EE550 – Computational Biology

Spring 2022 Final

due 10.6.2022

Instructions This is a take-home test. You are asked to return your answers in a single PDF document with your name and student identification number in the file name. Correct answers to wrong questions are not creditable. Verbal discussions and idea sharing on the test questions are encouraged, but sharing text and/or results in any manner, shape or form is absolutely not and should be avoided at all costs. Your name on your test acts as a statement that you have neither received nor given any unauthorized help. Show your work.

ID	**Comparison**
293088009	skin metastasis vs. all
293088002	local metastasis in the breast vs. all
293088003	local metastasis in the breast vs. skin metastasis
293088006	liver metastasis vs. skin metastasis
293088010	regional lymph node metastasis vs. all
293095016	local metastasis in the breast vs. liver metastasis
283085045	bone metastasis vs. all
283097043	bone metastasis vs. skin metastasis
293085009	liver metastasis vs. all
283085006	local metastasis in the breast vs. regional lymph node metastasis
293085014	liver metastasis vs. regional lymph node metastasis

Question 1 (50 points) For this question, locate the microarray dataset with accession number GDS4761 entitled "Breast cancer metastases from different anatomical sites: fine-needle aspiration biopsies" at the Gene Expression Omnibus website.

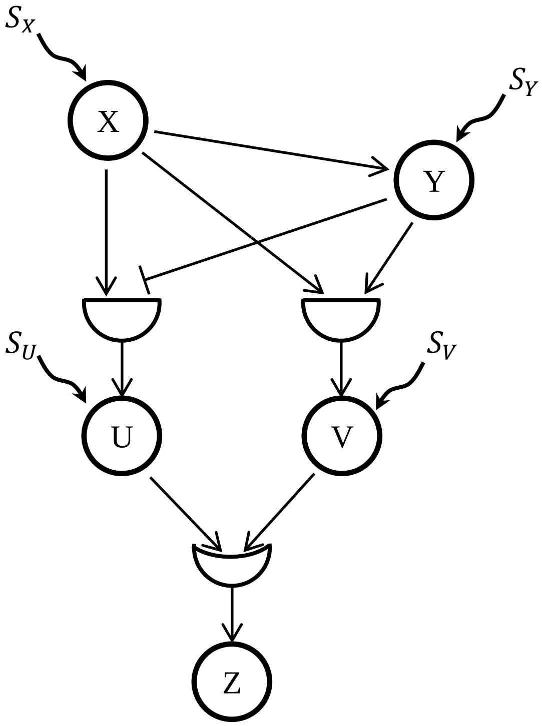
a) (10 points) Briefly describe the aims of the study and the contents of the dataset.

b) (20 points) Identify the genes that are differentially expressed between the two groups assigned to your student identification number in the table above.

c) (20 points) On the KEGG database, identify the pathways that are populated separately by the positively expressed and the negatively expressed genes and discuss the possible effects of metastasis site on breast cancer cells.

Question 2 (50 points) Consider the hypothetical gene transcription network shown in the figure below, where

- $S_X(t) = u(t)$ while $S_Y(t) = S_U(t) = S_V(t) = 1$ for all t, and
- $([X])(0) = X_{st}$ while $([Y])(0) = ([U])(0) = ([V])(0) = ([Z])(0) = 0$.



a) (30 points) Carry out a dynamical evaluation of the gene transcription network to determine how the concentrations of the protein products of genes Y, U, V and Z would change for $t \geq 0$ until steady state.

(Hint: Assume that the activation thresholds are low enough to trigger the designated effect during the dynamic evaluation. Make sure to contrast the transient behavior of the concentrations are contrasted to a simple regulation behavior.)

b) (20 points) Write down the mathematical expressions for the instantaneous production rates of $[Y]$, $[U]$, $[V]$, and $[Z]$ for a potential kinetic modelling of the system.