University of Cagliari

MSc programs in: Computer Engineering, Cybersecurity and Artificial Intelligence – Electronic Engineering – Internet Engineering
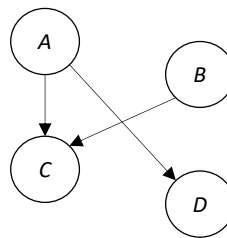

Artificial Intelligence

Academic Year: 2025/2026
Instructor: Giorgio Fumera

Exercises on Bayesian networks

1. Consider the Bayesian network (BN) shown below, whose nodes have been drawn in the following order: A, B, C, D.

   (a) What are the PDFs associated to its nodes? Write the expression of the joint probability distribution function (PDF) represented by the BN.

   (b) What are the independence assumptions encoded by this BN?

   (c) Assuming the variables are all Boolean, derive an expression for the conditional probability

   $$P(A = \text{true}|C = \text{true}, D = \text{true})$$

   using the *exact* inference procedure.

   (d) Describe how the *rejection sampling* algorithm estimates the same conditional probability, making an example of how a single sample is generated.

2. Consider the following Boolean random variables related to the state of a car: *Battery* (it equals *false* if the battery is dead), *Fuel* (it equals *false* if the fuel tank is empty), *Ignition* (it equals *true* if the ignition system works), *Moves* (it equals *true* if the car moves after one tries to start the engine), *Radio* (it equals *true* if the radio works when one tries to switch it on), *Starts* (it equals *true* if the engine fires when one tries to start it).

   (a) Identify causal dependencies between the events represented by these variables, and write the corresponding expression of their joint PDF using the chain rule.

   (b) Make suitable independence assumptions, *clearly motivating them*, and show how they affect the expression of the joint PDF given by the chain rule.

   (c) Draw a BN representing the expression of the joint PDF resulting from the independence assumptions.

   (d) How many probability values need to be estimated to specify the PDFs associated to the nodes of your BN? Can some of these probability values be set based on *a priori* causal knowledge about the corresponding events?

3. Two astronomers, in different parts of the world, look at the same region of the sky using their telescopes and count the number of stars they see. Their counts may be inaccurate for several reasons, including the fact that their telescopes can occasionally be out of focus.

    (a) Define a suitable set of random variables to describe the above domain.

    (b) Identify causal dependencies between them.

    (c) Draw a BN to represent their joint PDF, making suitable independence assumptions.

    (d) Do you see any reasonable constraint that some of the PDFs associated to the nodes of your BN should satisfy, based on *a priori* causal knowledge about the corresponding events?

4. Mary's car has an alarm that sounds when a motion sensor detects someone entering the car. The alarm and the sensor are powered by two distinct batteries, which can be occasionally dead.

    (a) Define a set of random variables to represent the above domain.

    (b) Identify causal dependencies between the events they represent.

    (c) Draw a BN to represent their joint PDF, making suitable independence assumptions.

5. Headache and fever are among the symptoms of several health problems, including influenza and food poisoning.

    (a) Define a probabilistic model of the above domain, introducing a suitable set of random variables, identifying causal dependencies among the corresponding events, and drawing a BN to represent their joint PDF, making suitable independence assumptions.

    (b) Derive an expression of the probability that a person suffering from headache has caught influenza, using the *exact* inference algorithm.

    (c) Describe how the *rejection sampling* algorithm estimates the same probability, making an example of how a single sample is generated.

6. In a nuclear power station an alarm sounds and warning lights flash in the control room, when a sensor detects that the temperature of the core exceeds a given threshold. On rare occasions the sensor measurement may be incorrect, resulting in false positive or false negative detections, especially when the external temperature is very high. Occasionally, also the alarm and the warning lights can fail; to limit joint failures, they are implemented as physically separated systems.

    (a) Define a probabilistic model of the above domain, introducing a suitable set of random variables, identifying causal dependencies among the corresponding events, and drawing a BN to represent their joint PDF, making suitable conditional independence assumptions.

    (b) Derive an expression of the probability of a core overheating when warning lights are flashing in the control room, using the *exact* inference procedure.

    (c) Describe how the *rejection sampling* algorithm estimates the same probability, making an example of how a single sample is generated.

