>
402         <xs:complexContent>
403             <xs:extension base="SEDBase">
404                 <xs:sequence>
405                     <xs:element ref="listOfVariables" minOccurs="0" />
406                     <xs:element ref="listOfParameters" minOccurs="0" />
407                     <xs:element ref="math" />
408                 </xs:sequence>
409                 <xs:attribute name="target" use="required" type="xs:token" />
410             </xs:extension>
411         </xs:complexContent>
412     </xs:complexType>
413 </xs:element>
414 <!-- data generator -->
415 <xs:element name="dataGenerator">
416     <xs:complexType>
417         <xs:complexContent>
418             <xs:extension base="SEDBase">
419                 <xs:sequence>
420                     <xs:element ref="listOfVariables" minOccurs="0" />
421                     <xs:element ref="listOfParameters" minOccurs="0" />
422                     <xs:element ref="math" />
423                 </xs:sequence>
424                 <xs:attributeGroup ref="idGroup" />
425             </xs:extension>
426         </xs:complexContent>
427     </xs:complexType>
428 </xs:element>
429 <xs:element name="curve">
430     <xs:complexType>
431         <xs:complexContent>
432             <xs:extension base="SEDBase">
433                 <xs:attributeGroup ref="idGroup" />
434                 <xs:attribute name="yDataReference" type="SId"
435                     use="required" />
436                 <xs:attribute name="xDataReference" type="SId"
437                     use="required" />
438             </xs:extension>
439             <xs:attribute name="logY" use="required" type="xs:boolean" />
440             <xs:attribute name="logX" use="required" type="xs:boolean" />

```

