

Towards Evolutionary Network Reconstruction Tools for Systems Biology

T. Lenser T. Hinze B. Ibrahim P. Dittrich

{thlenser,hinze,ibrahim,dittrich}@cs.uni-jena.de

Bio Systems Analysis Group
Friedrich Schiller University Jena

www.minet.uni-jena.de/csb

5th European Conference on
Evolutionary Computation,
Machine Learning and
Data Mining in Bioinformatics



Outline

Towards Evolutionary Network Reconstruction Tools in Systems Biology

Introduction

Motivation, Cell Signalling, ESIGNET

Artificial Network Evolution

Two-Level Evolutionary Algorithm

Operators, Parameterisation, Fitting

Selection and Fitness Evaluation

Results

Evolving Arithmetic Functions \log , $\sqrt[3]{}$

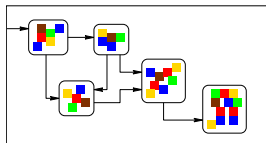
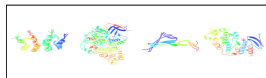
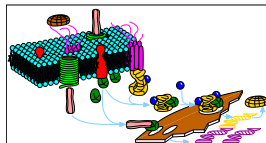
Effect of Duplication Operator

Case Study: Spindle Checkpoint

Biological Background

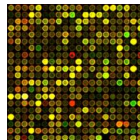
Modelling and Evol. Optimisation

Conclusions



Motivation

- Systems Biology deals with interplay of biological components rather than components themselves.
- Reconstructing nonlinear networks from (incomplete) data is a necessary but difficult task.



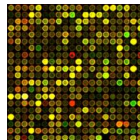
gene expression data
visualised by microarray
(TU Dresden, BIOTEC)

⇒ Evolutionary computing is well suited to this!

- Furthermore, bio-inspired algorithms provide a flexible, fault-tolerant, reliable computing paradigm.
⇒ Evolutionary computing can support design of such algorithms.
- Help in understanding emergence of biological complexity.
⇒ Evolution becomes observable.

Motivation

- Systems Biology deals with interplay of biological components rather than components themselves.
- Reconstructing nonlinear networks from (incomplete) data is a necessary but difficult task.



gene expression data
visualised by microarray
(TU Dresden, BIOTEC)

⇒ Evolutionary computing is well suited to this!

- Furthermore, bio-inspired algorithms provide a flexible, fault-tolerant, reliable computing paradigm.

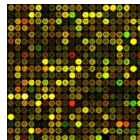
⇒ Evolutionary computing can support design of such algorithms.

- Help in understanding emergence of biological complexity.

⇒ Evolution becomes observable.

Motivation

- Systems Biology deals with interplay of biological components rather than components themselves.
- Reconstructing nonlinear networks from (incomplete) data is a necessary but difficult task.



gene expression data
visualised by microarray
(TU Dresden, BIOTEC)

⇒ Evolutionary computing is well suited to this!

- Furthermore, bio-inspired algorithms provide a flexible, fault-tolerant, reliable computing paradigm.

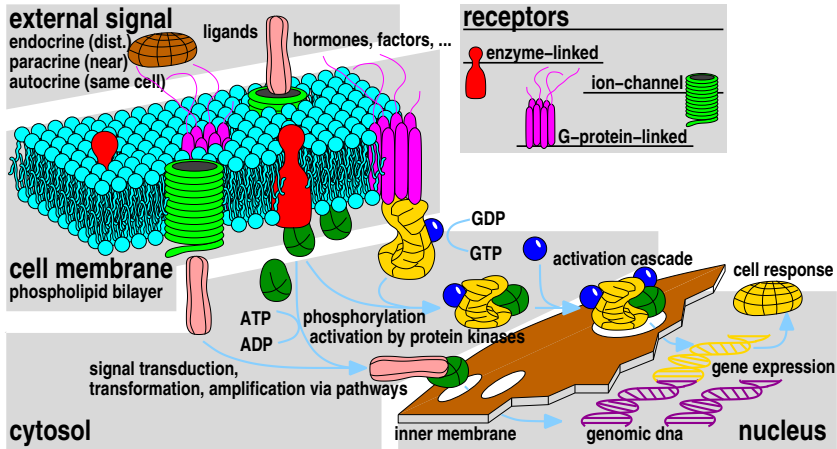
⇒ Evolutionary computing can support design of such algorithms.