Solution

1. (a) The expression of the joint PDF represented by the BN is:

$$P(D, C, B, A) = P(D|A)P(C|B, A)P(B)P(A) \ . \tag{1}$$

The PDFs in the right-hand side are the ones associated to the nodes of the BN.

(b) The general expression of the chain rule representing the joint PDF, following the same order among the variables, is:

$$P(D, C, B, A) = P(D|C, B, A)P(C|B, A)P(B|A)P(A) \ . \tag{2}$$

Comparing the corresponding PDFs in the right-hand side of expressions (1) and (2), it follows that the independence assumptions encoded by the BN are:

- B is (absolutely) independent of A: $P(B|A) = P(B)$;
- D is conditionally independent of C and B, given A: $P(D|C, B, A) = P(D|A)$.

The same independence assumptions can be directly derived by the structure of the BN, taking into account the order of its nodes:

- there is no edge from A to B, therefore B is (absolutely) independent of A: $P(B|A) = P(B)$;
- there are no edges from C and B to D, therefore D is conditionally independent of C and B, given A: $P(D|C, B, A) = P(D|A)$.

(c) The *exact* inference procedure derives the expression of any conditional probability by first using the definition of conditional probability, then applying the sum rule to its numerator and denominator (with respect to all the other variables), and finally rewriting the full joint probabilities using the chain rule corresponding to the BN:

$$
\begin{aligned}
P(A = \text{t}|C = \text{t}, D = \text{t}) \ &= \ \frac{P(A = \text{t}, C = \text{t}, D = \text{t})}{P(C = \text{t}, D = \text{t})} \\
&= \ \frac{\sum_b P(A = \text{t}, B = b, C = \text{t}, D = \text{t})}{\sum_{a,b} P(A = a, B = b, C = \text{t}, D = \text{t})} \\
&= \ \frac{\sum_b P(D = \text{t}|A = \text{t})P(C = \text{t}|B = b, A = \text{t})P(B = b)P(A = \text{t})}{\sum_{a,b} P(D = \text{t}|A = a)P(C = \text{t}|B = b, A = a)P(B = b)P(A = a)} \ .
\end{aligned}
$$

(d) The *rejection sampling* algorithm first generates a given number (say, N) of samples of all the random variables, according to their joint PDF represented by the BN, following the topological order A, B, C, D. This is an example of how a *single* sample is generated:

- a sample from $P(A)$ is drawn: assume $A = \text{false}$ is obtained;
- a sample from $P(B)$ is drawn: assume $B = \text{false}$ is obtained;
- a sample from $P(C|B = \text{false}, A = \text{false})$ is drawn: assume $C = \text{true}$ is obtained;
- a sample from $P(D|A = \text{false})$ is drawn: assume $D = \text{false}$ is obtained.

The corresponding sample is $A = \text{false}, B = \text{false}, C = \text{true}, D = \text{false}$. Note that it *does not* agree with the evidence $C = \text{true}, D = \text{true}$.

All the samples that do not agree with the evidence are rejected. Let $N' \leq N$ denote the number of the retained samples; among them, let $N'' \leq N'$ denote the number of samples that correspond to the event of interest, $A = \text{true}$. The probability $P(A = \text{true}|C = \text{true}, D = \text{true})$ is finally estimated as N''/N'. Note that this is a *frequentist* estimate of the conditional probability of interest:

$$P(A = \text{t}|C = \text{t}, D = \text{t}) = \frac{P(A = \text{t}, C = \text{t}, D = \text{t})}{P(C = \text{t}, D = \text{t})} \approx \frac{N''/N}{N'/N} = \frac{N''}{N'} \ .$$

2. (a) The state of the battery and of the fuel tank can be considered as the "root causes". It is reasonable to assume they do not directly influence each other. The battery state directly affects the radio and the ignition system. It is also reasonable to assume that the radio and the ignition system do not directly influence each other. The states of the fuel tank and of the ignition system directly determine whether the engine fires or not. The state of the engine directly determines whether the car moves.