```

441         </xs:extension>
442     </xs:complexContent>
443 </xs:complexType>
444 </xs:element>
445 <xs:element name="surface">
446     <xs:complexType>
447         <xs:complexContent>
448             <xs:extension base="SEDBase">
449                 <xs:attributeGroup ref="idGroup" />
450                 <xs:attribute name="yDataReference" type="SId"
451                     use="required" />
452                 <xs:attribute name="xDataReference" type="SId"
453                     use="required" />
454                 <xs:attribute name="zDataReference" type="SId"
455                     use="required" />
456                 <xs:attribute name="logY" use="required" type="xs:boolean" />
457                 <xs:attribute name="logX" use="required" type="xs:boolean" />
458                 <xs:attribute name="logZ" use="required" type="xs:boolean" />
459             </xs:extension>
460         </xs:complexContent>
461     </xs:complexType>
462 </xs:element>
463 <xs:element name="dataSet">
464     <xs:complexType>
465         <xs:complexContent>
466             <xs:extension base="SEDBase">
467                 <xs:attribute name="dataReference" type="SId"
468                     use="required"></xs:attribute>
469                 <xs:attribute name="label" use="required" type="xs:string" />
470                 <xs:attributeGroup ref="idGroup" />
471             </xs:extension>
472         </xs:complexContent>
473     </xs:complexType>
474 </xs:element>
475 </xs:schema>

```

Listing 52: *The SED-ML XML Schema definition*

D Examples

D.1 Le Loup Model (CellML)

The following example provides a SED-ML description for the simulation of the model based on the publication by Leoupy, Gonze and Goldbeter “Limit Cycle Models for Circadian Rhythms Based on Transcriptional Regulation in Drosophila and Neurospora” (PubMed ID: 10643740). The model source code is taken from the CellML Model Repository [Lloyd et al., 2008].

The original model used in the simulation experiment is referred to using a URL (http://models.cellml.org/workspace/leloupy_gonze_goldbeter_1999/@rawfile/d6613d7e1051b3eff2bb1d3d419a445bb8c754ad/leloupy_gonze_goldbeter_1999_b.cellml, ll. 15-16). In order to set up the model some pre-processing needs to be applied: Those are defined in the `listOfChanges` from ll. 17-25. All changes defined update particular parameter values in the model.

A second model is defined in l. 28 of the example, using `model1` as a source and applying even further changes to it, in this case updating two more model parameters.

One simulation setup is defined in the `listOfSimulations`. It is a `uniformTimeCourse` over 180 time units, using 1000 simulation points. The algorithm used is the CVODE solver, as denoted by the KiSAO ID `KiSAO:0000019`.

A number of `dataGenerators` are defined in ll. 42-92. Those are the prerequisite for defining the output of the simulation. The first `dataGenerator` named `tim1` in l. 45 maps on the `Mt` entity in the model that is used in `task1` which here is the model with ID `model1`. The second `dataGenerator` named `per-tim` in l. 57 maps on the `CN` entity in `model1`. Finally the third and fourth `dataGenerators` map on the `Mt` and `per-tim` entity respectively in the updated model with ID `model2`.

The `output` defined in the experiment consists of a 2D plot with two different curves (ll. 96-102). Both curves plot the `per-tim` concentration against the `tim` concentration. In the first curve the original parametrisation (as given in `model1`) is used, in the second curve the updated one is used (as given in `model2`).

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <sedML version="0.1" xmlns="http://www.biomodels.net/sed-ml"
3   xmlns:math="http://www.w3.org/1998/Math/MathML">
4   <notes><p xmlns="http://www.w3.org/1999/xhtml">Comparing Limit Cycles and strange attractors for
5     oscillation in Drosophila</p></notes>
6   <listOfSimulations>
7     <uniformTimeCourse id="simulation1" algorithm="KiSAO:0000019"
8       initialTime="0" outputStartTime="0" outputEndTime="180"
9       numberOfPoints="1000" >
10       <algorithm kisaoID="KiSAO:0000019"/>
11     </uniformTimeCourse>
12   </listOfSimulations>
13   <listOfModels>
