TDDC17 - Artificial Intelligence

Lab 3 Report

Group Members: Alim Amanzholov (aliam864)

Part 2. Inference in an existing Bayesian network

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(Meltdown) = 0.02578, P(Meltdown | IcyWeather) = 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 are an actual pump failure and water leak. What is the difference? The answers must be expressed as conditional probabilities of the observed variables, P(Meltdown|...).**

P(Meltdown | PumpFailureWarning, WaterLeakWarning) = 0.14535
P(Meltdown | PumpFailure, WaterLeak) = 0.20000

The difference is P(Meltdown | PumpFailure, WaterLeak) higher than P(Meltdown | PumpFailureWarning, WaterLeakWarning) because when there are actual pump failure and water leak the probability of meltdown should be higher than when there are just warnings.

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?**

It is difficult to get accurate numbers for conditional probabilities when the scenario is difficult to simulate to make observations. In the model, for example, P(Meltdown | Waterleak) would be difficult to estimate because we do not want to make it happen in the real world just to make observations.

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?**

If the “IcyWeather” is changed to the “Temperature” variable, we could have a continuous domain of temperatures or just divide temperature into different ranges (for example, hot, warm, cold). If we divide the variable into ranges, we would have a bigger probability table for P(WaterLeak | Temperature) and in case of the continuous domain, we would have a function that maps temperature to probability instead of a probability table.

6.

a) **What does a probability table in a Bayesian network represent?**

The probability table in a Bayesian network represents the conditional probabilities of possible states for a node given the states of its parent nodes.

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?**

The joint probability distribution is a probability distribution that gives the probability that each of the random variables falls in any particular range or discrete set of values specified for that variable. Bayesian Network is a representation of joint probability distribution.

P(Meltdown=F, PumpFailureWarning=F, PumpFailure=F, WaterLeakWaring=F, WaterLeak=F, IcyWeather=F)
=
P(Meltdown=F | PumpFailureWarning=F, PumpFailure=F, WaterLeakWaring=F, WaterLeak=F, IcyWeather=F)
* P(PumpFailureWarning=F | PumpFailure=F, WaterLeakWaring=F, WaterLeak=F, IcyWeather=F)
* P(PumpFailure=F | WaterLeakWaring=F, WaterLeak=F, IcyWeather=F)
* P(WaterLeakWaring=F | WaterLeak=F, IcyWeather=F)
* P(WaterLeak=F | IcyWeather=F)
* P(IcyWeather=F)

Using the semantics of the Bayesian Network we can rewrite the above equation as follows:

P(Meltdown=F, PumpFailureWarning=F, PumpFailure=F, WaterLeakWaring=F, WaterLeak=F, IcyWeather=F)
=
P(Meltdown=F | PumpFailure=F, WaterLeak=F)
* P(PumpFailureWarning=F | PumpFailure=F)
* P(WaterLeakWaring=F | WaterLeak=F)
* P(WaterLeak=F | IcyWeather=F)
* P(IcyWeather=F)
*P(PumpFailure=F)

= 0.999 * 0.95 * 0.95 * 0.9 * 0.95 * 0.9 = 0.69378

It is a common state for the nuclear 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!**

P(Meltdown | PumpFailure, WaterLeak) = 0.20000

Since in the Bayesian network, there is a direct relation between Meltdown and only PumpFailure and WaterLeak, knowing the state of any other variable does not matter.

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.**

P(Meltdown | PumpFailureWarning=F, WaterLeak=F, WaterLeakWarning=F, IcyWeather=F) =

alpha * P(Meltdown, PumpFailureWarning=F, WaterLeak=F, WaterLeakWarning=F, IcyWeather=F) =

alpha * $\Sigma(over\ y)$ (P(Meltdown, PumpFailureWarning=F, WaterLeak=F, WaterLeakWarning=F, IcyWeather=F, PumpFailure=**y**)) =

alpha * $\Sigma(over\ y)$ ((P(Meltdown | WaterLeak=F, PumpFailure=**y**) * P(PumpFailureWarning=F | PumpFailure=**y**) * P(WaterLeakWarning=F | WaterLeak=F) * P(WaterLeak=F | IcyWeather=F) * P(IcyWeather=F) * P(PumpFailure=**y**)) =

alpha * (0.001*0.95*0.95*0.9*0.95*0.9 + 0.15*0.1*0.95*0.9*0.95*0.1, 0.999*0.95*0.95*0.9*0.95*0.9 + 0.85*0.1*0.95*0.9*0.95*0.1) =

alpha * (0.0019, 0.7007);

alpha = 1 / (0.0019 + 0.7007) = 1.4233

=> alpha * (0.0019, 0.7007) = (0.0027, 0.9973)
Therefore there is **0.27%** chance of Meltdown when calculated by Exact Inference.

(There is 0.272% chance of Meltdown by inference in the applet)

Part 3. Extending a network

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?**

  Initially, P(Survives) = 0.99001, after his radio does not work,

  P(Survives | Radio=F) = 0.98116

- **The owner buys a new bicycle that he brings to work every day. The bicycle has the following properties:**
  - **P(bicycle_works) = 0.9**
  - **P(survives | ¬moves ∧ melt-down ∧ bicycle_works) = 0.6**
  - **P(survives | moves ∧ melt-down ∧ bicycle_works) = 0.9**

  How does the bicycle change the owner's chances of survival?

  Before having a bicycle, P(Survives) = 0.99001. After bicycle, P(Survives) = 0.99505

- **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?**

  It is possible to model any function in propositional logic with Bayesian Networks, which means that the complexity of exact inference in Bayesian Networks is exponential (i.e. NP-hard problem). Approximation algorithms can be used as an alternative to exact inference.

Part 4. More extensions

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?**

  Yes, it is possible to compensate for the lack of Mr H.S.'s expertise with a better pump in our model. If better pump is used probability of PumpFailure being true will decrease, which increases the survival chance of the owner. In our model, because Mr H.S. is quite incompetent, he adds very little value to the survival of the owner.

- **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?

To capture the disjunction, we need to add one more node to the network. Namely, we need to add AnyWarning node which will be true if either PumpFailureWarning or WaterLeakWarning is true. After adding this node and making an observation that it is true:

P(Survives | AnyWarning=T) = 0.98318

- **What unrealistic assumptions do you make when creating a Bayesian Network model of a person?**

  It is unrealistic to confine the behavior of person to 4 variables. Furthermore, the probability of variables that model a person can change from day to day depending on external factors. For example, isAsleep might depend on whether the person had enough sleep the night before and so on.

- **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.**

  We can model a more dynamic world by adding extra nodes and connecting it to existing nodes. For example, in the case of IcyWeather, we can add WasIcyWeatherYesterday node that as the name suggests indicates whether there was an icy weather the day before. We can then connect this new node to the IcyWeather node. If we want to have bigger sequence of days we can just create new nodes for that days and connect it to the next day and so on.