Accordingly, the considered random variables can be sorted from the "root causes" to the "end effects" as follows: $\{Fuel, Battery\} \to \{Radio, Ignition\} \to Starts \to Moves$. The order among variables inside curly brackets is not relevant. The following order among all the variables will be considered from now on: $Fuel, Battery, Radio, Ignition, Starts, Moves$. Denoting the random variables with their initial letters only, for the sake of brevity, the general expression of their joint PDF using the chain rule is:

$$P(M,S,I,R,B,F) = P(M|S,I,R,B,F)P(S|I,R,B,F)P(I|R,B,F)P(R|B,F)P(B|F)P(F) . \quad (3)$$

(b) To identify suitable independence assumptions, consider the conditional PDFs in the right-hand side of expression (3), from right to left:

- $P(B|F)$: the state of the battery and of the fuel tank can be considered (adsolutely) independent of each other: $P(B|F) = P(B)$;
- $P(R|B,F)$: the state of the radio is independent of the state of the fuel tank, given the state of the battery: $P(R|B,F) = P(R|B)$;
- $P(I|R,B,F)$: the state of the ignition system is independent of the state of the radio and of the fuel tank, given the state of the battery: $P(I|R,B,F) = P(I|B)$;
- $P(S|I,R,B,F)$: the state of the engine is independent of the state of the radio and of the battery, given the state of the fuel tank and of the ignition system (note that the battery affects the engine only *indirectly*, through the ignition system): $P(S|I,R,B,F) = P(S|I,F)$;
- $P(M|S,I,R,B,F)$: given the state of the engine, whether the car moves or not is independent of all the other factors: $P(M|S,I,R,B,F) = P(M|S)$.

The expression of the joint PDF becomes:

$$P(M,S,I,R,B,F) = P(M|S)P(S|I,F)P(I|B)P(R|B)P(B)P(F) . \quad (4)$$

(c) The BN representing the joint PDF (4) is shown below.
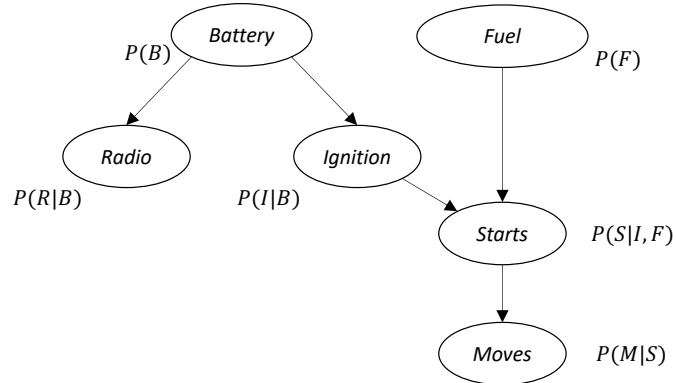
(d) Remember that, to specify the *prior* PDF $P(X)$ of any Boolean random variable X, only one value needs to be estimated: either $P(X = \text{true})$ or $P(X = \text{false})$, since the other value is determined by the constraint $P(X = \text{true}) + P(X = \text{false}) = 1$. Similarly, to specify the *conditional* PDFs $P(X|Y_1,\ldots,Y_n)$, when all variables are Boolean, either the probability that $X = \text{true}$ or the probability that $X = \text{false}$ need to be estimated, for *each* of the 2^n possible combinations of values of $Y_1,\ldots,Y_n$.

Therefore:

- $P(F)$ and $P(B)$ require the estimation of one probability value each, for a total of 2 values;
- $P(R|I)$ requires one value for $I = \text{true}$ and one for $I = \text{false}$; similarly for $P(R|B)$ and $P(M|S)$, for a total of 6 values;
- $P(S|I, F)$ requires one value for each of the four combinations of values of I and F, for a total of 4 values.

The joint PDF of the considered six Boolean variables can therefore be specified through $2+6+4 = 12$ probability values, thanks to the above independence assumptions, instead of $2^6 - 1 = 63$ values.

Some values of the PDFs associated to the BN can be set *a priori*, based on causal knowledge on the corresponding events. In particular, the ignition system and the radio cannot work, if the battery is dead, i.e., $P(I = \text{true}|B = \text{false}) = P(R = \text{true}|B = \text{false}) = 0$. Similarly, if the ignition system does not work or the fuel tank is empty, the engine cannot fire: $P(S = \text{true}|I = \text{false}, F) = 0$, and $P(S = \text{true}|I, F = \text{false}) = 0$.

