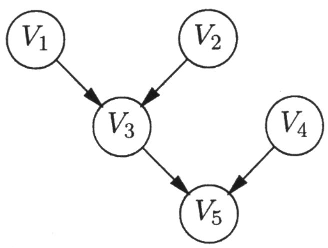
DEPARTMENT OF COMPUTER SCIENCE, FACULTY OF SCIENCE, UU.
MADE AVAILABLE IN ELECTRONIC FORM BY THE TBC OF A−Eskwadraat
IN 2005/2006, THE COURSE INFOPROB WAS GIVEN BY SILJA RENOOIJ.

Probabilistic Reasoning (INFOPROB)
6 January 2006

Question 1

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:



$\gamma(v_1) = 0.25$ $\quad$ $\gamma(v_2) = 0.6$

$\gamma(v_3 \mid v_1 \wedge v_2) = 0.2$ $\quad$ $\gamma(v_4) = 0.3$
$\gamma(v_3 \mid \neg v_1 \wedge v_2) = 0.1$
$\gamma(v_3 \mid v_1 \wedge \neg v_2) = 0.7$ $\quad$ $\gamma(v_5 \mid v_3 \wedge \neg v_4) = 0.5$
$\gamma(v_3 \mid \neg v_1 \wedge \neg v_2) = 0.35$ $\quad$ $\gamma(v_5 \mid \neg v_3 \wedge v_4) = 0.8$

a) Assume that the variables V_3 and V_4 exhibit a disjunctive interaction with respect to variable V_5 and that for specifying the assessment function for node V_5, the noisy-or gate is used. Complete the assessment function γ_{V_5} for node V_5. *(10 points)*

Let Pr be the probability distribution defined by probabilistic network B. Consider the five computation rules of *Pearl*'s data fusion algorithm.

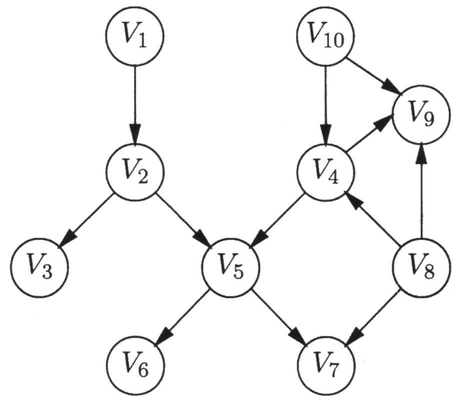
b) Assume that the observation $V_3 = true$ is entered into probabilistic network B. Illustrate Pearl's algorithm by computing the probability $\Pr^{v_3}(v_1)$. Clearly indicate which messages are passed and how they are computed; explicitly mention all assumptions you make. *(15 points)*

Now use Pearl's algorithm as a "black box": $\mathsf{Pearl}(B, \tilde{c}_E) \rightarrow \Pr^{\tilde{c}_E}(V_i)$ taking as input network B and a partial configuration $\tilde{c}_E$ of observations and returning for each node the probabilities of all its values, given the observations.

c) Explain how the probability $\Pr^{v_3}(v_1 \vee \neg v_4)$ can be efficiently computed from the network. *(10 points)*

Question 2

Consider a probabilistic network $B = (G, \Gamma)$, where $G = (V(G), A(G))$ is the following acyclic digraph with > 6 loops:



a) Give a loop cutset for graph G that can be found by applying the heuristic *Suermondt & Cooper* algorithm. *(10 points)*

b) Consider a set $C \subseteq V(G)$. Let G' be the graph that results from G by removing all *outgoing* arcs of all nodes in C, that is, $G' = (V(G), A'(G))$ with $A'(G) = A(G) \backslash \{(V_i, V_j) | V_i \in C\}$. Similarly, let G'' be a graph that results from G by removing all nodes in C together with their incident arcs, that is, $G'' = (V(G) \backslash C, A''(G))$ with $A''(G) = A(G) \backslash \{(V_i, V_j) | V_i \in C \text{ or } V_j \in C\} = A'(G) \backslash \{(V_i, V_j) | V_j \in C\}$.

A well-known property of G' is that if G' is a singly connected graph then set C is a loop cutset for G. Prove, or give a counter-example for, the following similar statement:

if G'' is a singly connected graph then set C is a loop cutset for G.

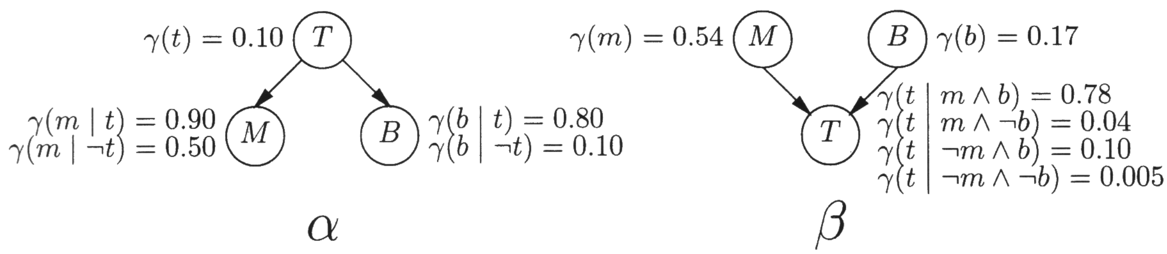
(15 points)

c) Suppose we subject network B to a one-way sensitivity analysis, where we are interested in the probability distribution $\Pr(V_5)$. More specifically, we are interested in the *most likely value* of variable V_5 and how this changes upon parameter variation. Given that there are no observations in the network and assuming that all variables in the network can only adopt two values, a one-way sensitivity analysis results in 116 sensitivity functions describing the effect of varying each of the 58 parameter probabilities on each of the two output values of V_5.

