Computational Intelligence 2008–2009
Tutorial I-b

Peter Lucas, Marina Velikova and Nivea Ferreira
Institute for Computing and Information Sciences
University of Nijmegen

Bayesian networks and naive probabilistic reasoning

Recall that a Bayesian network $\mathcal{B}$ is defined as a pair $\mathcal{B} = (G, P)$, where $G = (V(G), A(G))$ is an acyclic directed graph with set of vertices (or nodes) $V(G) = \{X_1, \ldots, X_n\}$ and arcs $A(G) \subseteq V(G) \times V(G)$, and a joint probability distribution P defined on the variables corresponding to the vertices $V(G)$, as follows:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid \pi(X_i))$$

where $\pi(X_i)$ stands for the set of parents (direct ancestors) of X_i. Another, more precise version of the definition of the joint probability distribution of a Bayesian network, is to say that $X_{V(G)}$ is the set of random variables corresponding to the vertices of G, where $V(G) = \{v_1, \ldots, v_n\}$ are the associated vertices. The joint probability distribution P is then defined as follows:

$$P(X_{V(G)}) = \prod_{v \in V(G)} P(X_v \mid X_{\pi(v)})$$

where $X_{\pi(v)}$ are the random variables that correspond to the parents of vertex $v \in V(G)$. However, although mathematically more beautiful, the latter definition is also slightly more difficult to understand. It is often used in mathematically oriented book, but less often in computer-science books.

Main exercises

Exercise 1

Consider the Bayesian network $\mathcal{B} = (G, P)$ shown in Figure 1.

a. Enumerate all conditional independence assumptions $U \perp\!\!\!\perp V \mid W$ for this Bayesian network.

b. Enumerate all the unconditional independence assumptions.

c. Enumerate the dependencies $U \not\perp\!\!\!\perp V|W$.

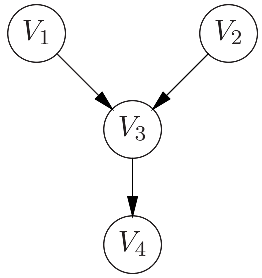

Figure 1: Bayesian network.

Exercise 2

Consider the Naive Bayes network given in Figure 2. Given the evidence: $\mathcal{E} = \{temp > 37.5\}$, compute the probability of *flu* using Bayes' rule, i. e., $P(flu \mid temp > 37.5)$.

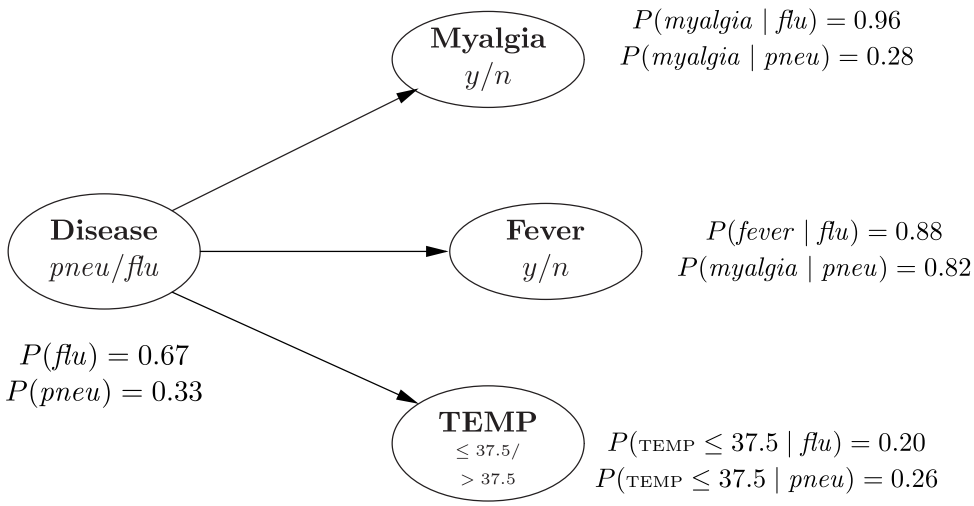

Figure 2: Naive Bayesian network.