On the other hand, the ignition system may not work even when the battery is *not* dead, due to several possible causes that one may not know or may not be willing to consider *explicitly*, such as a broken fuel pump or a fault in the ignition system itself. Therefore, $P(I = \text{true}|B = \text{true})$ should *not* be set to 1, to account for *any* other possible cause that prevents the ignition system from working. Similarly, the engine may not fire even when the ignition system *does* work and the fuel tank is *not* empty; therefore, $P(S = \text{true}|I = \text{true}, F = \text{true})$ should be lower than 1.

3. (a) Five random variables can be used to represent the relevant events:

- M_1 and M_2: the number of stars counted by the two astronomers: their domain is the set of natural numbers $\{0, 1, 2, \ldots\}$;
- N, the actual (unknown) number of stars in the region of the sky under observation: its domain is the set of natural numbers;
- F_1 and F_2: Boolean random variables representing whether the two telescopes are out of focus (true) or not (false).

(b) The actual number of stars N and the states of the two telescopes (F_1 and F_2) can be considered as "root causes"; it can also be assumed that they do not directly affect each other. The number of stars N_1 estimated by the first astronomer is directly influenced only by the actual number of stars N and by the state of his or her telescope, F_1, but not by the number of stars estimated by the other astronomer, N_2 (assuming they do not communicate with each other); similarly, N_2 is directly influenced only by N and F_2.

(c) Considering the random variables ordered as F_1, N, F_2, N_1, N_2 (in agreement with the above causal dependencies), the corresponding BN is shown below.
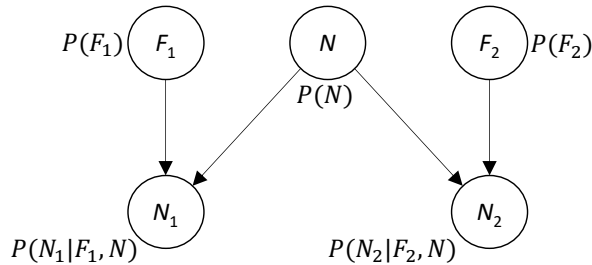
Formally, the independence assumptions made above and encoded by the BN are:

- N and F_1 are (absolutely) independent of each other: $P(N|F_1) = P(N)$;
- F_2 is (absolutely) independent of both N and F_1: $P(F_2|N, F_1) = P(F_2)$;
- N_1 is conditionally independent of F_2, given N and F_1: $P(N_1|F_2, N, F_1) = P(N_1|N, F_1)$;
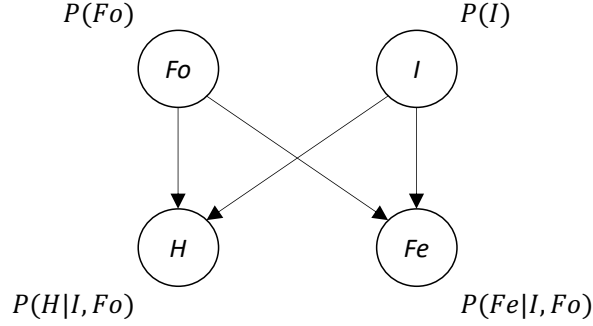- N_2 is conditionally independent of N_1 and F_1, given N and F_2: $P(N_2|N_1, F_2, N, F_1) = P(N_2|F_2, N)$.

(d) The conditional PDFs $P(N_i|N, F_i = \text{false})$, $i = 1, 2$ (i.e., the PDFs of the number of stars estimated by either astronomer when the corresponding telescope is not out of focus, for any given, actual number of stars), should be greater than zero when the estimate is *wrong*, i.e, when $N_i \neq N$. For instance, $P(N_i = 1000|N = 1050, F_i = \text{false}) > 0$. The reason is that the estimated number of stars may be wrong due to other possible causes, besides the telescope being out of focus, that are not explicitly taken into account in this probabilistic model (e.g., the sky may be not perfectly clear when the observation is made), and may be even unknown.

