Assignment 4: Bayesian Networks

Due Wednesday, 5 January 2022, 11pm

Design a small Bayesian network modeling a domain of your own choice, and run a few tests on it using SamIam.[1]

Domain Find a domain you find interesting. Here are some suggestions: diagnosing some malfunction (car, computer, application), medical diagnostics (causes and symptoms), plot prediction for TV shows, e-mail spam classifier, or weather forecast. Model this domain using 5–8 variables (nodes), with at least three layers of nodes. Supply all the values in conditional probability tables (CPTs).

You should adhere to the following design guidelines:

- Your network structure should reflect causal and probabilistic dependencies.

- Your CPT values should be reasonable (not randomly chosen). If you are finding it hard to come up with intuitive conditional probabilities for a certain node, it might be an indicator of wrong network structure, such as a symptom being a parent of a cause rather than the cause being the parent of the symptom.

- You should assume that your Bayesian network will be implemented in a real-life application. Therefore, you should to be able to explain how one could determine the CPT values in practice (e.g., some values could be provided by a domain expert and some could be estimated from data).

- When adding a variable, think of the role of that variable: is it an *evidence* variable (something we can observe), a *query* variable (something we want to know), or neither (something that helps us build a good network)?

Test cases Find two test cases for your network, that involve fixing the values of some evidence variables and computing (using SamIam) posterior probabilities of a few other variables of interest. One of these test cases should exhibit the explaining away phenomenon (i.e., knowing a symptom makes the increased probability of one of its causes lower the probabilities of the other causes).

Submit a pdf copy of the following:

- A diagram of the network and the CPT tables.

- A clear description of the network and what the domain it is modeling. Make sure that you describe all variables in the network, their types, what they represent, and the values they can be assigned (i.e., what does each value represent).

- Description of the two test cases, and the posterior probabilities you computed. Include an intuitive justification of the results, and identify the explaining away phenomenon.

Demost Make a demo of your solution on 9 January. The time and places will be announced at SUCourse+.

---

[1]http://reasoning.cs.ucla.edu/samiam/