Bioinformatics Center
– Mathematical Bioinformatics –

http://www.bic.kyoto-u.ac.jp/takutsu/index.html

Prof
AKUTSU, Tatsuya
(D Eng)

Assist Prof
HAYASHIDA, Morihiro
(D Inf)

Assist Prof
TAMURA, Takeyuki
(D Inf)

SPD (JSPS)
SIMONE, Marini
(Ph D)

PD
JIRA, Jindalertudomdee
(D Inf)

Students

BAO, Yu (D3)
LIU, Pengyu (D1)
KAWAKAMI, Yuko (M2)
FUKUSAKO, Yuta (M2)
LI, Ruiming (M1)
CAO, Yu (RS)
PI, Wenya (RS)

Guest Scholars

VINCHENT, Alexandra — University of Montpellier, France, 20 April–16 July
HOU, Wenpin (Ph D) — The University of Hong Kong, China, P.R., 23 May–22 August

Guest Res Assoc

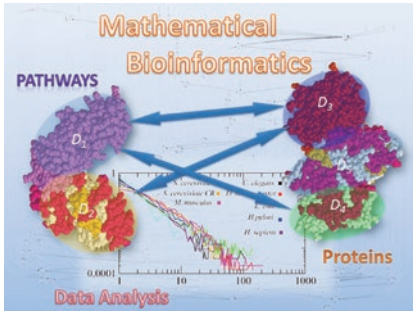
YANG, Jinn-Moon (Ph D) — National Chiao Tung University, Taiwan, 4 February–4 May
CHENG, Xiaoqing (Ph D) — The University of Hong Kong, China, P.R., 6 March–3 April

Scope of Research

Due to rapid progress of genome sequencing technology, whole genome sequences of organisms ranging from bacteria to human have become available. In order to understand the meaning behind the genetic code, we have been developing algorithms and software tools for analyzing biological data based on advanced information technologies such as theory of algorithms, artificial intelligence, and machine learning. We are currently studying the following topics: systems biology, scale-free networks, protein structure prediction, inference of biological networks, chemo-informatics, and discrete and stochastic methods for bioinformatics.

KEYWORDS

Scale-free Networks
Boolean Networks
Chemical Graphs
Grammar-based Compression
Protein Complexes



Selected Publications

Hou, W.; Tamura, T.; Ching, W. K.; Akutsu, T., Finding and Analyzing the Minimum Set of Driver Nodes in Control of Boolean Networks, *Advances in Complex Systems*, **19,** 1650006 (2016).

Takemoto, K.; Akutsu, T., Analysis of the Effect of Degree Correlation on the Size of Minimum Dominating Sets in Complex Networks, *PLoS ONE*, **11,** e0157868 (2016).

Jindalertudomdee, J.; Hayashida, M.; Akutsu, T., Enumeration Method for Structural Isomers Containing User-defined Structures Based on Breadth-first Search Approach, *J. Comput. Biol.*, **23,** 625-640 (2016).

Ishitsuka, M.; Akutsu, T.; Nacher, J. C., Critical Controllability in Proteome-wide Protein Interaction Network Integrating Transcriptome, *Sci. Rep.*, **6,** 23541 (2016).

Enumeration Method for Structural Isomers Containing User-defined Structures Based on Breadth-first Search Approach

Enumeration of chemical structures is useful for discovering new compounds and drugs, and elucidating chemical structures from mass spectrometry. We previously developed efficient algorithms, BfsSimEnum, BfsMulEnum, BfsBenNaphEnum for enumerating tree-like chemical compounds without and with multiple bonds, and compounds containing rings such as benzene and naphthalene, respectively. For many instances, the algorithms were able to enumerate chemical structures faster than other existing methods.

In this study, we propose a novel efficient enumeration algorithm, BfsStructEnum, which allows users to define desired cyclic structures and enumerates all nonredundant chemical compounds containing only user-defined structures as cyclic structures from a given chemical formula.

For evaluating the performance, we confirmed that the number of enumerated structures of BfsStructEnum was the same as that of an existing method, MOLGEN for several instances. Furthermore, we also found that BfsStructEnum was significantly faster than MOLGEN. It implies that compression of cyclic structures into single nodes makes the enumeration more efficiently.

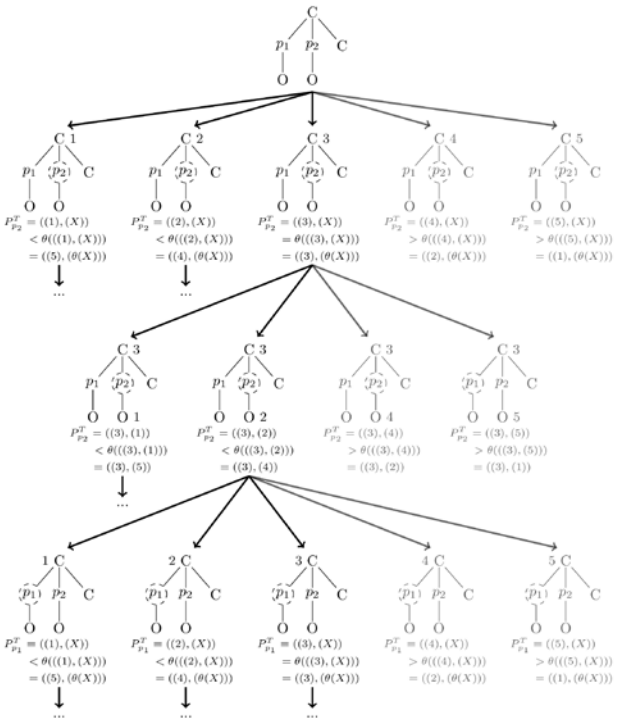


Figure 1. Example of the assignment process of atom position lists to pyridine nodes (denoted by p1 and p2).

Finding Influential Genes using Gene Expression Data and Boolean Models of Metabolic Networks

Selection of influential genes using gene expression data from normal and disease samples is an important topic in bioinformatics. In this research, we propose a novel computational method for the problem, which combines gene expression patterns from normal and disease samples with a mathematical model of metabolic networks. This method seeks a set of k genes knockout of which drives the state of the metabolic network towards that in the disease samples. We adopt a Boolean model of metabolic networks and formulate the problem as a maximization problem under an integer linear programming framework. We applied the proposed method to selection of influential genes using gene expression data from normal samples and disease (head and neck cancer) samples. The result suggests that the proposed method can select more biologically relevant genes than an existing P -value based ranking method can.

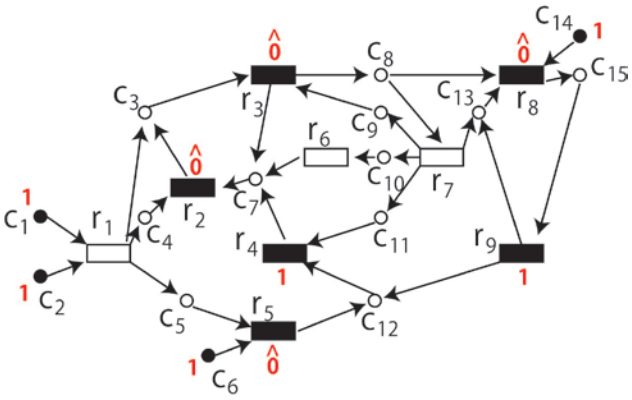


Figure 2. In Boolean metabolic networks, rectangles and circles represent reactions and metabolites, respectively. Each node is assigned either 0 or 1 to represent the status of the node.