4. (a) This domain can be described by five Boolean random variables denoting the presence of someone inside the car (I), the state (dead or not dead) of the batteries powering the sensor (B_s) and the alarm (B_a), the state (detection or no detection) of the sensor (S) and the state (sounding or not sounding) of the alarm (A).

(b) It is reasonable to assume that the "root causes" correspond to the presence of someone inside the car (I) and to the states of the two batteries (B_s and B_a), and that these events do not directly affect each other. The state of the sensor (S) is directly influenced only by the state of its battery (B_s) and by the presence of someone inside the car (I). The state of the alarm (A) is directly influenced only by the states of its battery (B_a) and of the sensor (S).

(c) Considering the random variables ordered as B_s, I, S, B_a, A (in agreement with the above causal dependencies), the corresponding BN is:
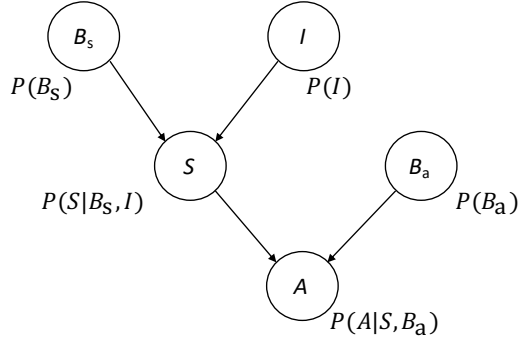
The independence assumptions represented by the BN are:

- I and B_s are (absolutely) independent of each other: $P(I|B_s) = P(I)$;
- B_a is (absolutely) independent of S, I and B_s: $P(B_a|S, I, B_a) = P(B_a)$;
- A is conditionally independent of I and B_s, given B_a and S: $P(A|B_a, S, I, B_s) = P(A|B_a, S)$.

5. (a) Four Boolean random variables can be used to represent the occurrence of the two symptoms (H for headache and Fe for fever) and of the two health problems (I for influenza and Fo for food poisoning). Influenza and food poisoning can be considered as the "root causes", and can be assumed not to directly affect each other. Each of the symptoms can be directly caused by both health problems. On the other hand, headache and fever can be assumed not to directly affect each other. A possible order between the random variables, in accord with the above causal dependencies, is: Fo, I, H, Fe. The corresponding BN is shown below.

Formally, the independence assumptions are:

- I is (absolutely) independent of Fo: $P(I|Fo) = P(I)$;
- Fe is conditionally independent of H, given I and Fo: $P(Fe|H, I, Fo) = P(Fe|I, Fo)$.

(b) The probability of interest is $P(I = \text{t}|H = \text{t})$. Using the exact inference procedure one obtains:

$$
\begin{aligned}
P(I = \text{t}|H = \text{t}) &= \frac{P(I = \text{t}, H = \text{t})}{P(H = \text{t})} \\
&= \frac{\sum_{fe,fo} P(Fe = fe, H = \text{t}, I = \text{t}, Fo = fo)}{\sum_{fe,i,fo} P(Fe = fe, H = \text{t}, I = i, Fo = fo)} \\
&= \frac{\sum_{fe,fo} P(Fe = fe|I = \text{t}, Fo = fo)P(H = \text{t}|I = \text{t}, Fo = fo)P(I = \text{t})P(Fo = fo)}{\sum_{fe,i,fo} P(Fe = fe|I = i, Fo = fo)P(H = \text{t}|I = i, Fo = fo)P(I = i)P(Fo = fo)} .
\end{aligned}
$$

(c) Following the topological order encoded by the BN a *single* sample is generated as follows:

- a sample from $P(Fo)$ is drawn: assume $Fo = \text{false}$ is obtained;
- a sample from $P(I)$ is drawn: assume $I = \text{true}$ is obtained;
- a sample from $P(H|I = \text{true}, Fo = \text{false})$ is drawn: assume $H = \text{true}$ is obtained;
- a sample from $P(Fe|I = \text{true}, Fo = \text{false})$ is drawn: assume $Fe = \text{true}$ is obtained.

The corresponding sample is $Fo = \text{false}, I = \text{true}, H = \text{true}, Fe = \text{true}$. This sample *agrees* with the evidence $H = \text{true}$, therefore it will *not* be rejected.