14     <model id="model1" name="Circadian Oscillations" language="urn:sedml:language:cellml" source="http://
15       models.cellml.org/workspace/leloupy_gonze_goldbeter_1999/@rawfile/
16       d6613d7e1051b3eff2bb1d3d419a445bb8c754ad/leloupy_gonze_goldbeter_1999_a.cellml" >
17       <listOfChanges>
18         <changeAttribute target="/cellml:model/cellml:component[@cmeta:id='MP']/cellml:variable[@name='vsP']/
19           @initial_value" newValue="1"/>
20         <changeAttribute target="/cellml:model/cellml:component[@cmeta:id='MP']/cellml:variable[@name='vmP']/
21           @initial_value" newValue="0.7"/>
22         <changeAttribute target="/cellml:model/cellml:component[@cmeta:id='P2']/cellml:variable[@name='vdP']/
23           @initial_value" newValue="2"/>
24         <changeAttribute target="/cellml:model/cellml:component[@cmeta:id='T2']/cellml:variable[@name='vdT']/
25           @initial_value" newValue="2"/>
26         <changeAttribute target="/cellml:model/cellml:component[@name='parameters']/cellml:variable[@name='k1
27           ']/@initial_value" newValue="0.6"/>
28         <changeAttribute target="/cellml:model/cellml:component[@name='parameters']/cellml:variable[@name='
29           K4P']/@initial_value" newValue="1"/>
30         <changeAttribute target="/cellml:model/cellml:component[@name='parameters']/cellml:variable[@name='
31           K4T']/@initial_value" newValue="1"/>
32       </listOfChanges>
33     </model>
34     <model id="model2" name="Circadian Chaos" language="urn:sedml:language:cellml" source="model1">
35       <listOfChanges>
36         <changeAttribute target="/cellml:model/cellml:component[@cmeta:id='MT']/cellml:variable[@name='vmT']/
37           @initial_value" newValue="0.28"/>
38         <changeAttribute target="/cellml:model/cellml:component[@cmeta:id='T2']/cellml:variable[@name='vdT']/
39           @initial_value" newValue="4.8"/>
40       </listOfChanges>
41     </model>
42   </listOfModels>
43   <listOfTasks>
44     <task id="task1" name="Limit Cycle" modelReference="model1" simulationReference="simulation1"/>
45   </listOfTasks>
46 </sedML>
```

```

35     <task id="task2" name="Strange attractors" modelReference="model2" simulationReference="simulation1"/
36     >
37 </listOfTasks>
38 <listOfDataGenerators>
39   <dataGenerator id="tim1" name="tim mRNA">
40     <listOfVariables>
41       <variable id="v0" taskReference="task1" target="/cellml:model/cellml:component[@cmeta:id='MT']" /
42       >
43     </listOfVariables>
44     <math:math>
45       <math:apply>
46         <math:plus />
47         <math:ci>v0</math:ci>
48       </math:apply>
49     </math:math>
50   </dataGenerator>
51   <dataGenerator id="per_tim" name="nuclear PER-TIM complex">
52     <listOfVariables>
53       <variable id="v1" taskReference="task1" target="/cellml:model/cellml:component[@cmeta:id='CN']" /
54       >
55     </listOfVariables>
56     <math:math>
57       <math:apply>
58         <math:plus />
59         <math:ci>v1</math:ci>
60       </math:apply>
61     </math:math>
62   </dataGenerator>
63   <dataGenerator id="tim2" name="tim mRNA (changed parameters)">
64     <listOfVariables>
65       <variable id="v2" taskReference="task2" target="/cellml:model/cellml:component[@cmeta:id='MT']" /
66       >
67     </listOfVariables>
68     <math:math>
69       <math:apply>
70         <math:plus />
71         <math:ci>v2</math:ci>
72       </math:apply>
73     </math:math>
74   </dataGenerator>
75   <dataGenerator id="per_tim2" name="nuclear PER-TIM complex">
76     <listOfVariables>
77       <variable id="v3" taskReference="task2" target="/cellml:model/cellml:component[@cmeta:id='CN']" /
78       >
79     </listOfVariables>
80     <math:math>
81       <math:apply>
82         <math:plus />
83         <math:ci>v3</math:ci>
84       </math:apply>
85     </math:math>
86   </dataGenerator>
87 </listOfDataGenerators>
88 <listOfOutputs>
89   <plot2D id="plot1" name="tim mRNA with Oscillation and Chaos">
90     <listOfCurves>
91       <curve id="c1" logX="false" logY="false" xDataReference="per_tim" yDataReference="tim1" />
92       <curve id="c2" logX="false" logY="false" xDataReference="per_tim2" yDataReference="tim2" />
93     </listOfCurves>
94   </plot2D>
95 </listOfOutputs>
96 </sedML>

