Bioinformatics Center
– Mathematical Bioinformatics –

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

Prof
AKUTSU, Tatsuya
(D Eng)

Assoc Prof
TAMURA, Takeyuki
(D Inf)

Assist Prof
MORI, Tomoya
(D Inf)

Guest Res Assoc*
SUN, Liangjie
(Ph D)

Program-Specific Res
LI, Ruiming (Ph D)

*JSPS Postdoctoral Fellowships for Research in Japan, China, 1 November 2023–31 October 2025

Students

TAKAGI, Motoshige (D3)	MU, Lixuan (D3)	MA, Yier (D2)	CUI, Yan (M2)
OHTOMO, Masahiro (D3)	SHIOTA, Koji (D3)	FUJITA, Satoki (D2)	IWAKI, Takuma (M1)
NAKASHIMA, Shogo (D3)	LIU, Chunting (D3)	GHAFOOR, Mamoona (D2)	WU, Chenyao (D1)

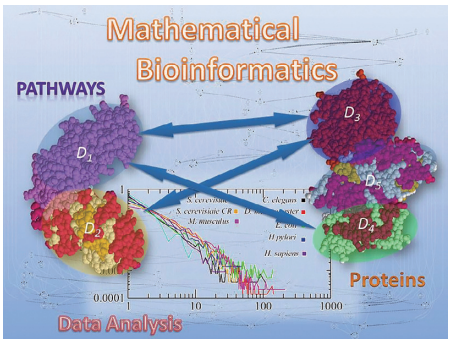
Guest Res Assoc

ZHAO, Jiaying	The University of Hong Kong, Hong Kong, 31 May 2023–21 August 2023
LI, Wei	Nankai University, China, 22 August 2023–31 May 2024
HAN, Mingjing	Nankai University, China, 22 August 2023–31 May 2024
ZHONG, Yezhao	National University of Ireland, Galway, Ireland, 6 November 2023–26 January 2024
HSU, Zi-Chu	National Yan Ming Chiao Tung University, Taiwan, 25 November 2023–15 December 2023
CHEN, Tze-Shin	National Yan Ming Chiao Tung University, Taiwan, 25 November 2023–15 December 2023
TUNG, Cheng-Yu	National Yan Ming Chiao Tung University, Taiwan, 25 November 2023–15 December 2023

Scope of Research

Due to the 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, the inference of biological networks, chemo-informatics, and discrete and stochastic methods for bioinformatics.

KEYWORDS
Complex Networks
Boolean Networks
Neural Networks
Chemical Graphs
Protein Informatics



Recent Selected Publications

Melkman, A. A.; Guo, S.; Ching, W-K.; Liu, P.; Akutsu, T., On the Compressive Power of Boolean Threshold Autoencoders, *IEEE Transactions on Neural Networks and Learning Systems*, **34,** 921-931 (2023).

Matsuda, K.; Shirakami, A.; Nakajima, R.; Akutsu, T.; Shimono, M., Whole-Brain Evaluation of Cortical Microconnectomes, *eNeuro*, **10,** ENEURO.0094-23.2023 (2023).

Mori, T.; Takase, T.; Lan, K-C.; Yamane, J.; Alev, C.; Kimura, A.; Osafune, K.; Yamashita, J. K.; Tatsuya, A.; Kitano, H.; Fujibuchi, W., eSPRESSO: Topological Clustering of Single-Cell Transcriptomics Data to Reveal Informative Genes for Spatio-Temporal Architectures of Cells, *BMC Bioinformatics*, **24,** 252 (2023).

Cao, Y.; Pi, W.; Lin, C-Y.; Münzner, U.; Ohtomo, M.; Akutsu, T., Common Attractors in Multiple Boolean Networks, *IEEE ACM Transactions on Computational Biology and Bioinformatics*, **20,** 2862-2873 (2023).

Tamura, T., Trimming Gene Deletion Strategies for Growth-Coupled Production in Constraint-Based Metabolic Networks: TrimGdel, *IEEE ACM Transactions on Computational Biology and Bioinformatics*, **20,** 1540-1549 (2023).