After generating N samples, for a given value of N, all the ones that do not agree with the evidence are rejected; then, $P(I = \text{t}|H = \text{t})$ is estimated as the fraction, among the retained samples, of the ones for which $I = \text{t}$.

6. (a) This domain can be represented using five Boolean random variables: *Core* (the core temperature exceeds the safety threshold), *Sensor* (the sensor detects a core overheating), *Heat* (the external temperature is higher than some predefined value), *WarningLights* (warning lights are flashing) and *Alarm* (the alarm is sounding).

The "root causes" are the core temperature and the external temperature. They can be assumed not to directly infuence each other. On the other hand, they both directly influence only the sensor measurement. In turn, the sensor measurement directly affects both the warning lights and the alarm, whereas the latter two can be assumed not to directly influence each other, since they are implemented as physically separated systems. Accordingly, a possible order between the random variables is: $Heat, Core, Sensor, WarningLights, Alarm$. The corresponding BN is shown below.
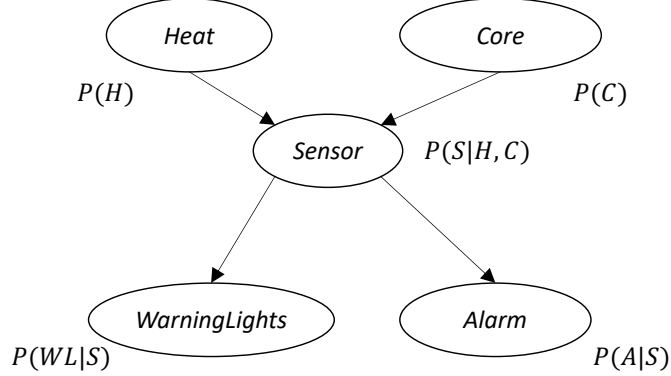
Denoting the variables with their initial letters, the formal independence assumptions are:

- C is (absolutely) independent of H: $P(C|H) = P(C)$;
- WL is independent of both C and H, given S: $P(WL|S, C, H) = P(WL|S)$;
- A is independent of WL, C and H, given S: $P(A|WL, S, C, H) = P(A|S)$.

(b) The probability of interest is $P(C = \text{true}|WL = \text{true})$. Using the exact inference procedure one obtains:

$$
\begin{aligned}
P(C = \text{t}|WL = \text{t}) &= \frac{P(C = \text{t}, WL = \text{t})}{P(WL = \text{t})} \\
&= \frac{\sum_{a,s,h} P(A = a, WL = \text{t}, S = s, C = \text{t}, H = h)}{\sum_{a,s,c,h} P(A = a, WL = \text{t}, S = s, C = c, H = h)} \\
&= \frac{\sum_{a,s,h} P(A = a|S = s)P(WL = \text{t}|S = s)P(S = s|C = \text{t}, H = h)P(C = \text{t})P(H = h)}{\sum_{a,s,c,h} P(A = a|S = s)P(WL = \text{t}|S = s)P(S = s|C = c, H = h)P(C = c)P(H = h)} .
\end{aligned}
$$

(c) According to the topological order of the nodes of the BN, a *single* sample is generated as follows:

- a sample from $P(H)$ is drawn: assume $H = \text{true}$ is obtained;
- a sample from $P(C)$ is drawn: assume $C = \text{false}$ is obtained;
- a sample from $P(S|C = \text{false}, H = \text{true})$ is drawn: assume $S = \text{false}$ is obtained;
- a sample from $P(WL|S = \text{false})$ is drawn: assume $WL = \text{false}$ is obtained.
- a sample from $P(A|S = \text{false})$ is drawn: assume $A = \text{false}$ is obtained.

The corresponding sample is $H = \text{true}, C = \text{false}, S = \text{false}, WL = \text{false}, A = \text{false}$. This sample *does not* agree with the evidence $WL = \text{true}$, therefore it will be rejected.

After generating N samples, for a given value of N, all the ones that do not agree with the evidence are rejected; then, $P(C = \text{true}|WL = \text{true})$ is estimated as the fraction, among the retained samples, of the ones for which $C = \text{true}$.