Reverse Engineering by Artificial Network Evolution

Thomas Hinze[1], Ines Heiland[1], Benedict Schau[1], Thorsten Lenser[2], Stefan Schuster[1]

Friedrich Schiller University, [1]Department of Bioinformatics at School of Biology and Pharmacy, [2]Bio Systems Analysis Group, E.-Abbe-Platz 1-4, D-07743 Jena, Germany

thomas.hinze@uni-jena.de

Reverse engineering of biochemical reaction networks with incomplete or inconsistent topological information is a promising and challenging field at the crossroads of optimisation and model selection. Evolutionary algorithms have a long-established history as heuristic optimisation techniques [1]. Recently, methodologies adapted from this field have been used to evolve biochemical network candidates capable of performing arithmetic calculations or specific behaviours such as oscillations or switching [2]. Others employed similar techniques to reconstruct metabolic pathways from time series data of biochemical species [3]. While these attempts were successful for small networks, they also highlighted the complexity of evolving larger networks.

Within an ongoing study, we complement a control system-based specification of a circadian oscillator with artificial evolution of its modular components (e.g. controller, actuator, plant, sensor). Control systems benefit from a strict modularisation that allows a clear decomposition of a complex system into functional units interconnected by signalling cascades. Here, each of these units can be represented by one or more reaction network candidates exhibiting a desired input/output interdependency in terms of arithmetic functions. Input as well as output data have been expressed by corresponding species concentrations over time. In this way, each component of a control system can be independently reconstructed by providing numerous, topologically different network candidates. Finally, the arrangement of these candidates leads to valid models of the entire system. Coping with the complexity of evolving large networks, the search space is significantly reduced while keeping a high probability of heuristical success by our modular approach.

In order to successfully manage the technical process of artificial network evolution, we propose a separation of structural network evolution from kinetic parameter fitting which yields a pronounced increase in the algorithm's fitness performance [4]. Our studies show that this separation helps to prevent premature convergence when evolving networks executing arithmetic calculations. To this end, we have built the SBMLevolver [4], an open source software tool implementing our approach of artificial network evolution. Seven specific operators affecting the network topology as well as stoichiometry enables an effective structural evolution along with consecutive mutational modification of kinetic parameters. Currently, we exclusively employ mass-action kinetics. Several case studies illustrate the practicability of our modular approach for reverse engineering of building blocks with non-linear transmission behaviour.

[1] W. M. Spears. Evolutionary Algorithms. Springer Verlag, 1998

[2] P. Francois, V. Hakim. Design of Genetic Networks with Specified Functions by Evolution *in silico*. PNAS 101:580-585, 2004

[3] J.R. Koza, W. Mydlowec, G. Lanza, J. Yu, M.A. Keane. Automatic Synthesis of Both the Topology and Sizing of Metabolic Pathways using Genetic Programming. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001), pp. 57–65, Morgan Kaufmann, 2001

[4] T. Lenser, T. Hinze, B. Ibrahim, P. Dittrich. Towards Evolutionary Network Reconstruction Tools for Systems Biology. In E. Marchiori, J.H. Moore, J.C. Rajapakse (Eds.), Proceedings Fifth European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics (EvoBIO2007), Series Lecture Notes in Computer Science 4447:132-142, Springer Verlag, 2007