CPSC 322, Practice Exercise
Solutions to Conditional Independence

---

1 Directed Questions

- Describe the difference between *marginal* independence and *conditional* independence.

  Answer: For variables X and Y, if $P(X|Y) = P(X)$ then X is marginally independent of Y. That is, knowing the value of Y doesn't change the probability of X. For conditional independence, if $P(X|Y,Z) = P(X|Z)$ then X is conditionally independent of Y given Z. In other words, if you know the value of Z, knowing the value of Y doesn't tell you anything more about X.

- Why are we more often interested in conditional independence?

  Answer: Conditional independence is more common - complete independence between variables is relatively rare in a given domain. Also, in any given domain, we typically know *something*, and independence queries should be conditional on that something.

- What are the three key components of a belief network?

  Answer: 1. A directed acyclic graph, with each node a random variable. 2. A domain for each of the random variables. 3. A set of conditional probability distributions giving $P(X|parents(X))$ for each variable X.

- What is the independence assumption in a belief network?

  Answer: Each random variable is conditionally independent of its non-descendants given its parents.

2 Belief Networks and Conditional Independence

A bank has asked us to write a program that allows customers to check on the status of their credit card accounts. In particular, customers whose cards have been disabled should be able to use the software to determine why the card was disabled, and to assess which activities will likely lead to their card being disabled.

A card can become disabled when the account is over its limit, when the card has expired, or when a fraud alert has occurred. Fraud alerts often occur when a card is used internationally or when the card is used at a new vendor, e.g. a store where the customer has never shopped before.

This problem is represented in Figure **??**. All of the variables are boolean. Note that *newVendor* is dependent on *internationalActivity*, as the probability of the vendor being new is much higher when the vendor is international.

- Is *cardholderContacted* conditionally independent of *internationalActivity* given *fraudAlert*? How can you tell from the belief network? **Answer:** Yes, because *internationalActivity* is a non-descendant of *contactCardholer* and *fraudAlert* is a parent of *contactCardholder*. We know that a variable is conditionally independent of its non-descendants given its parents.

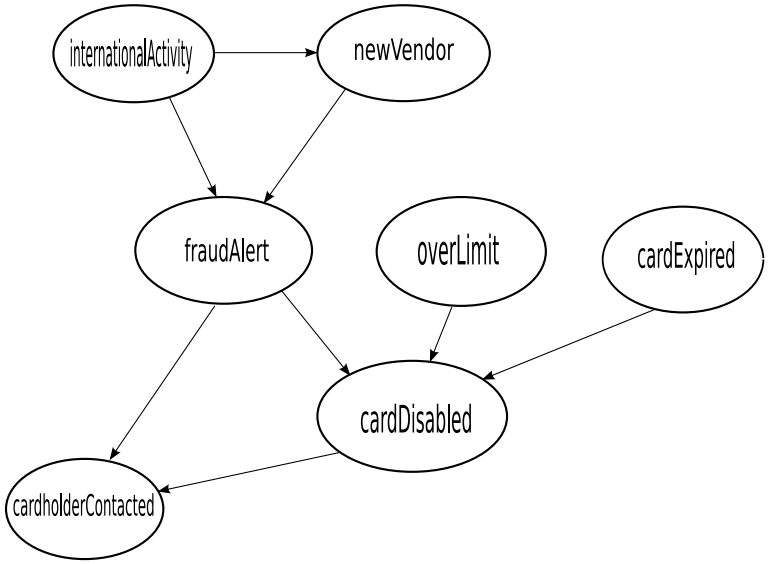

Figure 1: Belief Network for Credit Card Fraud

- Is *fraudAlert* conditionally independent of *newVendor* given *internationalActivity*? How can you tell from the belief network? **Answer:** No, it is not; *newVendor* is a parent of *fraudAlert*.

- In AISpace, load the associated xml file `credit_card_fraud.xml` using the 'Belief and Decision Networks' applet at `http://aispace.org/bayes/`. Examine the probability tables for this problem. Let's say you are travelling in Europe and have used your card there several times, and then the card stops working. You check your (perfectly reliable) voicemail and there is no message from the bank. What is the most likely cause for the card being disabled? In other words, what are the probabilities $P(overLimit = true)$,$P(fraudAlert = true)$ and $P(cardExpired = true)$ given the observations?

  Answer: You have three observations, $internationalActivity = true$, $cardholderContacted = false$, and $cardDisabled = true$. If we query $fraudAlert$, we see that $P(fraudAlert = true) = 0$, because you were not contacted and the bank always contacts when there is a fraud alert. We also find that $P(overLimit = true) = 0.71$ and $P(cardExpired = true) = 0.34$. So the most likely cause for the card being disabled is that you went over your credit limit.

- Consider a separate scenario. You are preparing for an overseas trip where you plan to use your credit card, and you want to know how likely it is that using your card abroad will lead to your card being disabled. If it's a high probability, you should inform your credit card company of your travel plans in advance. Calculate the probability of your card being disabled.

  Answer: We observe that $internationalActivity = true$ since that is the travel plan. The probability of the card being disabled given this observation is approximately 0.95, so it is best to inform the credit card company of your plans.

- Spend some time doing conditional independence quizzes in this applet (the right-most button in 'Solve' mode). This is a great way of getting familiar with conditional independence. You can use the credit card fraud network or any other network for the independence quiz.

  Spend some additional time playing around with this problem in AISpace, making observations and querying variables. The goal for now is to understand basic belief net structure and to have a good grasp of conditional independence. In the next practice exercise we'll go over the algorithm for making these inferences (variable elimination) in detail.

3 Learning Goals

You can:

- Define and use marginal independence
- Define and use conditional independence
- Build a belief network for a simple domain