

T-61.5110 Modeling biological networks

Exam, December 16, 2011

You are NOT allowed to use calculators or any other additional equipments/material in the exam. Please write your answers in English. Please write carefully.

1. Network summary statistics can be used to compare large biological networks. Explain your favorite global network statistic (e.g. node degree, degree distribution, average clustering coefficient, average diameter, etc. but explain only one!) and use your chosen statistic to compare the two networks shown below. (6 points)

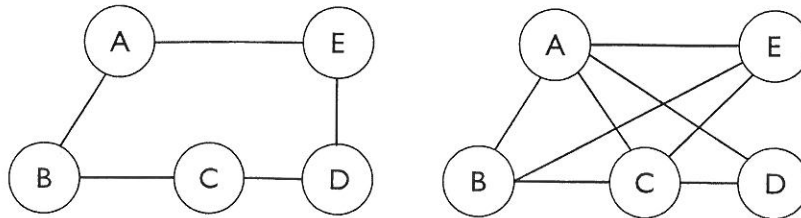


Figure 1: Two networks/undirected graphs.

2. Assume you are given a PSFM model θ for a DNA binding protein A . Assume the width of the PSFM (and thereby the width of the binding site/motif) is n .
 - a) You are given two DNA sequences, $X = (x_1 x_2 \dots x_n)$ and $Y = (y_1 y_2 \dots y_n)$, but only one of them contains a binding site for A . Describe how you can computationally decide which one of the sequences, X or Y , contains a binding site for A . (3 points)
 - b) You are given a DNA sequence $Z = (z_1 z_2 \dots z_\ell)$, where $\ell > n$. Explain how you can assess the statistical significance that Z contains a binding site for protein A . (3 points)
3. Explain the Gillespie algorithm for simulating coupled chemical reactions. (You can assume general stochastic rate constants c_1, \dots, c_v and hazard functions $h_1(\mathbf{x}, c_1), \dots, h_v(\mathbf{x}, c_v)$ for all reactions.) In addition to the Gillespie simulation algorithm itself, briefly explain the connection between Gillespie algorithm and the theory of continuous-time Markov processes. (6 points)
4. Explain how the maximum likelihood principle and the steepest descent (i.e., gradient descent) numerical optimization method can be applied to estimate parameters of ordinary differential equation (ODE) models, given a parametric ODE model and time-series measurements. Please explain all steps necessary for the estimation. (6 points)
5. Modeling approaches can be categorized into deterministic (e.g. ODEs) and stochastic (e.g. coupled chemical reactions) modeling methods. Briefly discuss benefits and drawbacks of deterministic and stochastic approaches in the context of simulation and modeling of biological networks. (6 points)