```

Listing 53: *LeLoup Model Simulation Description in SED-ML*

D.2 The IkappaB-NF-kappaB signaling module (SBML)

The following example provides a SED-ML description for the simulation of the IkappaB-NF-kappaB signaling module based on the publication by Hoffmann, Levchenko, Scott and Baltimore “The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. ” (PubMed ID: 12424381)

This model is referenced by its SED-ML ID **model1** and refers to the model with the MIRIAM URN [urn:miriam:biomodels.db:BIOMD0000000140](http://miriam.org/urn:miriam:biomodels.db:BIOMD0000000140). Software applications interpreting this example know how to dereference this URN and access the model in BioModels Database [Le Novère et al., 2006].

The simulation description specifies one simulation **simulation1**, which is a uniform timecourse simulation that simulates the model for 41 hours. **task1** then applies this simulation to the model.

As output this simulation description collects four parameters: **Total_NFkBn**, **Total_IkBbeta**, **Total_IkBeps** and **Total_IkBalpha**. These variables are to be plotted against the simulation time and displayed in four separate plots, as shown in Figure 52.

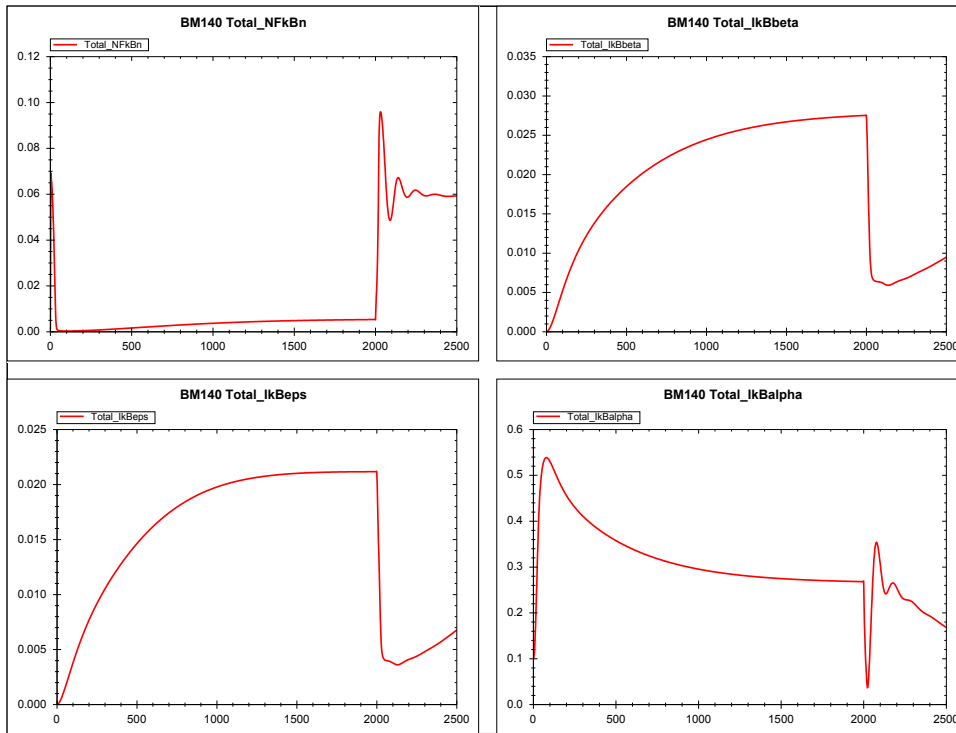


Figure 52: The simulation result gained from the simulation description given in listing 54

The SED-ML description of the simulation experiment is given in listing 54.

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <sedML xmlns="http://www.biomodels.net/sed-ml">
3   <listOfSimulations>
4     <uniformTimeCourse id="simulation1"
5       initialTime="0" outputStartTime="0" outputEndTime="2500"
6       numberOfPoints="1000" >
7       <algorithm kisaoID="KISAO:0000019"/>
8     </uniformTimeCourse>
9   </listOfSimulations>
10  <listOfModels>
11    <model id="model1" language="urn:sedml:language:sbml" source="urn:miriam:biomodels.db:BIOMD0000000140"
12      />
13  </listOfModels>
14  <listOfTasks>
15    <task id="task1" modelReference="model1"
16      simulationReference="simulation1"/>
17  </listOfTasks>
18  <listOfDataGenerators>
19    <dataGenerator id="time" name="time">
20      <listOfVariables>
21        <variable id="time1" taskReference="task1" symbol="urn:sedml:symbol:time"/>
22      </listOfVariables>
23    </dataGenerator>
24  </listOfDataGenerators>
25 </sedML>

```

