MASARYKOVA UNIVERZITA

Annex No. 11 to the MU Directive on Habilitation Procedures and Professor Appointment Procedures

Habilitation Thesis Reviewer's Report

Masaryk University	
Faculty	Faculty of Informatics
Procedure field	Informatics
Applicant	RNDr. David Šafránek, Ph.D.
Applicant's home unit, institution	Faculty informatics, Masaryk University
Habilitation thesis	Formal Methods for Analysis of Biological Systems under Parameter Uncertainty
Reviewer	Dr. François Fages
Reviewer's home unit, institution	Inria Saclay - Île-de-France research centre, France

[Review text] See letter enclosed

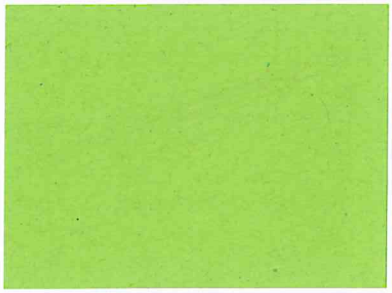
Reviewer's questions for the habilitation thesis defence (number of questions up to the reviewer)

... See questions at the end of my letter

Conclusion

The habilitation thesis entitled "*Formal Methods for Analysis of Biological Systems under Parameter Uncertainty*" by David Šafránek ***fulfils*** ~~***does not fulfil***~~ requirements expected of a habilitation thesis in the field of Informatics.

In Palaiseau on May 5th 2019



on May 5th 2019

from François Fages,
Inria Saclay IdF, Palaiseau, France
to the Habilitation Committee
of the Faculty of Informatics
of Mazaryk University, Czech Rep.

Object Review on the Habilitation Thesis of Dr. David SAFRANEK

David SAFRANEK presents an Habilitation Thesis on the development of Formal Methods for the Analysis of Biological Systems under Parameter Uncertainty. This Thesis belongs to a multidisciplinary research field named Computational Systems Biology which gathers computer scientists, mathematicians, biologists, chemists, physicists, on the understanding of the high-level functions of the cell from their molecular basis at the level of molecular interaction networks. This research trend appeared in the late 90's after the end of the Human Genome Project to launch a similar research effort on post-genomic data and the elucidation of information processing and biochemical processes in the cell.

The Thesis is representative of the application of concepts and tools form fundamental Computer Science to Biology. The Thesis extends previous work on the formal verification of temporal logic properties in biochemical reaction networks in two main directions: by considering the stochastic interpretation of reaction networks by Continuous Time Markov Chains (CTMC) instead of their continuous interpretation by Ordinary Differential Equations, and by considering the case where kinetic parameters are defined by intervals instead of precise values. These extensions complete the corpus of computational methods to tackle useful systems biology questions. The technical difficulties associated to these advances show the perfect mastering of this research field by David Safranek. Furthermore, most of the presented results have been implemented with colleagues in software available on the web showing the seriousness of the work.

The first chapter is an introduction to Systems Biology and some lessons learned by the author from his participation in European projects in this domain.

The second chapter presents the necessary background on, on the one hand, the formalisms used for modelling biochemical systems, i.e. chemical reaction networks (CRN) and influence networks with their different interpretations, Boolean, discrete, stochastic and continuous, and on the other hand, the temporal logics originated from Computer Science for modelling the behaviours of the systems, i.e. the cells, together with the associated model-checking tools.

The third chapter introduces the three main technical contributions of the habilitation thesis with respect to the state-of-the-art, concerning the problems of

- parameter synthesis with a symbolic representation of the sets of solutions,
- parameter exploration for probabilistic model-checking with an approximation of the

CENTRE DE RECHERCHE
SACLAY-ÎLE-DE-FRANCE
1 rue Honoré d'Estienne d'Orves
Bâtiment Alan Turing
Campus de l'École Polytechnique
91120 Palaiseau Cedex France
Tél. : +33 (0)1 72 92 59 00
Fax : +33 (0)1 74 85 42 42

landscape of the probability of satisfaction of temporal logic properties for each parameter sets,

- and the robustness analysis of formal properties in Signal Temporal Logic STL and probabilistic temporal logic CSL for stochastic CRN.

