Exam Probabilistic Reasoning

11 November 2011

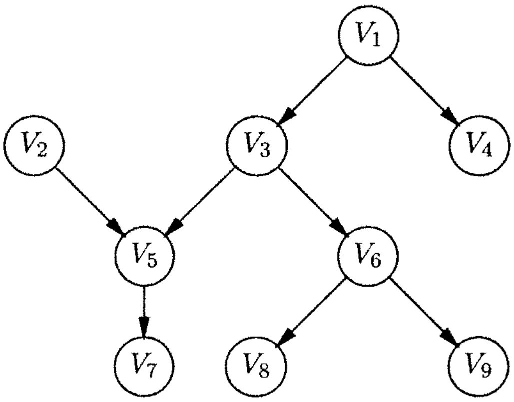
The exam consists of four problems, for each of which the number of credits per question is given. In total, a maximum of 100 credits is awarded. Read the questions very carefully; you may answer them in English and/or Dutch.

Good luck!

Reminder: please fill out the online course evaluation form. Thanks!

Problem 1 (a: 15 pts., b: 10 pts., total: 25 pts.)

Consider a probabilistic network $B = (G, \Gamma)$, where $G = (V(G), A(G))$ is the following acyclic digraph and $\Gamma = \{\gamma_{V_i} \mid V_i \in V(G)\}$ is given by:



$\gamma(v_1) = 0.65$

$\gamma(v_2) = 0.7$

$\gamma(v_3 \mid v_1) = 0.75$
$\gamma(v_3 \mid \neg v_1) = 0.4$

$\gamma(v_4 \mid v_1) = 0.1$
$\gamma(v_4 \mid \neg v_1) = 0.25$

$\gamma(v_5 \mid v_2 \wedge v_3) = 0.35$
$\gamma(v_5 \mid \neg v_2 \wedge v_3) = 0.1$
$\gamma(v_5 \mid v_2 \wedge \neg v_3) = 0$
$\gamma(v_5 \mid \neg v_2 \wedge \neg v_3) = 0.9$

$\gamma(v_6 \mid v_3) = 0.8$
$\gamma(v_6 \mid \neg v_3) = 0.6$

$\gamma(v_7 \mid v_5) = 0.5$
$\gamma(v_7 \mid \neg v_5) = 0.45$

$\gamma(v_8 \mid v_6) = 0.1$
$\gamma(v_8 \mid \neg v_6) = 0.15$

$\gamma(v_9 \mid v_6) = 0.9$
$\gamma(v_9 \mid \neg v_6) = 0.2$

Let Pr be the probability distribution defined by probabilistic network B. Consider the five computation rules of *Pearl*'s data fusion algorithm, given in the attached formula sheet. Assume that the sequence of observations $V_1 = true$ and $V_6 = false$ is entered into probabilistic network B.

a. Illustrate Pearl's algorithm by computing the probability $Pr^{v_1, \neg v_6}(v_3)$. Clearly indicate which messages are passed and how they are computed; explicitly mention all assumptions you make.

b. Suppose that for the domain of application, the probability $Pr^{v_1, \neg v_6}(v_2 \wedge v_5)$ is relevant. Clearly explain how you would compute this probability from the network. In doing so, you may consider Pearl's algorithm as a black box which returns probabilities of the form $\Pr^e(v_i)$.

Problem 2 (a: 15 pts., b: 10 pts., total: 25 pts.) Consider the independence relation defined by the following list of independence statements and assume that probability distribution $\Pr(A \wedge B \wedge C \wedge D \wedge E)$ satisfies this complete list:

$I_{\Pr}(\{A\}, \{B, C\}, \{D\})$
$I_{\Pr}(\{A\}, \{B, C, E\}, \{D\})$
$I_{\Pr}(\{A\}, \{C\}, \{E\})$
$I_{\Pr}(\{A\}, \{B, C\}, \{E\})$
$I_{\Pr}(\{A\}, \{C, D\}, \{E\})$
$I_{\Pr}(\{A\}, \{B, C, D\}, \{E\})$
$I_{\Pr}(\{B\}, \{A\}, \{C\})$
$I_{\Pr}(\{B\}, \{A, E\}, \{C\})$
$I_{\Pr}(\{B\}, \{A\}, \{E\})$
$I_{\Pr}(\{B\}, \{C\}, \{E\})$
$I_{\Pr}(\{B\}, \{A, C\}, \{E\})$
$I_{\Pr}(\{B\}, \{C, D\}, \{E\})$
$I_{\Pr}(\{B\}, \{A, C, D\}, \{E\})$
$I_{\Pr}(\{D\}, \{C\}, \{E\})$
$I_{\Pr}(\{D\}, \{B, C\}, \{E\})$
$I_{\Pr}(\{D\}, \{A, C\}, \{E\})$
$I_{\Pr}(\{D\}, \{A, B, C\}, \{E\})$

a. Draw all directed graphs that are I-maps of the above independence relation.

b. Give for each of the digraphs the corresponding factorisation of Pr.

Problem 3 (a: 15 pts., b: 10 pts., total: 25 pts.) Assume that a probabilistic network is developed for one of the hospitals in the Netherlands. The network will be used for the interpretation of x-rays. Since a large number of patient data are available, the developers have decided to use a learning algorithm for automated comstruction of the network from data. Unfortunately, after having applied the B heuristic with the MDL quality measure, a digraph with quite a large number of arcs has resulted.

a. Consider the MDL quality measure, given in the attached formula sheet, and the B search heuristic. Clearly explain how the MDL measure can be adapted to return a graph with fewer arcs than the graph which results using the original measure.

b. Suppose that the B search heuristic is again applied to the above mentioned dataset, but now with the adapted MDL measure from part a. Will the resulting graph be a subgraph from the graph that was found using the original MDL measure? Clearly explain your answer.