- Help in understanding emergence of biological complexity.

⇒ Evolution becomes observable.

Biological Principles of Cell Signalling

Information Processing in Living Cells



ESIGNET – Research Project

Evolving Cell Signalling Networks (CSNs) *in silico*

European interdisciplinary research project

- University of Birmingham (Computer Science)
- TU Eindhoven (Biomedical Engineering)
- Dublin City University (ALife Lab)
- University of Jena (Bio Systems Analysis)



SIXTH FRAMEWORK
PROGRAMME



Objectives

- Study the computational properties of CSNs
- Developing new ways to model and predict real CSNs
- Gain new theoretical perspectives on real CSNs



TU/e

Computing Facilities

- Cluster of 33 workstations
(two Dual Core AMD Opteron™ 270 processors, Rocks Linux)



ESIGNET – Research Project

Evolving Cell Signalling Networks (CSNs) *in silico*

European interdisciplinary research project

- University of Birmingham (Computer Science)
- TU Eindhoven (Biomedical Engineering)
- Dublin City University (ALife Lab)
- University of Jena (Bio Systems Analysis)



SIXTH FRAMEWORK
PROGRAMME



TU/e

Objectives

- Study the computational properties of CSNs
- Developing new ways to model and predict real CSNs
- Gain new theoretical perspectives on real CSNs

Computing Facilities

- Cluster of 33 workstations
(two Dual Core AMD Opteron™ 270 processors, Rocks Linux)



ESIGNET – Research Project

Evolving Cell Signalling Networks (CSNs) *in silico*

European interdisciplinary research project

- University of Birmingham (Computer Science)
- TU Eindhoven (Biomedical Engineering)
- Dublin City University (ALife Lab)
- University of Jena (Bio Systems Analysis)



SIXTH FRAMEWORK
PROGRAMME



Objectives

- Study the computational properties of CSNs
- Developing new ways to model and predict real CSNs
- Gain new theoretical perspectives on real CSNs



TU/e

Computing Facilities

- Cluster of 33 workstations
(two Dual Core AMD Opteron™ 270 processors, Rocks Linux)

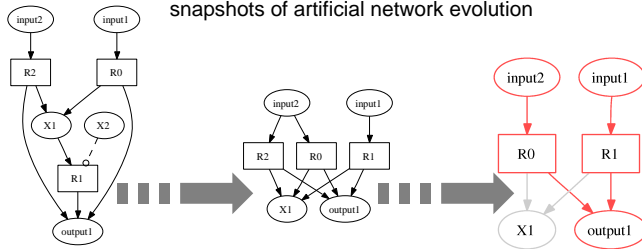


Artificial Network Evolution

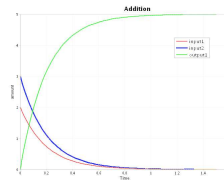
Introductory Example

Task: addition of two positive real numbers

snapshots of artificial network evolution



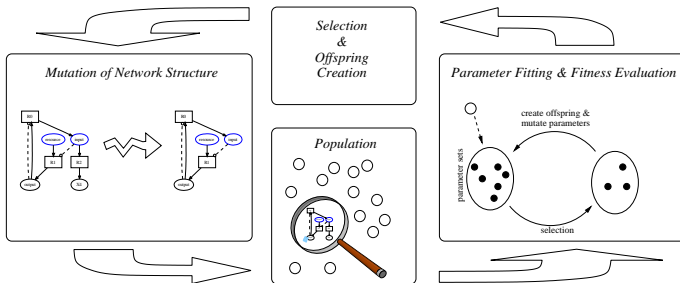
- **R0, R1, R2** identify reactions
- **input1, input2, output1**: distinguished species
- **X1, X2**: auxiliary species
- Stepwise modification of network structure and kinetic parameters



