Knowledge and Reasoning

Licenciatura em Engenharia Informática: 2º ano - 2º semestre

Practical Class nº 3

Bayesian Networks: GENI E

1. Objectives

This work consists of the implementation of some Bayesian Networks (Belief-Nets) designed to exemplify inferences of the causal and diagnostic types.
For its implementation, GENIE is used, a freeware application in the academic field from the University of Pittsburg and now made available by *BayesFusion*.
You can consult the documentation and tutorials at: https://support.bayesfusion.com/docs/GeNIe/

2. Bibliography

Moodle slides
Documentation Genie: https://support.bayesfusion.com/docs/GeNIe/

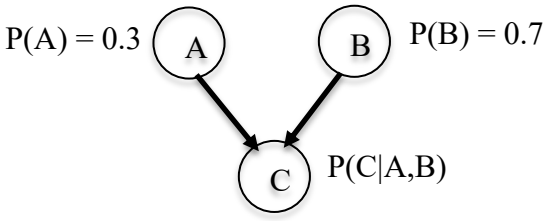
3. Installing Genie

In Moodle is the installation file (only available for Windows).
Installation on MacOS must be done using

- Wine:https://support.bayesfusion.com/docs/GeNIe/introduction_geniemac.html

4. Bayesian Networks

In many real problems, there is no complete information about the environment, either due to failure to obtain data, imprecision in the measuring devices or impossibility in obtaining it. In these situations, there is a need to resort to probabilistic reasoning techniques. Bayesian networks fall into this type of technique and use the theory of probabilities and the theory of Bayes, which considers probability as the degree of certainty of the occurrence of an event.
A Bayesian Network is represented by an acyclic and directed graph, in which the nodes represent the random variables (attributes) and the arcs the dependencies between the attributes. An arc is directed from causes to effects. An attribute A may have no dependencies, in which case it has a probability table of type P (A); If you have dependencies, your probability table will be of type P (C | A, B). See an example, in the following figure:



Designing a Bayesian network involves two phases:

- Design the network
- Estimate the conditional probabilities involved (use histories, experts, etc.)

A Bayesian network allows three types of calculations:

- Calculate the joint probability of any event
- Perform causal inference calculations (from causes to effects)
- Perform diagnostic inference calculations (from effects to causes)

5. Tasks to perform

5.1. Exercise 1: (Tuberculosis)

Create a Bayesian network and save it under the name **Tuberculosis**.
Build the network that models the following situation.
Suppose that in a given country:

1. 1% of the population has tuberculosis
2. An X-ray is positive (indicates the presence of the disease) in 95% of the cases in which the disease was actually contracted
3. An x-ray is interpreted as positive in 0.5% of the cases that, after all, were found not to be actually tuberculosis

Use Genie to answer the following questions:

- In the actual presence of tuberculosis (evidence), what is the probability of a positive radiograph? Is it negative?
    - **Answer:** Positive: 0.95; Negative: 0.05

- In the absence of tuberculosis (evidence), what is the probability of a positive radiograph? Is it negative?
    - **Answer:** Positive: 0.005; Negative: 0.995

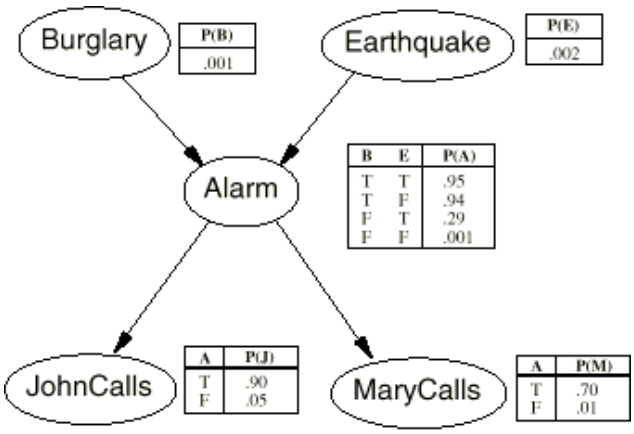
- Radiography is positive (evidence). What is the probability of tuberculosis?
    - **Answer:** 0.6574

NOTES:

1. To indicate that a radiograph is positive or negative, use the **Set Evidence** option on the node you build for the radiographs.
2. To calculate the probability of absent or present tuberculosis, use the **Network Update** option
3. To check the resulting probabilities in the Tuberculosis node, use the **Set Evidence** option in that node, for consultation only.
4. Use **Clear Evidence** to clear the previous evidence for each node

5.2. Exercise 2 - Alarm

A house has an **alarm** that sounds when there are **burglaries**, but sometimes also when there is an **earthquake**. When it rings, neighbors John and Mary **call the owner**, according to the probabilities indicated in the Bayesian network in the figure.



a) Create a new network in Genie and save it with the name **Alarm**
b) Create the Bayesian network and assign the probabilities according to the figure
c) Calculation of **joint probabilities**: use GeNie to calculate the probability that John and Mary both call, the Alarm rings and there is no burglary or earthquake, that is:

P(JohnCalls, MaryCalls, Alarm, ¬Burglary, ¬Earthquake)

- **Select the evidence indicated above**
- **Choose the option "Network – Probability of Evidence".**
- **Must obtain the value 0.0006281**

d) **Causal inference:** What is the probability that Maria will call, knowing that there has been a robbery and the alarm has gone off?

P(MaryCalls | Burglary, Alarm)

- **Clear previous evidence**
- **Select the evidence indicated above**
- **Choose the "Network – Update Beliefs" option.**
- **Should get the value 0.7**

e) **Diagnostic inference**: what is the probability that an earthquake occurred, knowing that Mary and John called and the alarm rang?