Problem 4 (a: 5 pts., b: 10 pts., c: 10 pts., total: 25 pts.) One in a thousand people is susceptible to a particular heart disease. There is a test to detect this disease. The test is 100% accurate for people who have the disease and is 95% accurate for those who do not (this means that 5% of people who do not have the disease will be wrongly diagnosed as having it).

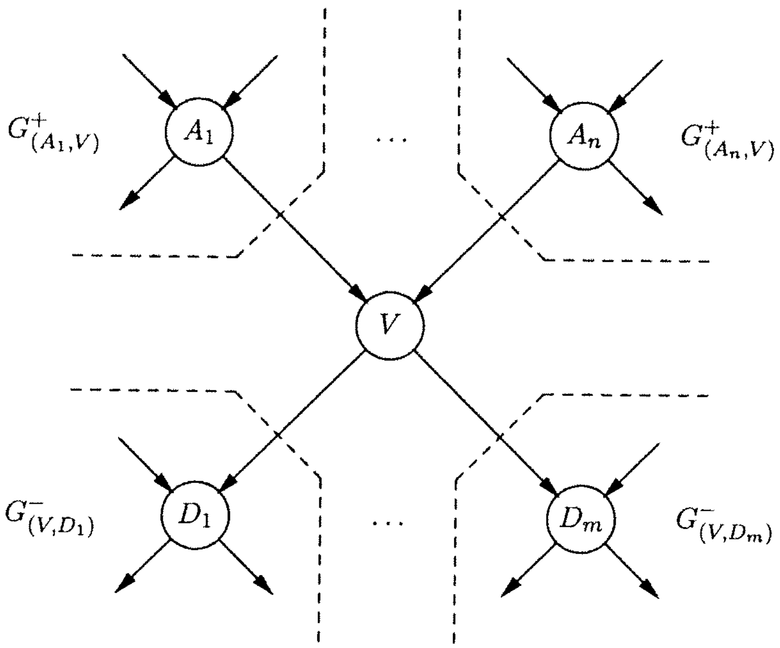
a. Draw a probabilistic network that reflects the above described relationship between disease and test-result; give all assessment functions, or conditional probability tables, as well.

b. Establish the sensitivity function $f_{\Pr(\text{Heart-disease}=yes \mid \text{Test}=yes)}(x)$ for $x = \gamma(\text{Heart-disease} = yes)$.

c. Determine the sensitivity value and the admissible deviation for $x_0 = 0.001$. Comment on their interpretation in terms of the domain under consideration.

Formulas (Probabilistic Reasoning Exam Nov 2011)

Pearl in a singly connected digraph



Consider a node V in a probabilistic network $B = (G, \Gamma)$. Let $\rho(V) = \{A_1, \ldots, A_n\}$ be the set of direct ancestors (parents) of V in G, and let $\sigma(V) = \{D_1, \ldots D_m\}$ be the set of its direct descendants (children). With Pearl's algorithm, node V computes the following parameters:

$$\pi(V) = \sum_{c_{\rho(V)}} \Big(\gamma(V \mid c_{\rho(V)}) \cdot \prod_{i=1,\ldots,n} \pi_V^{A_i}(c_{A_i}) \Big)$$

$$\lambda(V) = \prod_{j=1,\ldots,m} \lambda_{D_j}^V(V)$$

$$\pi_{D_j}^V(V) = \alpha \cdot \pi(V) \cdot \prod_{\substack{k=1,\ldots,m \\ k \neq j}} \lambda_{D_k}^V(V)$$

$$\lambda_V^{A_i}(A_i) = \alpha \cdot \sum_{c_V} \lambda(c_V) \cdot \sum_{c_{\rho(V)\setminus\{A_i\}}} \Big(\gamma(c_V \mid c_{\rho(V)\setminus\{A_i\}} \wedge A_i) \cdot \prod_{\substack{k=1,\ldots,n \\ k \neq i}} \pi_V^{A_k}(c_{A_k}) \Big)$$

in order to perform *data fusion*: $\alpha \cdot \pi(V) \cdot \lambda(V)$, which results in V's (prior or posterior) probability distribution.

The MDL quality measure

Let $G = (V(G), A(G))$ be an acyclic digraph and let D be a dataset over N cases. Let $P(G)$ be a probability distribution over the set of acyclic graphs with node set V. Then, the MDL quality measure for graph G is given by

$$\begin{aligned} Q(G,D) &= \log P(G) - N \cdot H(G,D) - \frac{1}{2}\log N \cdot \sum_{V_i \in V} 2^{|\rho(V_i)|} \\ &= \log P(G) + \sum_{V_i \in V} q(V_i, \rho(V_i), D) \end{aligned}$$

where

$$-N \cdot H(G,D) = \sum_{V_i \in V} \sum_{c_{V_i}} \sum_{c_{\rho(V_i)}} N(c_{V_i} \wedge c_{\rho(V_i)}) \cdot \log\Big(\frac{N(c_{V_i} \wedge c_{\rho(V_i)})}{N(c_{\rho(V_i)})}\Big)$$

and $q(V_i, \rho(V_i), D)$ is the quality of node V_i.