**Tasks for part II**

The owner asks you to start experimenting with the network and find out how safe the plant is.

1. Start the applet and choose "*Edit/ViewEdit Text representation (.xml format)*". A window with the XML representation of the current (empty) network will pop up. This window can be used to edit or save the current network.
2. Copy and paste the network part of the file [bayesian\_net/nuclear\_power\_station.xml](http://www.ida.liu.se/%7ETDDC17/info/labs/bayesian_net/nuclear_power_station.xml) into the corresponding network tag in the empty network.
3. Press the "*Update*" button. You should now see that a network has been created in the original applet window.
4. Set the "*Network Options / Decimal Places for Monitoring*" to 5. This means that you will be able to see small changes in probabilities.
5. Use the applet and the loaded Bayesian Network to answer the following questions:
   * 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?
     1. If no observations have been done the monitoring shows 0.02578 probability for a meltdown
     2. If the weather has the added observation “Icy weather=true” the risk for a meltdown increases to 0.03472
   * b) Suppose that both warning sensors indicate failure. What is the risk of a melt-down 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? (The answer must be expressed in conditional probabilities)
     1. PumpFailureWarning=True; WaterLeakWarning=True  
        The risk for meltdown is 0.14535
     2. PumpFailure=True; WaterLeak=True;  
        Risk for meltdown is 0.20000
     3. The difference is 0.05465
   * 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 probabilites in the model of the plant do you think are difficult or impossible to estimate?
     1. The more rare or volatile a state is, the harder it is to get accurate statistics for it, or in other words, the more observations the smaller the standard deviation becomes.
        1. IceWeather: Reasonably predictable considering the level of accuracy in weather prognosis. (This is by itself a great accomplishment, but due to a lot of research in the subject of weather science we take fairly accurate predictions for granted)
        2. Water Leak/Pump failers: This would be fairly difficulty to predict. Stress tests outside the system might aid in assessing the probabilities, but it might be generally difficult to decide anything with any certainty when the unit is added to the system.
        3. PumpFailureWarning/WaterLeakWarning: The accuracy of these two should be fairly easy to predict, as repeated stress tests and simulated failures on the equipment they are connected to could easily be repeated.
        4. Meltdown: This must be considered as a difficult state to predict. The repeatability is very low, and this system is very dependent on a lot of other systems, thus adding to the complexity and in turn affecting the predictability negatively.
   * 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?
     1. Icy Weather most likely affects the water-leaks by water reaching a specific temperature, thus making the water freeze. If we assume that icy weather doesn’t always indicate that water is frozen, while water being frozen would always indicate icy weather, the change from icy weather 🡪 temperature would make reduce the probability of water leaks occurring; the result would be this:

|  |  |  |
| --- | --- | --- |
| Icy Weather | P(WaterLeak=T) | P(WaterLeak=F) |
| T | Increases | Decreases |
| F | Decreases | Increases |

1. To guide your understanding of how Bayesian networks work we provide a few theory questions below that should be answered using the lecture slides and/or book. When asked to calculate something manually **show your calculations**. You may use the applet for verification where possible.
   * a) What does a probability table in a Bayesian network represent?
   * 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?
   * 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!
   * 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). Small differences compared to the applet is expected due to rounding errors.

**Part III: Extending a network**

**Scenario**

The owner of the nuclear plant is quite selfish and wants to optimize profit of the plant at the expense of safety, yet he is very worried about his future survival in case a problem with the plant arises. Instead of increasing the safety of the plant which is costly, he decides to analyse his chances of escaping from the plant in case of a melt-down. As an apprentice to the owner, the owner wants you to investigate the properties of his escape vehicle (his car). Your model of the car is the same as in exercise 14.7, figure 14.21 (p. 560) in the course book with the extension of the "*IcyWeather*" variable, which of course is the same as we already have in our model of the plant. After a year of observations and subjective assumptions you come up with the following conditional and prior probabilities:

* P(battery | icyWeather) = 0.8
* P(battery | ¬icyWeather) = 0.95
* P(radio | battery) = 0.95
* P(ignition | battery) = 0.95
* P(gas) = 0.95
* P(starts | gas ∧ ignition) = 0.95
* P(moves | starts) = 0.95
* P(survives | moves ∧ melt-down) = 0.8
* P(survives | moves ∧ ¬melt-down) = P(survives | ¬moves ∧ ¬melt-down) = 1.0
* P(survives | ¬moves ∧ melt-down) = 0.0

Fill in the rest of the probabilities by using common sense reasoning about the domain.

**Task for part III**

1. Model the car with the applet tool and integrate it with the model of the plant.
2. Answer the following questions:
   * 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?
   * 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?

* + 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?

**Part IV: More extensions**

**Scenario**

After your excellent analysis of the plant's Bayesian Network and the creation of the model of the owner's car, he realizes that he is still in great danger. He comes to the conclusion that he needs to hire someone who is in charge of the plant's safety. After some job interviews he decides that a Mr H.S. is the most suitable person for the job because he practically works for free. But Mr H.S. has some less appealing properties:

* He sleeps a lot during work which means that he can not react that rapidly to warning signals and so on.
* He is very incompetent. Even if he is awake it doesn't mean that he knows what to do in case of an emergency.

You volunteer for the task of modeling Mr H.S.'s behavior with the help of the "Bayes Applet" tool.

**Task for part IV**

1. Using your own modeling creativity, create a model of Mr H.S. with the Applet tool. The requirements of the model are the following:
   * It must be integrated with the model of the nuclear plant and must depend on the "WaterLeakWarning" and "PumpFailureWarning" stochastic variables.
   * It must contain at least four stochastic variables.
   * It should match fairly well with the informal description of Mr H.S.'s properties given in the scenario description.
2. Answer the following questions using your model:
   * 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?
   * 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? (notice that this question involves a disjunction which can not be answered by querying the network directly)
   * 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.

**How to report**

The examination for this lab will consist of the following:

1. Demonstrate your networks to the lab assistant. The lab assistants may have some extra questions for you that you have to answer correctly during the demonstration.
2. Write a report where you answer all the questions in part II-IV. In those answers where the applet is used for obtaining the result, you must clearly describe what variables you choose to observe and why. The theory questions in part II should likewise be accompanied by reasoning and the steps of any calculations. [Submission instructions](http://www.ida.liu.se/%7ETDDC17/info/labs.en.shtml).

**Acknowledgements**

This lab has been developed by Per Nyblom.