Bioinformatics III

Prof. Dr. Volkhard Helms — Saarland University
Daria Gaidar, Markus Hollander, Duy Nguyen, Thorsten Will — Chair of Computational Biology
Summer Semester 2018

Exercise Sheet 6

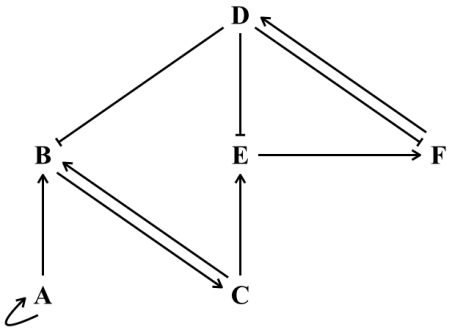
Due: 05.06.2018 10:15

Submit your solutions on paper, hand-written or printed at the *beginning* of the lecture or in building E21, Room 3.03. Alternatively you may send an email with a single PDF attachment. If possible, please include source code listings. Additionally hand in all source code via mail to duy.nguyen@bioinformatik.uni-saarland.de.

Boolean Networks and Differential Expression Analysis

Exercise 6.1: Boolean Networks (50 points)

Consider the following network, which describes the mutual regulation of the hypothetical genes **A** to **F**. A line with an arrowhead denotes an activation while a flat end denotes an inhibition, i.e., if **A** is high, **B** is activated, whereas high levels of **D** inhibit the expression of **B**.



To investigate the behavior of this network use a dynamic simulation as introduced in lecture 12 with a synchronous update scheme. Assume that an activation has a weight of 1, while an inhibition is always 3 times stronger than an activation. Set all threshold to 0.

(a) **Weighted Interactions (10)**

Set up the propagation matrix that relates the states of the genes **A** to **F** in the next iteration to the current state.

(b) **Implementation (20)**

Write a program to simulate the Boolean network.

To enumerate the initial states, convert the binary levels of the genes into an integer where **A** determines the least significant bit and **F** the most significant one. An initital state where, e.g., only **A**, **C**, and **D** are on high levels would translate into 1 + 4 + 8 = 13.

(1) When does it make sense to stop the propagation and why?

(2) Which sequences do you get when you start from states 1, 4, 21, and 33?

(c) **Periodic Orbits (20)**

To determine the attractors and the corresponding basins of attraction, go through all possible initial states and save at which state the Boolean network closes its orbit.

(1) List these orbits with their respective lengths and basins of attraction.

(2) Give the relative coverages of the state space by the basins of attraction.

(d) **Interpretation (10)**

(1) Give the attractors in terms of active genes and characterize them with a few words.

(2) Which are the special genes and what are their respective effects on the behavior of the network? For this, explain what is determining the period of the orbits. Further, compare the two shorter orbits which each other. Which gene is responsible for the difference?

Exercise 6.2: Differential Expression Analysis (50 points)

In this exercise, you will get familiar with differential expression analysis of proteomics data (from Mass Spectrometry experiment) using samr package (R package of SAM method, see lecture 10). More details about samr package can be found on https://cran.r-project.org/web/packages/samr/samr.pdf.
The supplementary file (ms_data.txt) contains intensity values, which exhibit the abundance of the corresponding proteins, of 3 experimental conditions in the first 9 columns: 1 control and 2 siRNA-induced samples (control, rna1 and rna2). Each condition has 3 replicates. Your task is finding which proteins are significantly affected by siRNA-induced knockout (choose 1 out of 2 siRNA samples) by comparing their intensity values in siRNA samples against control samples. The task is composed of the following steps:

(1) Log2-transform intensity values.

(2) Normalise the data using quantile normalisation (using package `preprocessCore`).

(3) Differential expression analysis using `SAM` function of `samr` package.

Play around with various configuration of `fdr.output` (q-value threshold) and `nperms` (number of permutations) to see how they affect the result. List 10 down-regulated and 10 up-regulated proteins and their fold change (Hint: look into the `siggenes.table` component of `SAM` output).

Have fun!