The key concepts underlying these contributions are the ones of coloured model-checking and parameterized uniformisation presented in the following two chapters.

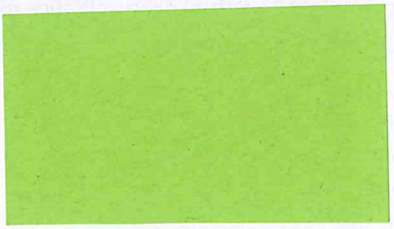
In Chapter 4, the concept of Coloured model-checking for dealing with a symbolic representation of sets (colours) of parameter values is presented, together with algorithms for computing those sets for both temporal logics LTL and CTL, and with some short mention of the performance of the prototype implementations. The complexity of these algorithms comes from the computation of solution sets for satisfying a formula in a bottom-up fashion for all sub-formulae and for all states.

In Chapter 5, the method of parameterized uniformisation is presented to approximate the minimum and maximum functions of the formula satisfaction landscape in a continuous CRN. This method is somehow reminiscent to interval arithmetic and constraint solvers such as RealPaver to approximate non-linear functions or the behaviour of non-linear dynamical systems, but here it is used to approximate the landscape probability of satisfaction of bounded time CSL properties. This has been implemented on top of PRISM, the reference CSL model-checker.

Chapter 6 shows how these results can be used to measure the robustness of temporal logic properties in both continuous and stochastic CRN in respectively Signal Temporal Logic STL and Continuous Stochastic Logic CSL.

Chapter 7 presents the application of those results to the analysis of several small size yet representative biological systems. Unfortunately, this evaluation has not been done for comparing the different interpretations in a systematic way. For instance, the Lotka-Volterra example used to illustrate the deterministic continuous interpretation in LTL (oscillations), could be interestingly used to show the very different behaviour in the stochastic interpretation in CSL (almost sure extinction). Such a comparison is done for the cell cycle model of G1/S transition to illustrate the computation overhead of the stochastic interpretation by one or two orders of magnitude.

In conclusion, David Safranek has shown in his Habilitation Thesis the excellent skills of a researcher in informatics, and his ability to contribute at the cutting edge of international research in computational methods for Systems Biology. His scientific achievements certainly fulfill the requirements expected of a habilitation thesis in informatics.



François Fages
Research Director
Resp. Project-team Lifeware
http://lifeware.inria.fr/~fages
Inria Saclay – Ile de France

Inria

François Fages's questions for the habilitation thesis defence

1. In the motivation of your thesis you write that the structure of the models is usually available and that the main question is the estimation of parameter values. However, in my experience, the most interesting contributions of parameter synthesis methods to biological knowledge happened when the methods failed to fit the data. Indeed, this lead us precisely to revisit the structure of the network, for instance by considering a reaction that was neglected and that revealed essential to fit some mutant data, or by hypothesizing a different mechanism of interaction to explain data that could not be fitted with the initial structure of the model. In those situations, it is an empty set of parameter values that was the most informative lesson of the model. Have you encountered a similar situation?
2. You cite the work of Sicun Gao and Edmund Clarke on SAT modulo ODEs and delta-reachability. They combine reachability analysis (based on SAT solver) with interval arithmetic (based on RealPaver constraint solver) to propagate reachability constraints in hybrid systems. Is there any link between these techniques and your computation of min-max approximation of the landscape function for reachability properties?
3. When written with reactions, the prey-predator model of Lotka and Volterra can be interpreted in both the continuous semantics by ODEs and the stochastic semantics by CTMC. In the first interpretation, the model exhibits sustained oscillations, while in the second it leads almost surely to the extinction of the predator (and the prey for most parameter settings). Can it be shown with your tool based on PRISM and parameter uniformization?