It is possible to use these sensitivity functions to establish whether or not an *observation* for one of the variables in the network could change the most likely value of variable V_5? Clearly motivate your answer. *(10 points)*

Question 3

Consider modelling the disease tuberculosis, together with its symptoms and the different tests that are used to diagnose the disease, in a probabilistic network. We focus on the variable T, with possible values t and $\neg t$ that describe whether or not a patient has tuberculosis, and two test variables M and B with values m and $\neg m$, and b and $\neg b$, respectively. Variable M models the outcome of a Mantoux test and variable B models the outcome of a bloodtest.
Consider the following two possible probabilistic networks:



a) Clearly explain which of the following statements is or are correct:

A. Network α cannot diagnose the patient, but can only predict the outcome of the two tests M and B for patients for whom we know whether or not they have tuberculosis (T).

B. Network β needs the outcome of both tests B and M to diagnose whether or not a patient has tuberculosis (T).

C. Networks α and β can both predict whether or not a patient has tuberculosis based upon ≤ 0 test-results; both networks can in addition predict the outcome of the two tests M and B for patients for whom we know for sure whether or not they have tuberculosis (T).

Observation: the two networks do not capture the same independence relation: in network α we have for example that $I(\{M\},\{T\},\{B\})$ and $\neg I(\{M\},\emptyset,\{B\})$, whereas in network β we find that $\neg I(\{M\},\{T\},\{B\})$ and $I(\{M\},\emptyset,\{B\})$.
Argument: if the test results are somewhat reliable, then, for example, a positive outcome for test M would make a positive outcome for test B more likely.

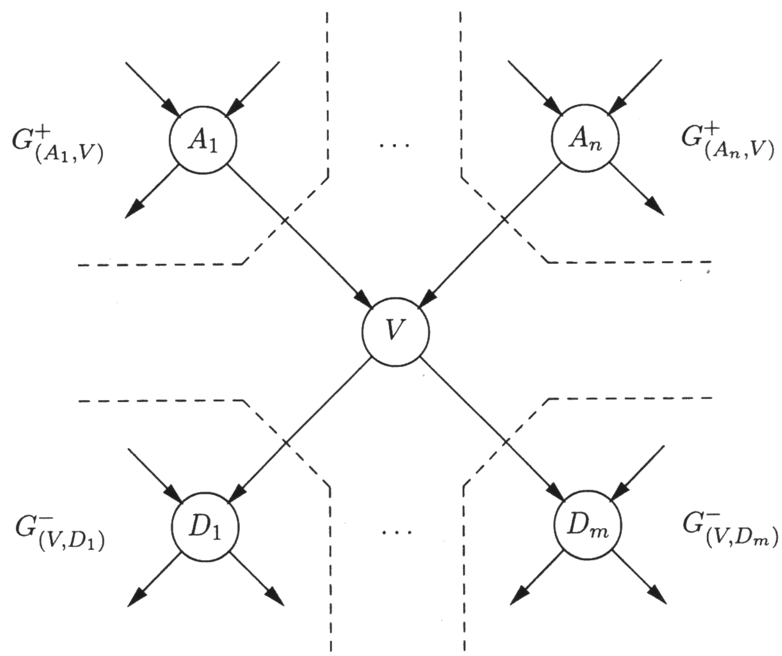
b) Assume that you agree with *Argument*, would you then prefer network α or network β? Explain your answer.

Suppose that we change the network by summarising the two testresults in an additional intermediate variable I in such a way that $\Pr(TMB)$ remains unchanged. This intermediate variable is placed between the desease variable and the two test variables, resulting in two new networks α' and β', with $A(G_{\alpha'}) = \{T \to I, I \to M, I \to B\}$ and $A(G_{\beta'}) = \{M \to I, B \to I, I \to T\}$.

c) For your network of preference from part b), are the (in)dependences among the variables T, M and B stated in *Observation* above still valid in the extended network? If not, would you prefer the simple network, or the extended one? Clearly motivate your answer(s).

Formulas

Pearl in a singly connected graph



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:

$$\begin{aligned}
\pi(V) &= \sum_{c_{\rho(V)}} \left(\gamma(V \mid c_{\rho(V)}) \cdot \prod_{i=1,\ldots,n} \pi_V^{A_i}\,(c_{A_i}) \right) \\
\lambda(V) &= \prod_{j=1,\ldots,m} \lambda_{D_j}^V(V) \\
\pi_{D_j}^V(V) &= \alpha \cdot \pi(V) \cdot \prod_{\substack{k=1,\ldots,m \\ k\neq j}} \lambda_{D_k}^V(V) \\
\lambda_V^{A_i}(A_i) &= \alpha \cdot \sum_{c_V} \lambda(c_V) \cdot \sum_{c_{\rho(V)\setminus\{A_i\}}} \left(\gamma\left(c_V \mid c_{\rho(V)\setminus\{A_i\}} \wedge A_i\right) \cdot \prod_{\substack{k=1,\ldots,n \\ k\neq i}} \pi_V^{A_k}\,(c_{A_k}) \right)
\end{aligned}$$