## Assignment No. 2 Course AI, Spring 2021

# Bayesian Network based Modeling of Gene Regulatory Network of Cancerous Cellular Regulatory Process

**Prepared by Dr. Rauf Ahmed Shams Malick**

**Start time 11:30pm, 5/5/2021, End time 11:30pm, 8/5/2021 Marks 30**

Living cells work as continuous process and can be considered as dynamic machines. There are mainly two fundamental types of proteins in our cells one is regulatory and other is functional. Both types of proteins are formed (transcribed) from genes and their source genes are called regulatory and functional respectively. Regulatory proteins/genes regulates the switching off/on process of functional genes, thereby considered as regulatory/control machinery within cellular process.

This regulatory mechanism from regulatory protein/gene to functional protein/gene can be modeled as a directed graph. A regulatory protein/gene may also regulates other regulatory proteins/genes. At one instance a regulatory protein/gene can be represented as a vertex (source) and a regulated protein/gene can be treated as another vertex (target), can be connected through directed edge.

You have to formulate a graphical probabilistic model for the Gene Regulatory Network for the given data.

**Data**

You are provided two excel sheets including 1. GRN target source Network.xls 2. GRN source target TFs.xls. The sheet containing source target data will be used to build the directed graph at step 1. The resultant directed graph may have cycles, therefore, remove the cycles. Once you have the directed graph prepared then pick any one centrality score from the file 1 i.e. GRN target source Network Analyzer file. Associate the centrality score with the genes/vertices in the directed graph. You need to interpret the centrality score and must associate the meaning with the probabilistic model in terms of cause and effect. In third column, interaction type is provided, activation means switching on whereas repressor means switching off the target gene.

Now you need to build a graphical model, by considering the cascaded effect (can be seen in the provided image).

Answer the following:

Q1. Design and develop a probabilistic formulation, that how probabilities of cause and effect can be measure between interaction of two vertices i.e. regulatory vs regulatory/functional genes. 10 Marks

Q2. Design a cascade based graphical model. Answer how cascade effect can be represented through probabilities and outcome can be inferred through the source as parent and grandparent. 10 Marks

Answer the following queries:

1. Find the probability of switching on of the target gene ERBB2 given that gene AR is up (on). 5 Marks
2. Find the probability of switching off of the gene ATM given that genes ARNT and CTBP1 are off and on respectively. 5 Marks

Show the graphical model and calculate the probabilities.