

Integrated System for Inference of Gene Expression Network

Masahiko Nakatsui¹

nakatsui@brs.kyushu-u.ac.jp

Takanori Ueda²

ueda@brs.kyushu-u.ac.jp

Masahiro Okamoto¹

okahon@brs.kyushu-u.ac.jp

¹ Laboratory for Bioinformatics, Graduate School of Systems Life Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

² Laboratory for Applied Biological Regulation Technology, Graduate School of Biore-source and Bioenvironmental Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

Keywords: inverse problem, numerical optimization, S-system formalism, real-coded genetic algorithm, uniform normal distribution crossover, generation alternation models, gene expression network

1 Introduction

Recent advances of powerful new technologies such as DNA microarrays provide a mass of gene expression data on a genomic scale. One of the most important projects in post-genome-era is the system identification of gene expression networks by using these observed data.

We previously proposed an efficient numerical optimization technique by using time-course data of system components, which is based on real-coded genetic algorithm (RCGAs [1, 2]) to estimate the interrelated coefficients among system components of a dynamic network model called S-system [3] that is a type of power-law formalism and is suitable for description of organizationally complex systems such as gene expression networks and metabolic pathways. This technique with the combination of unimodal normal distribution crossover (UNDX) [1], which is one of the crossover operators for RCGAs, with the alternation of generation model called minimal generation gap (MGG) [2] showed superiority to the conventional simple genetic algorithm [4, 5].

In this study we shall describe on the integrated inferring system involving the GUI program for real-time visualization of the inferred network structures by using distributed parallel computer systems.

2 Integrated Inferring System

Integrated System consists of two programs, that is inferred engine module and visualization module, respectively.

2.1 Inferred Engine

Inferred engine can be driven by an efficient algorithm [4, 5] for narrowing down the network candidates that possibly explain the observed time series of expression profiles within the immense huge searching space of parameter values. At starting of the inference, experimentally observed time-course data and conditions for optimization are transmitted to the distributed parallel computer system as arguments of a command. Inferred engine were developed by C++. MPI-2 was used for starting of program, and for communication between each process of distributed parallel machines.

2.2 GUI Program

The GUI program has been developed for real time visualization of inferred network structures. A snapshot of GUI program is shown in Figure. 1.

Upper half of Fig. 1 shows transient optimizing progress of all processes in parallel machines; the ordinate shows averaged error value between calculated values and experimental ones, the abscissa is generation number of genetic algorithm. Lower half shows inferred network structure over generation on an arbitrary assigned process, which are automatically updated according to optimizing progress. Right sided upper small window shows a common core interactions of all inferred network structures which were obtained by all processes. This information is very useful for reliability of inference of interactions and shows minimum requirement for realization of experimentally given observed time-courses. We developed the GUI program using Java. In order to communicate with inferred engine by MPI-2, the communicating part is coded by C++ and is called through JNI(Java Native Interface) from the GUI program.

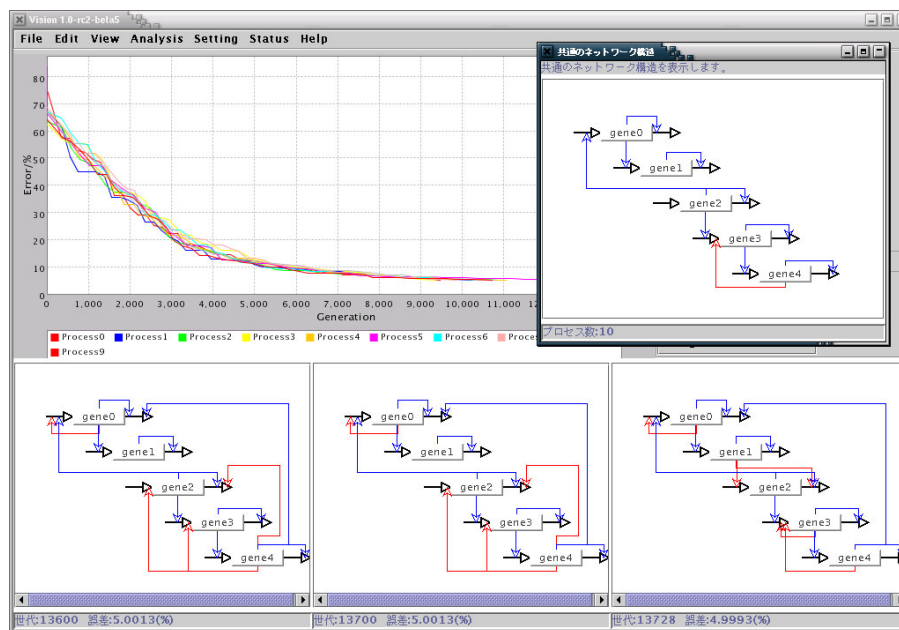


Figure 1: A snapshot of Integrated Inferring System.

References

- [1] Ono, I. and Kobayashi, S., A real-coded genetic algorithm for function optimization using uni-modal distribution crossover, *Proc. 7th ICGA*, 249–253, 1997.
- [2] Sato, H., Ono, I., and Kobayashi, S., A new generation alternation model of genetic algorithm and its assessment, *J. of Japanese Society for Artificial Intelligence*, 12(5):734–744, 1997.
- [3] Savageau, M.A., *Biochemical Systems Analysis: A study of function and design in molecular biology*, Addison–Wesley, Reading, 1976.
- [4] Ueda, T., Koga, N., and Okamoto, M., Efficient numerical optimization technique based on real-coded genetic algorithm, *Genome Informatics*, 12:451–453, 2001.
- [5] Ueda, T., Ono, I., and Okamoto, M., Development of system identification technique based on real-coded genetic algorithm, *Genome Informatics*, 13:386–387, 2002.