Field of the Master: Computer science, Bioinformatics, Statistical physics

Level: M1/M2

Required skills: computer programming, statistics

Duration: 3 to 6 months

Period: any periods between January and July 2016

Title of the research project: Development and implementation of an efficient Bayesian inference scheme for systems biology

---

Name of the supervisor: Daniel Jost (daniel.jost@imag.fr)

Laboratoire TIMC-IMAG, CNRS, Université de Grenoble Alpes.

Context: Nowadays, experimental biology generates an important mass of quantitative data. Thus, the development of numerical tools to analyze and interpret these data is strongly needed in all fields of biology. In particular, in systems biology, a big challenge is to understand the regulatory interactions between genes from various sources of data. A possible strategy is to build quantitative mathematical models that describe the dynamics of each entity in the network, and to infer the multiple parameters of such models. However when the number of parameters is large and/or the number of data is small and/or the uncertainties on the data points are significant, standard inference schemes might not be adapted and may lead to an incomplete or an erroneous description of the data.

Objectives: In collaboration with the experimental group of Gael Yvert (LBMC, ENS Lyon), we have recently developed a Bayesian inference scheme that allows to infer the parameters of a dynamical network model from time-course expression and mRNA data. The scheme is based on an advanced Monte-Carlo Markov Chain (MCMC) algorithm that is intrinsically parallelizable. In this internship, we aim to improve the algorithm, to implement it on different platforms, to test its efficiency and to develop a modular user-friendly interface that allows researchers with similar problems to use the code on their model of interest and with their own data.

mathematical model of a gene network + experiments → network properties $(k_1, \alpha_1, \beta_1, \mu_1, ...)$
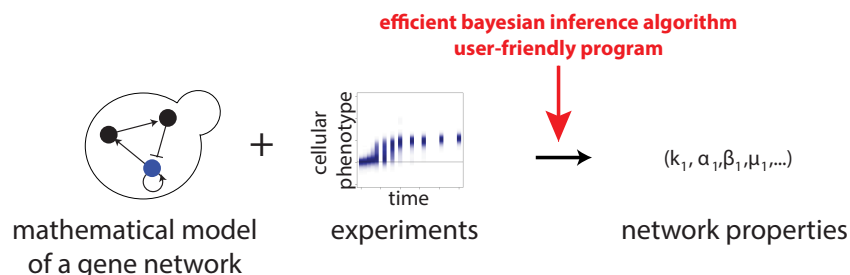
Expected results: The internship will lead to the development of an open-access efficient parallelized program that infer the distribution of the most-likely parameters from a given set of experimental data and for a given user-defined dynamical model.