Adding bio-logic to query complete database compendiums

January 31, 2014

Receiving team: MeForBio Formal Methods for Bioinformatics
Research laboratory: IRCCyN, Ecole Centrale de Nantes
Head of laboratory: Malabre Michel (Michel.Malabre@irccyn.ec-nantes.fr)
Postal address: 1, rue de la Noë, BP 92101, 44321 NANTES CEDEX 3
Website: `http://www.irccyn.ec-nantes.fr`
Student mentor: Carito Guziolowski (carito.guziolowski@irccyn.ec-nantes.fr)

1 Project description

This project has as main objective to confront existing repositories of signaling and gene regulations with experimental data. It proposes a classical, and yet not so easy to implement, scenario when studying biological systems: model generation - analyses - experimental validation. In particular we wish to formalize and query complete database compendiums. The starting point of this project, will be to recover automatically regulatory knowledge from public databases of signaling and gene regulation. In a second step, we will formalize this knowledge by using frameworks such as Process Hitting (Paulevé *et al.*, 2011) or Answer Set Programming (Gebser *et al.*, 2010) in order add logic to the system. Once the data formalized, we are interested on formulating nontrivial queries to the system, such as: which node is the main regulator/controller of the expression of a set of genes, which molecule is key for a certain biological process, how to inverse the observed/measured logic of the system?
To illustrate this work-flow we will use as case study the Hepatocyte Growth Factor signaling network, relevant for the study of skin regeneration and liver cancer; for which an interaction graph has been compiled from public repositories in Guziolowski *et al.* (2012). Experimental measures of some essential genes of the system have been carried out by the Institute of Pathology (Kai Breuhahn team) of the University Hospital of Heidelberg, Germany, and concern measures upon the inhibition of relevant upstream molecules in this pathway. The objective of this internship will be to propose (automatic) ways to recover a model to verify these measures or complete the model in order to verify experiments.

References

Gebser, M., Guziolowski, C., Ivanchev, M., Schaub, T., Siegel, A., Thiele, S., and Veber, P. (2010). Repair and prediction (under inconsistency) in large biological networks with answer set programming. In F. Lin, U. Sattler, and M. Truszczynski, editors, *KR*. AAAI Press.

Guziolowski, C., Kittas, A., Dittmann, F., and Grabe, N. (2012). Automatic generation of causal networks linking growth factor stimuli to functional cell state changes. *FEBS J.*, **279**(18), 3462–3474.

Paulevé, L., Magnin, M., and Roux, O. (2011). Tuning Temporal Features within the Stochastic π-Calculus. *IEEE Transactions on Software Engineering*, **37**(6), 858–871.