Assignment 2: Causal Inference

Giorgia Adorni (giorgia.adorni@usi.ch)

1 Structure of the Network

List at least 10 binary variables all measurable involved in the problem

State which is the objective of the network: for instance, highlight a couple of situations in which decision making could be difficult and in which the graph could provide valuable indications.

Explaining how you decide the arcs orientation, in case they are not self- explaining.

Which arrows can be reversed without being detectable by a statistical test? Explain why.

D-Separation

Identify at least 4 couple of nodes (the node of each couple should be not directly linked to each other) and analyse their d-separation properties possibly conditioning on others.

Discuss how d-connected variables are in fact dependent in the real problem, while d-separated variables are instead independent in the real problem.

2 Conditional Probability Tables

Explain how do you fill the probability tables for the nodes. For instance: a. you have retrieved information from the internet (or other sources); b. you have estimated the CPTs from a database; c. you have relied on your personal experience/common sense.

3 Causal Inference

Choose one pair of variables. The pair must be made up of a variable X with at least one parent and another variable Y of the graph such that there is (at least) a causal path from X to Y. For the pair (X,Y) perform: • Calculate the causal effect of X on Y. • Identify possible confounders between X and Y. • Would it be practically possible in your specific problem to perform also a randomized controlled study to disentangle the causal effect between the variables from their correlation? • Compute the ACE of X on Y. Choose another pair of variable (X,Y) (it can be also the previous one) and: • Choose another variable C such that it is possible to calculate the c-specific effect of X on Y and calculate it. • Identify a minimal set of variables that must be measured in order to estimate the c-specific effect of X on Y. • Choose a function g and compute the effect of the conditional

intervention of X=g(C) on Y. Choose another pair of variable (X,Y) (it can be also the previous one) and: • Identify possible mediating variables between X and Y and calculate the CDE of Y changing the value of X.

4   Simulation

Suppose that you can't measure some parents of variable X chosen in every point of "Causal Inference". Repeat the "Causal Inference" part of the exercise considering this new situation.

5   Comment on the Results

What kind of experience have you got with this model? E.g., is the causal model responding in a sensible way to your queries? What should be changed/modified to make it more realistic?