```

23     <ci>time1</ci>
24   </math>
25 </dataGenerator>
26 <dataGenerator id="Total_NFkBn" name="Total_NFkBn">
27   <listOfVariables>
28     <variable id="Total_NFkBn1" taskReference="task1"
29       target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='Total_NFkBn']"/>
30   </listOfVariables>
31   <math xmlns="http://www.w3.org/1998/Math/MathML">
32     <ci>Total_NFkBn1</ci>
33   </math>
34 </dataGenerator>
35 <dataGenerator id="Total_IkBbeta" name="Total_IkBbeta">
36   <listOfVariables>
37     <variable id="Total_IkBbeta1" taskReference="task1"
38       target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='Total_IkBbeta']"/>
39   </listOfVariables>
40   <math xmlns="http://www.w3.org/1998/Math/MathML">
41     <ci>Total_IkBbeta1</ci>
42   </math>
43 </dataGenerator>
44 <dataGenerator id="Total_IkBeps" name="Total_IkBeps">
45   <listOfVariables>
46     <variable id="Total_IkBeps1" taskReference="task1"
47       target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='Total_IkBeps']"/>
48   </listOfVariables>
49   <math xmlns="http://www.w3.org/1998/Math/MathML">
50     <ci>Total_IkBeps1</ci>
51   </math>
52 </dataGenerator>
53 <dataGenerator id="Total_IkBalpha" name="Total_IkBalpha">
54   <listOfVariables>
55     <variable id="Total_IkBalpha1" taskReference="task1"
56       target="/sbml:sbml/sbml:model/sbml:listOfParameters/sbml:parameter[@id='Total_IkBalpha']"/>
57   </listOfVariables>
58   <math xmlns="http://www.w3.org/1998/Math/MathML">
59     <ci>Total_IkBalpha1</ci>
60   </math>
61 </dataGenerator>
62 </listOfDataGenerators>
63 <listOfOutputs>
64   <plot2D id="plot1" name="BM140 Total_NFkBn">
65     <listOfCurves>
66       <curve id="c1" logX="false" logY="false" xDataReference="time"
67         yDataReference="Total_NFkBn"/>
68     </listOfCurves>
69   </plot2D>
70   <plot2D id="plot2" name="BM140 Total_IkBbeta">
71     <listOfCurves>
72       <curve id="c2" logX="false" logY="false" xDataReference="time"
73         yDataReference="Total_IkBbeta"/>
74     </listOfCurves>
75   </plot2D>
76   <plot2D id="plot3" name="BM140 Total_IkBeps">
77     <listOfCurves>
78       <curve id="c3" logX="false" logY="false" xDataReference="time"
79         yDataReference="Total_IkBeps"/>
80     </listOfCurves>
81   </plot2D>
82   <plot2D id="plot4" name="BM140 Total_IkBalpha">
83     <listOfCurves>
84       <curve id="c4" logX="false" logY="false" xDataReference="time"
85         yDataReference="Total_IkBalpha"/>
86     </listOfCurves>
87   </plot2D>
88 </listOfOutputs>
89 </sedML>

```

Listing 54: *IkappaB-NF-kappaB signaling Model Simulation Description in SED-ML*

References

- D. A. Beard, R. Britten, M. T. Cooling, A. Garny, M. D. Halstead, P. J. Hunter, J. Lawson, C. M. Lloyd, J. Marsh, A. Miller, D. P. Nickerson, P. M. Nielsen, T. Nomura, S. Subramaniam, S. M. Wimalaratne, and T. Yu. CellML metadata standards, associated tools and repositories. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367(1895):1845–1867, May 2009. ISSN 1364-503X. doi: 10.1098/rsta.2008.0310.
- D. Bell. UML basics, Part III: The class diagram. IBM, the rational edge, 2003. http://download.boulder.ibm.com/ibmdl/pub/software/dw/rationaledge/nov03/t_modelinguml_db.pdf.
- T. Berners-Lee, R. Fielding, and L. Masinter. Uniform Resource Identifier (URI): Generic Syntax, 2005. URL <http://www.ietf.org/rfc/rfc3986.txt>.
- T. Bray, J. Paoli, C.M. Sperberg-McQueen, E. Maler, F. Yergeau, and J. Cowan. Extensible Markup Language (XML) 1.1 (Second Edition), 2006. URL <http://www.w3.org/TR/xml11/>.
