Medical Data Science, WS 2023/2024

Prof. Dr. Nico Pfeifer

05.12.2023

Oğuz Ata Çal 6661014

Edward Beach 5451904

Assignment 4

Problem 1

(a) Bayesian networks are generative models because you can generate new data from the distribution from the fitted Bayesian network. Dependencies are represented by edges between the nodes. The nodes are random variables.

(b) **a)** False. Because when you find a single path between A to B that is active, you can say that A and B are not d-separated. You only need to check all paths if all of them are blocked and A and B are d-separated. So, you don't always need to check all paths between A and B.

b) True. Because if the path exists from A to B, then it also exists from B to A since we are checking undirected paths. Furthermore, the conditions for the nodes on the path are also the same regardless of the direction of A to B or B to A. This is because head-to-head nodes are always head-to-head nodes regardless of direction and the head-to-tail / tail-to-head nodes are also independent of direction. So, if A is d-separated from B then B is also d-separated from A by C.

c) True. If A and B are d-separated and B and C are d-separated, then there is no active path between A and B and no active path between B and C. Since there is no active path between B and C, the path between A and C must be blocked. So, A and C are d-separated

(c)

- $A \perp\!\!\!\perp B$ is not d-separated so they are dependent. Number of paths checked 1. (A,B)
- $A \perp\!\!\!\perp G|C,E,B$ is d-separated so independent. Number of paths checked 2. (A,B) and (A,C) Because all paths start with either A,B or A,C
- $D \perp\!\!\!\perp C|E,G$ is not d-separated so they are dependent. Number of paths checked 1. (D,F,G,C)
- $G \perp\!\!\!\perp B|A,E$ is not d-separated so they are dependent. Number of paths checked 3. (G,C,E,B) (G,C,A,B) (G,E,B)

(d) **P(A|B)** represents the probability of A given that we observe B in the world, while **P(A|do(B))** represents the probability of A given that we actively intervene on B. This means that we actively set the value of B to a particular value, rather than observing it in the world.