若手研究者インターナショナル・トレーニング・プログラム(ITP)

バイオインフォマティクスとシステムズバイオロジーの国際連携教育研究プログラム 応募書類

Name: Yuki Narita
Title: Metrization of information entropies and their algebraic applications on biological networks
Institute: Bioinformatics center, Institute for Chemical Research, Kyoto University
Partner institute of your choice: AG Mathematics in Life Sciences, Department of Mathematics and Computer Science, Free University Berlin
Duration of your choice: Jun. 4—Sep. 1, 2012

Plan :

Background.

The partner institute, AG mathematics in life sciences (Bockmayr's group) in the Free university of Berlin, as well as the mathematical bioinformatics lab.(Akutsu lab.) in the bioinformatics center, Kyoto university, is one of the leading research team studying on the mathematical and computer-scientific aspects and the related computer simulations on the biological networks. Especially, while Akutsu lab. is working on discrete systems and algorithmic aspects, Bockmayr's group is rather motivated by the viewpoints of logic and computation, and focusing on the hybrid systems of the discrete and continuous models. In addition to this, Bockmayr's group belongs to IRTG which includes not only the research groups in systems biology and bioinformatics, there are also flexibilities to communicate and collaborate with those in the field of experimental biology.
Although mathematical biology already has more than a half century history and its research objectives are tremendously varied, on the other hand, at the same time experimental biologists have overwhelmingly enriched their methods and scopes. Therefore simply applying the existing theoretical methods seems to be inadequate and there are needs to explore novel approaches that capture some newly discovered biological features.
Since I have been interested in and working on mathematical models appearing in biological systems, and have learned especially mathematical issues including topics in discrete geometry related to bioinformatics under supervision of Prof. Akutsu, it is greatly expected in joining Bockmayr's group that I can improve my mathematical skills, communication skills with experimental biologists, and widen my view on mathematical biology; so I strongly hope to join them via ITP program.

Research Objective.

Recent advances in molecular biology have provided vast varieties of intra, extra or inter-cellular molecular communicatory networks governing biological phenomena such as metabolic networks, gene regulatory networks, signal transduction networks, …, etc.. Not only that there are already huge amount of such data sets, existing theoretical frameworks to approach them are also quite varied and composite. However, simply putting these frameworks to capturing the total cellular picture, each of the existing approaches confronts some limitation in scales: for example, naively simulating the full chemical processes in a cell by simultaneous differential equations is obviously either practically or mathematically impossible.
Within this context during the ITP program I am going to work on the following theoretical framework which concerns a type of variation of information entropies:

Plan (Continued)

Research Topic*: *Metrization of Information Entropies.

The concept of information entropy is extensively used in studying biological systems. On the other hand, while we are often rather interested in interactions, dependencies and correlations among several biophysical compartments, Kullback-Leibler divergence (Boltzmann's relative entropy) have been the most currently adopted and also the almost only tool as an extended form of information entropy to examine mutual dependencies between two probability measures whereas the underlying two measurable spaces should share a common one (usually $\mathbf{R}_{\geq 0}$, interpreted as "amounts of substrates"). However, from some viewpoints this may be non-realistic at all that: for example, not only the amount but also conformations, configurations and spatial and temporal distributions of the substrates affect the cellular behavior, although the resulting discriminations and transmissions of biological information seems to be comparably simple (such as ON/OFF switching).
In this respect I have developed a novel extended version of information entropies reflecting the probabilistic features of signal transductions between two given probability measures. This version of entropies shows properly different features (when it reaches to 0; when it increases, …, etc.) from Kullback-Leibler divergence and notably the underlying measurable spaces are not necessarily assumed to be the same. However, this abstraction is too strong for actual application that:

- Any model will fit but no one can tell which model is biologically precise.
- The mathematical restriction is too weak that it seems hard to find a new interesting property for biological applications any more.

To overcome these weaknesses I'm going to restrict this framework into the cases in metric spaces, and at the same time conversely extend this where the general case can be obtained as a special example. It is strongly expected that strategies in this direction involve the following notable advantages:

- Analysis on metric measure spaces without any priori smooth structure itself has been in much mathematical progress in recent years so that there might be meets in mathematical interests with existing studies and possibly applications of known results.
- At Akutsu lab. I have already surveyed geometrical issues in metric spaces (especially low-distortion embedding problem of finite metric spaces into normed spaces), and many of these issues are known to have close relationship with biological problems, it may be a good opportunity to make use of these knowledge and experiences on metric spaces.
- One of the main issues treated in Bockmayr's group has been the hybridizations and integrations of discrete and continuous spaces to study biological systems. Since metric structure is one of the most frequently assumed structures regardless of discrete or continuous, it is expected that this research will propose some flexibility which allows us to replace discrete models to continuous ones and vice versa.
- Bockmayr's group has also been concerned with "constraint-based methods" in systems biology: that is, accepting the fact that biological data are usually incomplete so that it is often impossible to make exactly correct predictions, we can still predict what *might* occur regardless of their correctness. Although the framework I have already worked with still seems to be too abstract to give such possible predictions, the study I'm going to take on might also give some insight in studies in setting precise restriction.