- Business Process Technology group. BPMN – business process modeling notation. poster, 2009. URL <http://bpt.hpi.uni-potsdam.de/Public/BPMNCorner>.
- D. Carlisle, P. Ion, R. Miner, and N. Poppelier. Mathematical Markup Language (MathML) version 2.0. *W3C Recommendation*, 21, 2001.
- J. Clarke and S. DeRose. XML Path Language (XPath) Version 1.0, 1999. URL <http://www.w3.org/TR/xpath/>.
- M. Courtot et al. Controlled vocabularies and semantics in systems biology. *in preparation*, 2011.
- J.O. Dada, I. Spasić, N.W. Paton, and P. Mendes. SBRML: a markup language for associating systems biology data with models. *Bioinformatics (Oxford, England)*, 26(7):932–938, April 2010. ISSN 1367-4811. doi: 10.1093/bioinformatics/btq069.
- M.B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, January 2000.
- D.C. Fallside, P. Walmsley, et al. XML schema part 0: Primer. *W3C recommendation*, 2, 2001.
- N. Goddard, M. Hucka, F. Howell, H. Cornelis, K. Skankar, and D. Beeman. Towards NeuroML: Model Description Methods for Collaborative Modeling in Neuroscience. *Phil. Trans. Royal Society series B*, 356:1209–1228, 2001.
- S. Hoops, S. Sahle, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer. COPASI - a COmplex PATHway Simulator. *Bioinformatics (Oxford, England)*, 22(24):3067–3074, December 2006. ISSN 1460-2059. doi: 10.1093/bioinformatics/btl485.
- M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J. H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, March 2003. ISSN 1367-4803. doi: 10.1093/bioinformatics/btg015.
- M. Hucka, F.T. Bergmann, S. Hoops, S. Keating, S. Sahle, and D.J. Wilkinson. The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core (Release 1 Candidate). *Nature Precedings*, January 2010. ISSN 1756-0357. doi: 10.1038/npre.2010.4123.1.
- N. Le Novère, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J. L. Snoep, and M. Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*, 34(Database issue), January 2006. ISSN 1362-4962.

- C. Li, M. Donizelli, N. Rodriguez, H. Dharuri, L. Endler, V. Chelliah, L. Li, E. He, A. Henry, M. Stefan, J. Snoep, M. Hucka, N. Le Novère, and C. Laibe. BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4(1):92+, June 2010. ISSN 1752-0509. doi: 10.1186/1752-0509-4-92.
- C.M. Lloyd, M.D.B. Halstead, and P.F. Nielsen. CellML: its future, present and past. *Prog Biophys Mol Biol*, 85:433–450, 2004.
- C.M. Lloyd, J.R. Lawson, P.J. Hunter, and P.F. Nielsen. The CellML model repository. *Bioinformatics*, 24(18):2122, 2008.
- OMG. *UML 2.2 Superstructure and Infrastructure*, February 2009. URL <http://www.omg.org/spec/UML/2.2/>.
- S. Pemberton et al. XHTML 1.0: The Extensible HyperText Markup Language—W3C Recommendation 26 January 2000. *World Wide Web Consortium (W3C)(August 2002)*, 2002.
- W3C. XML Schema Part 1: Structures Second Edition. W3C Recommendation, October 2004. URL <http://www.w3.org/TR/xmlschema-1/>.
- D. Waltemath, R. Adams, D.A. Beard, F.T. Bergmann, U.S. Bhalla, R. Britten, V. Chelliah, M.T. Cooling, J. Cooper, E. Crampin, A. Garny, S. Hoops, M. Hucka, P. Hunter, E. Klipp, C. Laibe, A. Miller, I. Moraru, D. Nickerson, P. Nielsen, M. Nikolski, S. Sahle, H. Sauro, H. Schmidt, J.L. Snoep, D. Tolle, O. Wolkenhauer, and N. Le Novère. Minimum information about a simulation experiment (MIASE). *submitted*, 2011.
- S.A. White et al. Business process modeling notation (BPMN) version 1.0. *Business Process Management Initiative, BPMI. org*, 2004.