# Towards Evolutionary Network Reconstruction Tools for Systems Biology

Thorsten Lenser, Thomas Hinze, Bashar Ibrahim, and Peter Dittrich

Friedrich Schiller University Jena
Bio Systems Analysis Group
Ernst-Abbe-Platz 1–4, D-07743 Jena, Germany
{thlenser,hinze,ibrahim,dittrich}@minet.uni-jena.de

**Abstract.** Systems biology is the ever-growing field of integrating molecular knowledge about biological organisms into an understanding at the systems level. For this endeavour, automatic network reconstruction tools are urgently needed. In the present contribution, we show how the applicability of evolutionary algorithms to systems biology can be improved by a domain-specific representation and algorithmic extensions, especially a separation of network structure evolution from evolution of kinetic parameters. In a case study, our presented tool is applied to a model of the mitotic spindle checkpoint in the human cell cycle.

## 1 Introduction

Reverse engineering of biochemical networks, making sense of rapidly growing molecular proteomics data, is a promising and important field at the crossroads of optimisation and model selection. Supplementing human-curated models with automatically generated, data-based models will enhance our understanding of the function of cells as a whole, which is at the core of systems biology.

Evolutionary algorithms (EA) and especially genetic programming (GP) have a long-established history as heuristic optimisation techniques [2,13,15]. Recently, methodologies adapted from this field have been used to evolve artificial biochemical networks, capable of performing arithmetic calculations [7] or specific behaviours such as oscillations and switching [10,17]. Others have used similar techniques to reconstruct metabolic pathways from time series data of chemical species [14]. While these attempts were successful for small networks, they also highlighted the complexity of evolving larger networks.

To expand our capability of evolving networks, improvements on these algorithms have to be investigated. In this contribution, we propose a separation of structural evolution of the network from kinetic parameter evolution, which yields a pronounced increase in the algorithm's fitness performance. Our studies show that this separation helps to prevent premature convergence when evolving networks performing arithmetic calculations. We suppose this happens because parameter fitting after each structural mutation smoothes their effect, which is usually rather strong. In this way, network parameters can adapt to a new topology before this topology is evaluated.

E. Marchiori, J.H. Moore, and J.C. Rajapakse (Eds.): EvoBIO 2007, LNCS 4447, pp. 132–142, 2007.

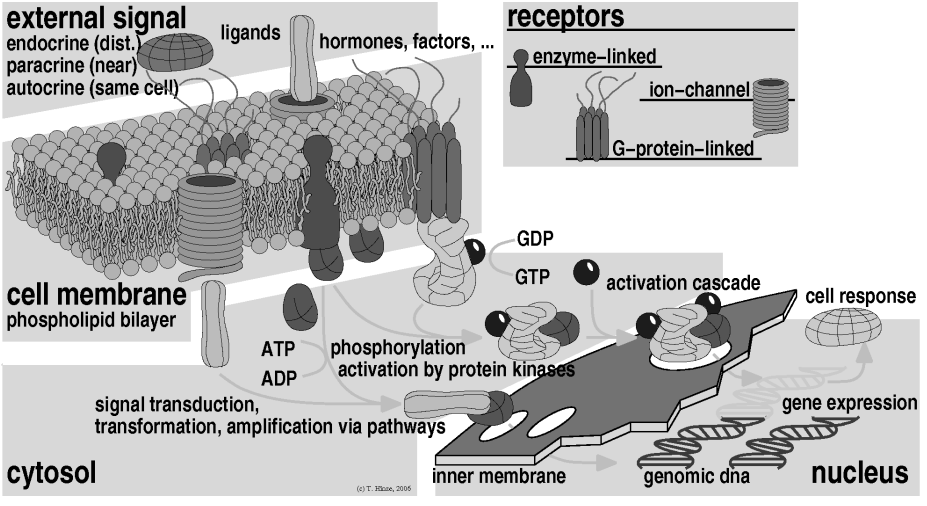


**Fig. 1.** Biological principle of signalling in eukaryotic cells: from arriving stimuli to specific cell response

Two other ideas are also investigated: the biologically-inspired mutation operator of species duplication, and the use of Akaike's Information Criterion (AIC) as a fitness function to evolve parsimonious models. By using the markup language SBML, the tool described here can work directly on systems biological problems, aiming at applications throughout this growing community.

In a case study, we apply our algorithm to a model of the human mitotic spindle checkpoint. By allowing the algorithm to introduce additional reactions, the performance of the model can be increased in comparison to a mere optimisation of parameters. Although biological plausibility is not considered, the example serves as a proof of concept for further investigations.

## 2 Modelling and Evolving Biochemical Networks

Biochemical reaction networks found in pro- and eukaryotic cells represent important components of life. Despite their high degree of complexity, they are hierarchically arranged in modular structures of astonishing order. The function of a cell emerges from the interplay of connected reaction processes. Three essential types of biochemical networks can be distinguished: metabolic, cell signalling (CSN), and gene regulatory (GRN) networks [1]. While metabolism consists of coupled enzymatically catalysed reactions supplying energy, CSNs and GRNs perform information processing of external and internal signals [6]. Malfunctions or perturbations within these networks are the cause of many diseases.

We have built a software tool implementing an evolutionary algorithm that evolves artificial biochemical networks performing pre-specified tasks. As a representation format, the systems biology standard SBML [9] is used, the most