Simulation Modelling Practice and Theory 68 (2016) 33–53

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Reusing simulation experiment specifications to support developing models by successive extension

Danhua Peng[a,*], Tom Warnke[a], Fiete Haack[a,b], Adelinde M. Uhrmacher[a]

[a] Institute of Computer Science, University of Rostock, Rostock 18059, Germany
[b] Research Unit "Functional Genome Analysis", Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf 18196, Germany

ARTICLE INFO

Article history:
Received 5 March 2016
Revised 25 July 2016
Accepted 28 July 2016

Keywords:
Model extension
Stochastic modeling
Simulation experiments
Experiment specification
Experiment generation and execution

ABSTRACT

Model development is a successive process of validating, revising, and extending models, and requires iterative execution of simulation experiments. While developing a model by extension, executing similar simulation experiments to those performed with the original model reveals important behavioral insights into the extended model. An automatic generation and execution of these simulation experiments can provide valuable support in the process of developing models. A prerequisite is an explicit specification of simulation experiments. Therefore, we annotate models with simulation experiments that are specified in a declarative domain specific language SESSL (Simulation Experiment Specification via a Scala Layer). Based on experiment specifications of the original model, we introduce a mechanism to automatically generate and execute simulation experiments for the extended model with necessary adaptations. Furthermore, as we experiment with stochastic models, we exploit statistical model checking and specify the expected model behavioral properties, against which the simulation results are checked. Thereby, when a model is extended, the original experiment specifications are reused, adapted, and applied to the extended model. Accordingly, the generated simulation trajectories are probed to check whether the expected properties hold with a certain probability or not. Thus, more fast and frequent feedback during model development can be provided to the modeler. Based on a model of membrane related dynamics, we show how the developed approach can be used in successively extending models.

1. Introduction

Developing models presents itself as an intricate process and requires iterative experimentation, validation and refinement, as described by different life-cycles, e.g., in [1]. Often, models are generated incrementally. As soon as one part of the model satisfies certain requirements (after iteratively executing experiments and revising the model), new parts can be added. When a model is extended, it would be interesting to know whether the extended model still behaves the same way as before, and which "properties" of its behavior remain unchanged and which do not. Simulation experiments can provide key insights into model behavior. Therefore, to analyze the impact that the extension has on the model behavior, similar simulation experiments to those that have been conducted with the original model are conducted with the extended model.

* Corresponding author.
E-mail addresses: danhua.peng2@uni-rostock.de, pengdanhua26@gmail.com (D. Peng), tom.warnke@uni-rostock.de (T. Warnke), fiete.haack@uni-rostock.de (F. Haack), adelinde.uhrmacher@uni-rostock.de (A.M. Uhrmacher).

http://dx.doi.org/10.1016/j.simpat.2016.07.006

However, re-conducting prior simulation experiments on the extended model, especially by hand, is non-trivial and error-prone. First of all, the information of each simulation experiment conducted on the original model needs to be recorded, along with corresponding simulation results, for example in document files or spreadsheets. Based on this information, simulation experiments are manually designed and executed for the extended model. For each experiment configuration, the results from different model versions are analyzed and compared. The situation is aggravated when models are successively extended and many simulation experiments are conducted for each model version during this process. Therefore, in this paper, we present an approach to support the automatic generation and execution of experiments upon the extended model by reusing the information about simulation experiments of the original model. Our goal is to provide fast and frequent feedback during model development to the modeler.

Our work is based on explicit, unambiguous experiment specifications for executing simulation experiments. To interweave modeling and experimentation, models are annotated with specifications of simulation experiments that have been executed, so that those experiment specifications can be reused in revising and extending models. At the moment a model is extended, the original simulation specifications are reused, adapted, and applied to the extended model. As we focus on stochastic models and to make the properties to be checked explicit, we make use of well-established statistical model checking techniques to analyze simulation results, i.e., formalizing the interesting model behavior properties and checking simulation results against those properties.

Even though the individual aspects of our approach, i.e., explicit specification of simulation experiments (e.g., see [2]) and statistical model checking technique (e.g., see [3]), are not new, the contribution of our approach lies in how they are brought together to provide assistance in the process of model development by successively revising, extending, and validating models.

In the following, we first discuss different aspects in reusing simulation experiments, i.e., specifying simulation experiments, specifying properties of model behavior and specifying the data processing method. Next, we present the method for checking properties. Afterwards, we illustrate the automatic experiment generation for the extended model based on experiment specifications of the original model. For demonstration, we applied our approach to the development of receptor dynamics models. Finally, we discuss related work and the usability of our approach, and summarize the results.

2. Requirements for reusing simulation experiments

The essential idea of our approach is to reuse simulation experiments. To facilitate the reuse of simulation experiments, an explicit, unambiguous description of experiments is needed. For that, the information required has to be identified.

2.1. Structuring information about simulation experiments

An unambiguous and complete description of simulation experiments is of crucial importance for their reproducibility and reuse. Many efforts have been dedicated to define standards for describing simulation experiments, e.g., Minimum Information About a Simulation Experiment (MIASE) [4] and Minimum Simulation Reporting Requirements (MSRR) [5]. Following those guidelines, we distinguish four important aspects for specifying simulation experiments: model configuration, simulation configuration, data processing method and model behavioral property.

Model configuration defines how the model is used in the simulation experiment, including its location, its initial state, and configuration of model parameters. *Simulation configuration* describes the set-up of the simulation experiment (e.g., simulation algorithm and stopping rules). *Data processing method* describes the method used to process the output data of the simulation experiment, such as a smoothing method. *Model behavioral property* defines the requirements a model's behavior has to satisfy, which are reflected in the generated experiment output. For example, in a cell model, one property could be that the concentration of a certain protein should reach a steady state after a certain simulation time.

2.2. Specifying simulation experiments

Numerous approaches exist to support the specification of experiments. For instance, the Integrated Modeling Support Environment (IMSE) Experimenter, which is a graphical tool to support experimentation with performance models [6], and the Experiment Schema Extension (Ex-SE) of a framework [7], which combines performance model interchange formats and experiment specifications to execute and analyze performance experiments, are developed to facilitate the description, execution, and documentation of experiments based on formal models. In addition, some other work exploits general scientific workflow systems, such as Taverna [8], to specify simulation experiments, e.g., [9].

In the last decade, a series of *domain specific languages* (DSL) have been developed to express different aspects of a simulation experiment [10]. A DSL is a language specialized to a particular problem that "speaks" the language of the domain [11].

The Simulation Experiment Description Markup Language (SED-ML) [12] is an XML-based language to encode and document simulation experiment information required by MIASE, to facilitate exchange and reproduction of experiments in systems biology. It allows describing most frequent types of simulation experiments and is independent of concrete simulation systems. However, it only supports models described in XML and encodes the description of simulation experiments in XML, which makes it more machine-readable than human-readable, and therefore requires additional tools (e.g., SED-ED)