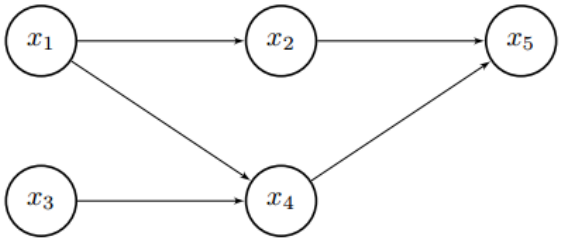
Homework 5

Introduction to Machine Learning
Fall 2018
Instructor: Anna Choromanska

Homework is due 12/14/2018.

Problem 1 (10 points): Bayesian Network Conditional Independence

Consider the Bayesian network below with binary variables representing the following: x_1 student is intelligent, x_2 student is good at taking tests, x_3 student is hard working, x_4 student understands the material, and x_5 student gets good grade.



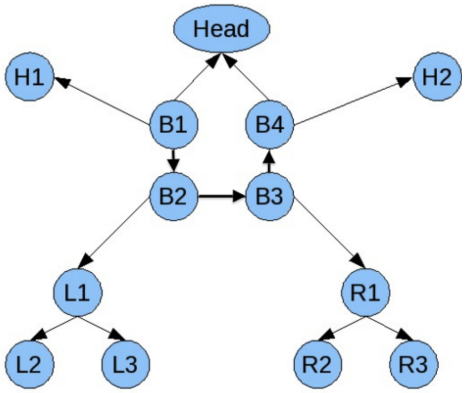
Write out the factorization of the probability distribution $p(x_1, ..., x_5)$ implied by this directed graph. Then, using the Bayes ball algorithm, indicate for each statement below if it is True or False and justify your answers

- x_2 and x_4 are independent.
- x_2 and x_4 are conditionally independent given x_1, x_3, and x_5.
- x_2 and x_4 are conditionally independent given x_1 and x_3.
- x_5 and x_3 are conditionally independent given x_4.
- x_5 and x_3 are conditionally independent given x_1, x_2, and x_4.
- x_1 and x_3 are conditionally independent given x_5.

- x_1 and x_3 are conditionally independent given x_2.
- x_2 and x_3 are independent.
- x_2 and x_3 are conditionally independent given x_5.
- x_2 and x_3 are conditionally independent given x_5 and x_4.

Problem 2 (20 points): Junction Tree

Eve is looking for WallE using her cameras but can't find WallE. Eve has small circuits for performing the junction-tree algorithm. Help her out by designing a junction-tree from the graph below which Eve has in her mind for WallE.



Problem 3 (20 points): Neural Networks

Look at the TensorFlow Playground as given in `http://playground.tensorflow.org/`. For each of the four data sets (3 points (first data set) + 4 points (second data set) + 5 points (third data set) + 8 points (fourth data set)), where each data set consists of data points from 2 different classes (they are shown on the left side of the screen), design an architecture that solves the binary classification problem. Try to design as small network as possible (minimum number of neurons). Play with the: architecture, learning rate, activation functions, and regularization. Use default values for the ratio of train to test data, noise, and batch size. In the final write-up show for each data set one screenshot (the screenshot should show your entire screen: the architecture, parameters, and the obtained output).