P(Earthquake | JohnCalls, MaryCalls, Alarm)

- **Clear previous evidence**
- **Select the evidence indicated above**
- **Choose the "Network – Update Beliefs" option.**
- **You should get the value 0.231**

5.3. Exercise 3 – Heart Disease

The practice of physical exercise and a healthy diet condition the appearance of heart disease. An inadequate diet can be responsible for the appearance of heartburn. Heart disease causes high blood pressure and chest pain. In turn, chest pain can also be caused by heartburn. According to your experience, the following probabilities have been established:

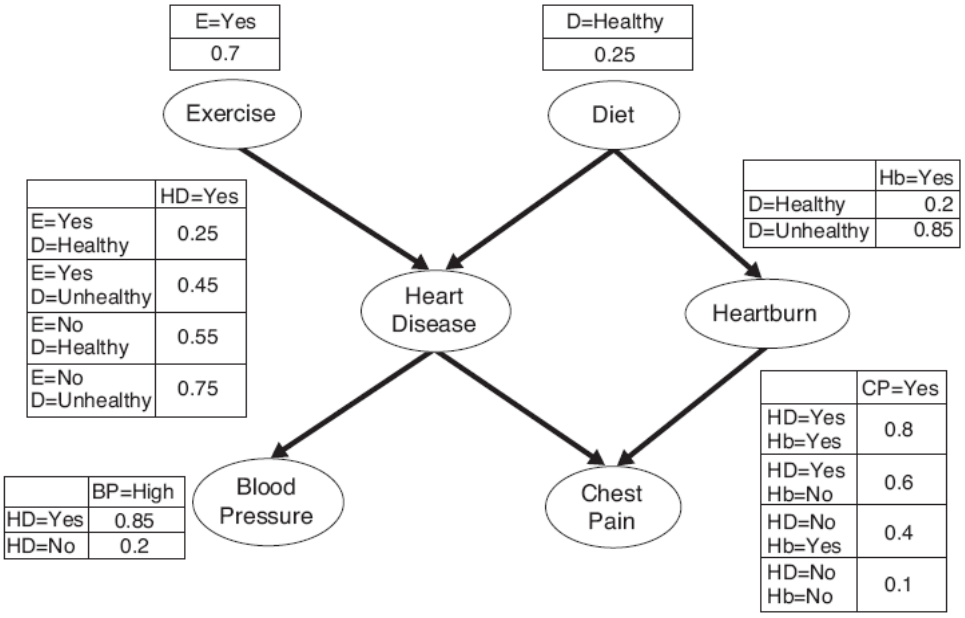


Figure 5.13. A Bayesian belief network for detecting heart disease and heartburn in patients.

a) Create a new network in Genie and save it under the name ***CardiacDisease***
b) Assign the probabilities according to the figure.

c) Use Genie to answer the following questions:

Causal Inference:

- Mary practices exercise. According to the Bayesian network in the figure, what is the probability that Mary is a candidate for heart disease?
  Should get the value 0.4

- Without any additional information, and in the scenario described by this network, what is the probability of a person having heart disease.
  Should get the value 0.49

- What is the probability of a person having high blood pressure, knowing that they exercise and have a healthy diet?
  Should get the value 0.3625

Diagnostic inference:

- What is the probability of a patient having heart disease, knowing that they have high blood pressure, do not exercise or have a healthy diet?
  You should get the value 0.9272

- How likely is it for a patient to have an unhealthy diet if they know they have chest pain and do not have heart disease?
  You should get the value 0.846

5.4 Exercise 4 - Fault Diagnosis

Create a Bayesian network and save it under the name **Faults**.
Build the network that models the following situation:

- An elevator system has 3 common causes of failure: the **control unit**, the **floor sensors** and the **engine**.
- These three failures give rise to three types of evidence: **burning smell**, **stopping on the wrong floor**, **unevenness on arrival**. The frequencies of occurrence, in discussion with the technicians, are:

	Engine OK	Engine OK	Engine Damaged	Engine Damaged
	Control Unit OK	Control Unit Damaged	Control Unit OK	Control Unit Damaged
Presence of Burning smell	Almost never	Sometimes	Frequently	Almost always

	Floor Sensors OK	Floor Sensors OK	Floor Sensors Damaged	Floor Sensors Damaged
	Control Unit OK	Control Unit Damaged	Control Unit OK	Control Unit Damaged
stopping on the wrong floor	Almost never	Frequently	Uncommonly	Almost always

	Floor Sensors OK	Floor Sensors OK	Floor Sensors Damaged	Floor Sensors Damaged
	Control Unit OK	Control Unit Damaged	Control Unit OK	Control Unit Damaged
unevenness on arrival	Almost never	Sometimes	Very often	Almost always

The linguistic terms were translated by the following (conditioned) probabilities:

Almost never	0.05
Sometimes	0.30
Uncommonly	0.60
Frequently	0.70
Very often	0.80
Almost always	0.95

The company has the following summary extracted from the failure history:

Cause	Nº of occurrences
Control Unit	300
Floor Sensors	200
Engine	500

Use Genie to build the network suitable for the fault diagnosis described above.
Try to answer the following questions:

- How likely is it that the engine, the control unit and the sensors will break down and there will be an unevenness on arrival?
- Given the network specifications, what is the probability of a stop on the wrong floor?
- Knowing that there was a burning smell and an unevenness on arrival, what are the probabilities of engine failure?