Exercise 3

Consider Figure 3, which displays a Bayesian network in which the three vertices A, B and C interact to cause effect D through a noisy-AND. The intermediate variables I_A, I_B and I_C are not indicated in the figure, but it is assumed that for computational purposes these intermediate variables are implicitly present. The following probabilities have been specified by the designer of the Bayesian network:

$P(i_A \mid a) = 0.7$ $\quad P(i_A \mid \neg a) = 0.9$
$P(i_B \mid b) = 0.4$ $\quad P(i_B \mid \neg b) = 0.8$
$P(i_C \mid c) = 0.3$ $\quad P(i_C \mid \neg c) = 0.3$

$P(a) = 0.4$ $\quad P(b) = 0.7$
$P(c) = 0.8$

$P(e \mid d) = 0.2$ $\quad P(e \mid \neg d) = 0.6$

a. Draw the full Bayesian network with the intermediate variables I_A, I_B, and I_C included. (See the slides from the lecture "Building Bayesian Networks".)

b. Compute $P^*(e) = P(e \mid a, b, c)$, i.e. the marginal probability of e given that $A = B = C = \mathit{true}$.

c. Compute $P^*(e) = P(e \mid a, b)$.

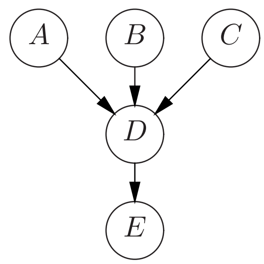

Figure 3: Bayesian network: noisy-AND.

Exercise 4

Consider again Figure 1.

a. Define a probability distribution P for this Bayesian network $\mathcal{B}$, i.e., supply numeric parameters such as $P(v_1) = 0.3$, $P(v_4 \mid \neg v_3) = 0.8$, etc.

b. Compute the marginal probability distribution $P(V_4)$.

c. Next, assume that we know that v_2 ($V_2 = \mathit{true}$) holds. What is the value of $P^*(v_2)$ and $P^*(\neg v_2)$? Compute $P^*(V_4)$, i.e. $P^*(v_4)$ and $P^*(\neg v_4)$, and also $P^*(V_1)$.

d. Finally, assume that $V_4 = \mathit{true}$ (hence, V_2 is again unknown). Compute the marginal probability distribution $P^*(V_2)$.

Additional exercises

Exercise 5

Consider the Bayesian network $\mathcal{B} = (G, P)$ shown in Figure 1. Based on the probability distribution P defined in Exercise 4, show that the following property holds:

$$P(V_1, V_2, V_3, V_4) = \prod_{i=1}^{4} P(V_i \mid \pi(V_i))$$

Exercise 6

Consider the Bayesian network $\mathcal{B} = (G, P)$ shown in Figure 4, where $G = (V(G), A(G))$ is the directed acyclic graph shown in the figure, and P is a probability distribution defined on the variables corresponding to the vertices $V(G) = \{V_1, V_2, V_3, V_4, V_5\}$. The following (local) probability distributions are defined for P:

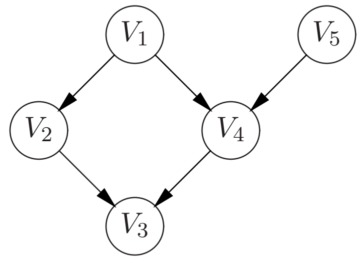

Figure 4: Bayesian network of Question 6.

$P(v_1) = 0.2$
$P(v_5) = 0.8$
$P(v_2 \mid v_1) = 0.5$ $\quad P(v_2 \mid \neg v_1) = 0.4$
$P(v_3 \mid v_2, v_4) = 0.4$ $\quad P(v_3 \mid \neg v_2, v_4) = 0.7$
$P(v_3 \mid v_2, \neg v_4) = 0.3$ $\quad P(v_3 \mid \neg v_2, \neg v_4) = 0.6$
$P(v_4 \mid v_1, v_5) = 0.6$ $\quad P(v_4 \mid \neg v_1, v_5) = 0.2$
$P(v_4 \mid v_1, \neg v_5) = 0$ $\quad P(v_4 \mid \neg v_1, \neg v_5) = 1$

a. What are the consequences of the fact that this Bayesian network contains a cycle in the underlying (undirected) graph?

b. Compute the probability $P(\neg v_5 \mid v_3)$, i.e. for V_5 equal to false given that V_3 is equal to true, if it is known that $P(v_3) \approx 0.5$.

Exercise 7

a. Mention at least two advantages of the Bayesian-network formalisms (or probabilistic graphical models in general) in comparison to the early ucertainty methods, such as the certainty-factor model and the subjective Bayesian model.

b. Mention at least one advantage of the certainty-factor model in comparison to Bayesian networks.

c. Give a counterexample showing the probabilistic incorrectness of one of the combination functions of the certainty-factor model

d. Give a counterexample showing the probabilistic incorrectness of one of the combination functions of the subjective Bayesian method.

(The certainty-factor model is explained in the "Introductory material on uncertainty reasoning" – Section 5)