Applying bisimulation and invariants to alternate pathways in a signalling cascade

Vashti Galpin
Laboratory for Foundations of Computer Science
School of Informatics, University of Edinburgh
Vashti.Galpin@ed.ac.uk

This presentation will describe how weak g-bisimulation where invariants are used to determine g, can be used to show that two alternative pathways in the MAPK signalling cascade have the same behaviour and hence are pairwise redundant [4].

Weak g-bisimulation is defined for the stochastic process algebra Bio-PEPA [1]. It applies specifically to Bio-PEPA with levels where the amount of each species is quantified as an natural number which represents a range of concentrations or molecule counts defined by a fixed step size.

It involves a function g which is defined over the labels of the capability relation and the target and source models of a transition in this capability relation, $P \xrightarrow{\theta}_c P'$. Each label θ has the form (α, w) where $\alpha \in \mathscr{A}$ a set of reaction names and the set w is defined by the grammar $w ::= [S : \ \mathtt{op}\ l, k] \mid w :: w$. Here S is a species, op is the prefix operator which indicates the role of the species in the reaction, $l \in \mathbb{N}$ is the level of S before the reaction and $k \in \mathbb{N}^+$ is the stoichiometry constant of S in reaction α. The term $w :: w$ represent list concatenation; however, since ordering is not important and for well-defined Bio-PEPA models, there can be no duplicates, these lists are sets and hence this operation can be viewed as set union [4].

The function g maps to an arbitrary set $X \cup \{\tau\}$ where τ is a distinguished action which indicates a reaction from which we wish to abstract. New transitions are defined as follows.

- $P \xrightarrow{g(\theta_1,P,P_1)\ldots g(\theta_n,P_{n-1},P')}_g P'$ for $P \xrightarrow{\theta_1}_c P_1 \xrightarrow{\theta_2}_c \ldots \xrightarrow{\theta_{n-1}}_c P_{n-1} \xrightarrow{\theta_n}_c P'$.
- $P \xrightarrow{\phi_1 \ldots \phi_n}_g P'$ for $P\,(\xrightarrow{\tau}_g)^* \xrightarrow{\phi_1}_g (\xrightarrow{\tau}_g)^* \ldots (\xrightarrow{\tau}_g)^* \xrightarrow{\phi_n}_g (\xrightarrow{\tau}_g)^* P'$ for $\phi_i \in X \cup \{\tau\}$, $i = 1, \ldots, n$.

Weak g-bisimulation has a similar definition to that of weak bisimulation [6]. It has been shown that weak g-bisimulation is a congruence for Bio-PEPA under certain reasonable conditions on g [4].

Weak g-bisimulation is applied to part of the model developed by Schoeberl *et al* [7] which describes the MAPK kinase signalling pathway. This pathway occurs when EGF (epidermal growth factor) binds to receptors on the external surface of the cell's membrane. The model is an ordinary differential equation (ODE) model and has been further considered by Gong and Zhang [5] in terms of redundant alternate pathways. An early part of this whole pathway (before the actual MAPK kinase cascade) has two alternate smaller pathways where one pathway is independent of the protein *Shc* and the other is dependent on *Shc*. The goal is to show using invariants and weak g-bisimulation that the two pathways are pairwise redundant – either pathway gives the same result. This is considered in a structural fashion and ignores specific rates.

Two Bio-PEPA models are constructed, one for each pathway. To show weak g-bisimulation, it is necessary to define a suitable function g that allows abstraction from irrelevant details. This is achieved by understanding the invariants of both pathways using the Bio-PEPA Plugin [2, 3]. With this knowledge it is possible to reassess the model in terms of input species, output species and other species.

Submitted to:
PASTA 2011

The function g is defined to capture a change in level of any of three specific species: the input species, the output species and a species which is an intermediate step between these two species. If there is no change in any of these species, the result of applying g is τ which indicates that the reaction that the transition represents is not relevant in assessing whether the two pathways have the same outcome.

Using numerical vector form to describe the Bio-PEPA models, a relation is defined over the two models where certain species from the two models are required to have the same levels, including the three species mentioned above. This relation is shown to be a weak g-bisimulation using an exhaustive case analysis.

Acknowledgements: The author was supported by the EPSRC SIGNAL Project, Grant EP/E031439/1.

References

[1] F. Ciocchetta & J. Hillston (2009): *Bio-PEPA: a framework for the modelling and analysis of biological systems*. Theoretical Computer Science, 410(33-34), pp. 3065–3084.

[2] Allan Clark, Stephen Gilmore, Maria Luisa Guerriero & Peter Kemper (2010): *On verifying Bio-PEPA models*. In: Proceedings of the 8th International Conference on Computational Models in Systems Biology (CMSB 2010), ACM Press, pp. 23–32.

[3] Allan Clark, Stephen Gilmore, Jane Hillston & Peter Kemper (2010): *Verification and testing of biological models*. In: Proceedings of the 2010 Winter Simulation Conference, pp. 620–630.

[4] V. Galpin (2011): *Equivalences for a biological process algebra*. Theoretical Computer Science, to appear, doi:10.1016/j.tcs.2011.07.006.

[5] Y. Gong & X. Zhao (2003): *Shc-dependent pathway is redundant but dominant in MAPK cascade activation by EGF receptors: a modeling inference*. FEBS Letters 554, pp. 467 – 472.

[6] R. Milner (1989): *Communication and concurrency*. Prentice Hall.

[7] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles & G. Muller (2002): *Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors*. Nature Biotechnology 20, pp. 270–375.