Report for Lab 3 in TDDC17, Artificial Intelligence
Bayesian Networks

Margareta Vi
marvi154

Markus Petersson
marpe163

October 5, 2016

2.5

a) What is the risk of melt-down in the power plant during a day if no observations have been made? What if there is icy weather?

$$P(\text{Meltdown | no observation}) = 0.02578$$
$$P(\text{Meltdown | icy weather }) = 0.03472$$

b) Suppose that both warning sensors indicate failure. What is the risk of a meltdown in that case? Compare this result with the risk of a melt-down when there is an actual pump failure and water leak. What is the difference?

If both sensors indicates true then

$$P(\text{Meltdown | both warning sensors false}) = 0.14535$$

If there actually is an pump failure and water leak then

$$P(\text{Meltdown | pump failure and water leak}) = 0.20000$$

c) The conditional probabilities for the stochastic variables are often estimated by repeated experiments or observations. Why is it sometimes very difficult to get accurate numbers for these? What conditional probabilities in the model of the plant do you think are difficult or impossible to estimate?

For example it is hard to observe several meltdowns from the same power plant at the same. We think that the conditional probability of meltdown is impossible to estimate. The conditional probability is also hard to estimate.

d) Assume that the "IcyWeather" variable is changed to a more accurate "Temperature" variable instead (don't change your model). What are the different alternatives for the domain of this variable? What will happen with the probability distribution of P(WaterLeak | Temperature) in each alternative?

The domains could be $\{T < 10, \quad -10 < T \leq -5, \quad -5 < T \leq 0, \quad 0 < T < 10, \quad T > 10\}$. The probability P(WaterLeak | Temperature) could increase for $\{T < 10, \quad -10 < T \leq -5, \quad -5 < T \leq 0\}$ since colder weather could result in higher probability of breakage of the water pipes. For temperatures above zero the probability ought to be the same for all reasonable (naturally occurring) temperatures, since it previously was the icy weather increasing the propability of a water leak.

2.6

a) What does a probability table in a Bayesian network represent?
It represents the conditional probabilities. I.e. $P(A|B)$, where A and B are events.

b) What is a joint probability distribution? Using the chain rule on the structure of the Bayesian network to rewrite the joint distribution as a product of P(child|parent) expressions, calculate manually the particular entry in the joint distribution of P(Meltdown=F, PumpFailureWarning=F, PumpFailure=F, WaterLeakWaring=F, WaterLeak=F, IcyWeather=F). Is this a common

state for the nuclear plant to be in?

A joint probability distribution is the probability that a collection of Stochastic variables take certain values at the same time, without any prior observations on any of them.

$$P(\neg MD, \neg PF, \neg WL, \neg WLW, \neg PFW \neg IW) =$$
$$= P(\neg MD|\neg PF, \neg WL)P(\neg PFW|\neg PF)P(\neg PF)P(\neg WLW|\neg WL)P(\neg WL|\neg IW)P(\neg IW) \approx$$
$$\approx 0.6938.$$

This is a common state for the nuclear power plant to be in.

c) What is the probability of a meltdown if you know that there is both a water leak and a pump failure? Would knowing the state of any other variable matter? Explain your reasoning!

If there actually is an pump failure and water leak then the probability is:

$$P(\text{Meltdown | pump failure and water leak}) = 0.20000.$$

No. If we know that there is both a pump failure and a water leak, knowing states of other variables would not matter, since the probability for meltdown is only dependent on pump failure and water leak.

d) Calculate manually the probability of a meltdown when you happen to know that PumpFailureWarning=F, WaterLeak=F, WaterLeakWarning=F and IcyWeather=F but you are not really sure about a pump failure. Hint: Use the Exact Inference formula near the end of the slides, or in sec. 14.4.1 in the book. This formula includes both conditioning on the variables you know (evidence) and marginalizing (summing) over the variable(s) you do not know (often called unobserved or hidden). You need to calculate this both for P(Meltdown=T|...) and P(Meltdown=F|..) and normalize them so that they sum to 1. This normalization factor is the alpha symbol in the equation. With this formula you could answer any query in the network, though it will be impractical for cases with many unobserved variables. A suggestion is to move the terms that do not involve the pump failure variable out of the sum over the two states pump failure can be in (T/F). You may use inference in the applet for verification purposes, but small differences is expected due to rounding errors.

The wanted probability is given by

$$P(MD|\neg WL, \neg PFW) = \sum_{PF} P(MD|\neg WL, PF)P(PF|\neg PFW) =$$

$$= P(MD|\neg WL, PF)P(PF|\neg PFW) + P(MD|\neg WL, \neg PF)P(\neg PF|\neg PFW),$$

where

$$P(PF|\neg PFW) = \frac{P(PF|\neg PFW)}{P(PF|\neg PFW) + P(\neg PF|\neg PFW)} \approx 0.01156,$$

which is due to Bayes' theorem. With

$$P(\neg PF|\neg PFW) = 1 - P(PF|\neg PFW)$$

we the get that

$$P(MD|\neg WL, \neg PFW) \approx 0.00272$$

3.2

During the lunch break, the owner tries to show off for his employees by demonstrating the many features of his car stereo. To everyone's disappointment, it doesn't work. How did the owner's chances of surviving the day change after this observation?

$$P(\text{survival | no observation }) = 0.99561$$
$$P(\text{survival | radio failure }) = 0.99343$$

The owner buys a new bicycle that he brings to work every day. How does the bicycle change the owner's chances of survival?

$$P(\text{survival | no observation on bike}) = 0.99525$$
$$P(\text{survival | bike works}) = 0.99561$$

It is possible to model any function in propositional logic with Bayesian Networks. What does this fact say about the complexity of exact inference in Bayesian Networks? What alternatives are there to exact inference?

The complexity of exact inference depends strongly on the structure of the network. If you have a singly connected network then the time and space complexity of this network is linear to the size of the network.
Instead of doing an exact inference you can do an approximate inference.

4.2

The owner had an idea that instead of employing a safety person, to replace the pump with a better one. Is it possible, in your model, to compensate for the lack of Mr H.S.'s expertise with a better pump?

In our model survival rate is pretty high, 0.96, though it could be higher so it is possible to replace the safety person with a better pump.

Mr H.S. fell asleep on one of the plant's couches. When he wakes up he hears someone scream: "There is one or more warning signals beeping in your control room!". Mr H.S. realizes that he does not have time to fix the error before it is to late (we can assume that he wasn't in the control room at all). What is the chance of survival for Mr H.S. if he has a car with the same properties as the owner? Hint: This question involves a disjunction (A or B) which can not be answered by querying the network as is. How could you answer such questions? Maybe something could be added or modified in the network.

This probability of survival for Mr H.S. includes that Mr H.S has a bike similar to the one that the owner has.

$$P(\text{survival | one or more warning sounds and not in the control room}) = 0.96408$$

What unrealistic assumptions do you make when creating a Bayesian Network model of a person? Describe how you would model a more dynamic world where for example the "IcyWeather"

is more likely to be true the next day if it was true the day before. You only have to consider a limited sequence of days.

A person is hard to make a correct model for, since there are so many underlying factors that affects each state. So each probability we used in our model is a simplification of what could affect each state. If it was *icyweather* the day before then the probability of it being *icyweather* the next day also should increase due to the weather being cold. So we would model it like this: the more days in a row the weather is being *icyweather* the higher the probability it is to be *icyweather* as well (for a limited sequence of days).