Two-Level Evolutionary Algorithm

Artificial Network Evolution in Detail

- Separation of structural evolution from parameter fitting
- Idea: parameters can adapt to mutated network structure



- Upper level: network structure, analogue to graph-GP
- Lower level: parameter fitting using standard Evolution Strategy

⇒ All networks handled as SBML models

Operators and Parameterisation

Artificial Network Evolution in Detail

EA used here employs eight different mutations

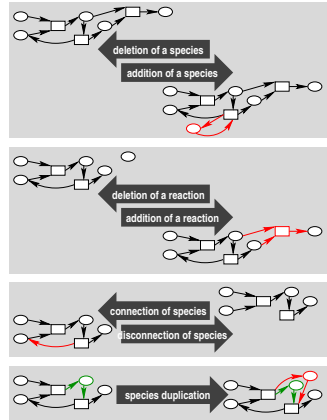
Operators for structural evolution

- Addition/deletion of a species
- Addition/deletion of a reaction
- Connection/removal of an existing species to/from a reaction
- Duplication of a species with all its reactions (discussed in detail later)

Operator for parameter evolution

- Mutation of a randomly selected kinetic parameter by addition of a Gaussian variable

Network size can be limited.



Parameter Fitting by an Evolution Strategy

Artificial Network Evolution in Detail

- **Small population size**
⇒ due to high computational costs of fitness evaluation
- **Non-overlapping generations** (comma-selection)
⇒ supports self-adaptation
- **Self-adaptation of strategy parameters**
⇒ balancing between exploration of search space and fine-tuning
- **Parameter settings copied from parent to offspring networks**
⇒ incremental parameter fitting
- **Initial parameters uniformly distributed between given minimal and maximal values**
⇒ no extra bias introduced

Selection and Fitness Evaluation

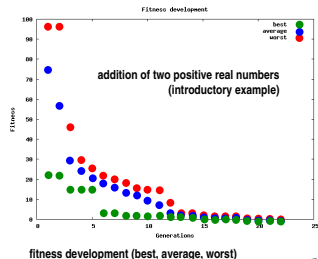
Artificial Network Evolution in Detail

Selection on structural level

- **Overlapping generations** (plus-selection), elitistic
⇒ good solutions cannot be lost
- **Fixed population size** (10 ... 100)
⇒ due to computational costs

Fitness evaluation

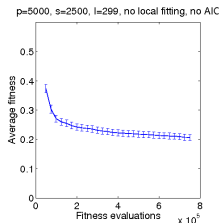
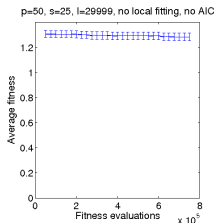
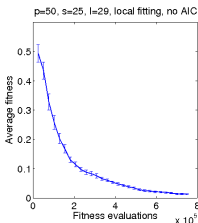
- Numerical integration of reaction network using ODE solver (SOSlib)
- Fitness measure given by weighted squared distance to target time course (output species)



Result: Logarithm Network

Evolving Arithmetic Functions

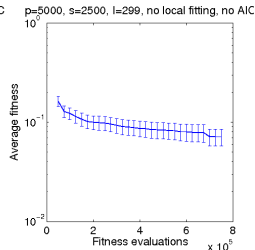
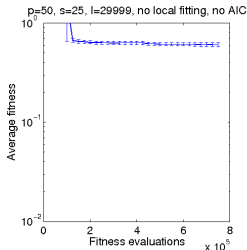
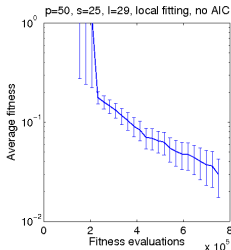
- We compare three settings:
 - two-level EA
 - one-level EA (simultaneous structural and parameter evolution) for many generations
 - one-level EA with a larger population
- Setup such that all approaches use same number of fitness evaluations (normalisation)
- Two-level approach clearly superior for this task
- Both other approaches converge prematurely