On the Compressive Power of Autoencoders Using Linear Threshold Activation Functions

Artificial neural networks have recently been extensively applied to bioinformatics. Among various models of artificial neural networks, autoencoders attract much attention because of their power to generate new objects such as protein sequences and chemical structures. An autoencoder is a layered neural network (Figure 1) consisting of an encoder which compresses an input vector to a lower dimensional vector, and a decoder which transforms the low-dimensional vector back to the original input vector (or one that is very similar). Although it is often mentioned that autoencoders perform dimensionality reduction, a kind of data compression, how data are compressed is not yet very clear. Therefore, we study the numbers of nodes and layers that are required to ensure that each vector in a given set of distinct input binary vectors is transformed back to its original using a autoencoder model with linear threshold activation functions. We show that for any set of distinct vectors there exists a seven-layer autoencoder with the optimal compression ratio, but that there is a set of vectors for which there is no three-layer autoencoder with a middle layer of the same size. We also study the numbers of nodes and layers required only for encoding, and the results suggest that decoding is more difficult than encoding.

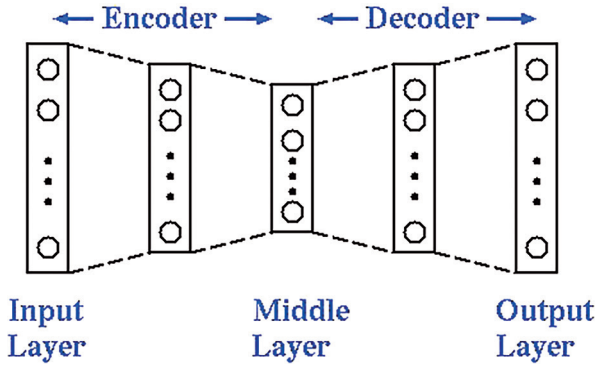


Figure 1. Architecture of an autoencoder.

Trimming Gene Deletion Strategies for Growth-Coupled Production in Constraint-Based Metabolic Networks: TrimGdel

When simulating genome-scale metabolite production using constraint-based metabolic networks, it is often necessary to find gene deletion strategies which lead to growth-coupled production, which means that target metabolites are produced when cell growth is maximized. One of the best current methods for this problem is the minimal cut set-based method, which utilizes the fact that minimal cut sets in the primal network are the elementary modes in the corresponding dual network. This method is effective when the number of gene deletions is relatively small, but when the number of required gene deletions exceeds 20, the time required for the calculation is often unfeasible. Therefore, a complementing algorithm that is effective even when the required number of gene deletions is approximately 20 to 40 would be helpful because the number of genes that can be deleted in a strain is increasing with advances in genetic engineering technology. In this study, the present author developed an algorithm, TrimGdel, which first computes a strategy with many gene deletions that results in growth-coupled production and then gradually reduces the number of gene deletions while maintaining the production rate and growth rate. The results of the computer experiments showed that, for 34.2% of the target metabolites in iML1515, the genome-scale constraint-based model of Escherichia coli, TrimGel could calculate gene deletion strategies of size 40 or less (23.6% for between 21 and 40) leading to growth-coupled production. TrimGdel can calculate stoichiometrically feasible gene deletion strategies, especially of sizes 21 to 40, which lead to growth-coupled production of target metabolites, which include useful vitamins such as biotin and pantothenate, for which existing methods could not.

The developed software is available on https://github.com/MetNetComp/TrimGdel

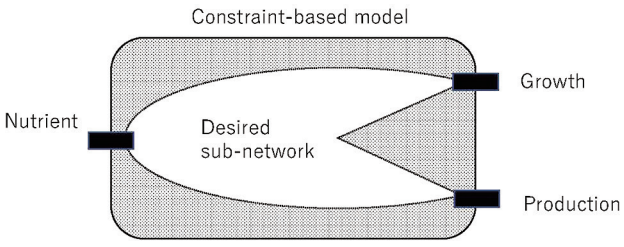


Figure 2. Each gene deletion strategy is evaluated by the least value of the target metabolite production rate when maximizing the cell growth rate.