Result: Third Root Network

Evolving Arithmetic Functions

- Same three settings as in logarithm example
- Depicted on logarithmic scale
- Result also confirms advantage of two-level EA
- Differences are not as pronounced as in logarithm example

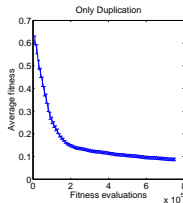
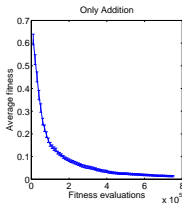
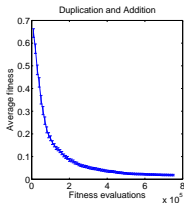


Result: Effect of Duplication Operator

Duplication of a species with all its reactions



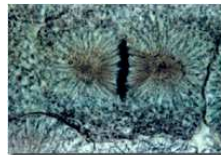
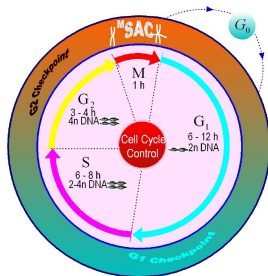
- Search for “soft” mutation operators
- Inspired by gene duplication in living organisms
- Adapted to evolution of arithmetic functions



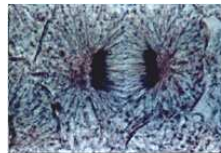
- In log example, duplication doesn't improve or worsen observed results
- However, we still regard duplication as potentially promising

Case Study: Human Spindle Assembly Checkpoint

Biological Background



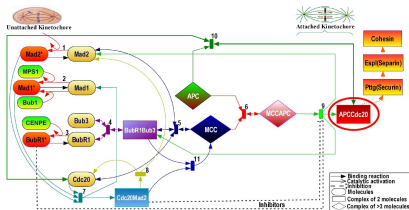
Spindle Assembly
Checkpoint (SAC)



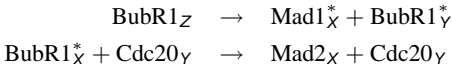
- Sequence of events starting from one cell leading to two daughter cells
- Focus on checkpoint mechanism in mitosis (cell division)
- SAC prevents cell cycle progression until all chromosomes are attached to mitotic spindle
- Defects lead to cell death, aneuploidy, aging, and cancer

Case Study: Human Spindle Assembly Checkpoint

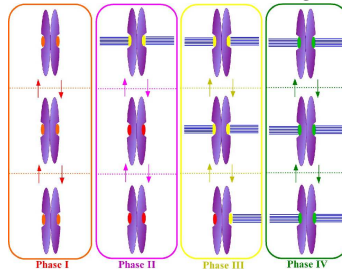
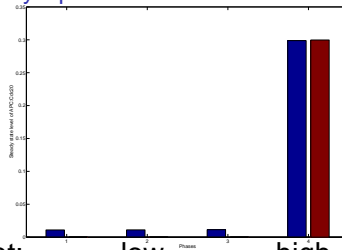
Modelling and Evolutionary Optimisation



- 17 species, 11 reactions
- Compartments represent chromosomes
- Structural evolution adds two (unrealistic) reactions:



target: ——— low ——— high ———



Conclusions and Outlook

- Two level approach enhances performance of EA for biological network reconstruction
- Duplication operator interesting and promising approach in general, although first results for the logarithm network show no convincing effect yet
- So far, we mostly tested evolution of networks for arithmetic functions
- Evolutionary method can improve and predict realistic complex networks exemplified here by human Spindle Assembly Checkpoint
- Further studies will target additional parameter settings as well as application to evolution of computing devices
- Interested in our software? thlenser@minet.uni-jena.de