Hacettepe Journal of Mathematics and Statistics
Volume 33 (2004), 69 – 76

DETERMINING THE CONDITIONAL PROBABILITIES IN BAYESIAN NETWORKS

Hülya Olmuş* and S. Oral Erbaş*

Received 22 : 07 : 2003 : Accepted 04 : 01 : 2005

Abstract

Bayesian networks are used to illustrate how the probability of having a disease can be updated given the results from clinical tests. The problem of diagnosis, that is of determining whether a certain disease is present, D, or absent, D', based on the result of a medical test, is discussed. Using statistical methods for medical diagnosis, information about the disease and symptoms are collected and the databases are used to diagnose new patients. How can we evaluate the diagnostic probability represented by $\Pr(D \setminus \text{evidence})$, where evidence is the result of a clinical test or tests on a new patient? The object of this article is to answer this question. Using the HUGIN software, diagnostic probabilities are analyzed using the Bayesian approach.

Keywords: Bayesian networks, Medical diagnosis, Conditional probability.

1. Introduction

Bayesian networks were introduced in the 1980's as a formalism for representing and reasoning with models of problems involving uncertainty, adopting probability theory as a basic framework [12]. Over the last decade, the Bayesian network has become a popular representation for encoding uncertain expert knowledge in expert systems [7]. The field of Bayesian networks has grown enormously over the last few years, with theoretical and computational developments in many areas. Bayesian networks are also known as belief networks, causal probabilistic networks, causal nets, graphical probability networks, and probabilistic influence diagrams.

Bayesian networks have proved useful in practical applications, such as medical diagnosis and diagnostic systems. The probability based expert systems for medical diagnosis that emerged during the 60's and 70's could be characterized by the following points: The sets of possible diseases a system could diagnose were mutually exclusive and collectively exhaustive, the evidence was assumed conditionally independent given any hypothesis,

*Gazi University, Department of Statistics, Teknikokullar, Beşevler, Ankara, Turkey.

and only one disease was assumed to exist in any patient. These assumptions were made in order to keep to a manageable size the problem of acquiring and calculating probabilities [10].

A Bayesian network is used to model a domain containing uncertainty in some manner. It is a graphical model for probabilistic relationships among a set of variables and is composed of directed acyclic graphs (DAGs) in which the nodes represent the random variables of interest, and the links represent informational or causal dependencies among the variables [16]. Here, each node contains the states of the random variable and it represents a conditional probability table. The conditional probability table of a node contains probabilities of the node being in a specific state given the states of its parents [2, 5, 9, 11, 13, 15, 20, 21]. Furthermore, edges reflect cause-effect relations within the domain. These effects are normally not completely deterministic (e.g. disease $\rightarrow$ symptom). The strength of an effect is modelled as a probability.

Bayesian networks help us answer questions such as: What is the probability that a random variable will be in a given state if we have observed the values of some other random variables. They can also suggest what could be the best choice for acquiring new evidence. Conditional probabilities are important for building Bayesian networks. But Bayesian networks are also built to facilitate the calculation of conditional probabilities, namely the conditional probabilities for variables of interest given the data (also called evidence) at hand [5].

The quantities of interest in a medical diagnostic procedure are the probabilities of having or not having a disease, i.e. the diagnostic probabilities [17, 18]. These quantities may change their values according to the diagnostic value of the observed evidence. Evidence is produced by responses (called indicants) to clinical questions (tests, signs or symptoms). The data structure is complicated by a number of factors. Studies of acquisition for this problem occur in the literature [6, 12].

The implementation of a Bayesian network is an excellent approach to creating a medical diagnostic system that realistically models the multiple symptoms and indicators (rather than just one particular test) that affect the conditional probability that a person has a particular disease which may be causing the symptoms and positive test results. Because each node in a Bayesian network can have multiple parent and child nodes, and thus multiple ancestor and descendant nodes, evaluating Bayesian networks is more complex than performing a single calculation with Bayes' theorem.

Inference in a Bayesian network means computing the conditional probability for some variables, given information (evidence) concerning other variables. This is easy when all available evidence is for variables that are ancestors of the variable(s) of interest. But when evidence is available on a descendant of the variable(s) of interest, we have to perform inference against the direction of the edges. To this end, we employ Bayes' Theorem:

$$\Pr(A\backslash B) = \frac{\Pr(B\backslash A)\Pr(A)}{\Pr(B)}.$$

2. The HUGIN System

During the early stages of the development of probabilistic expert system, several obstacles were encountered due to difficulties in defining the joint probability distribution of the variables. With the introduction of probabilistic network models, these obstacles have largely been overcome, and probabilistic expert systems have made a spectacular comeback during the last two decades or so. These network models, which include Markov and Bayesian networks, are based on a graphical representation of the relationships among

the variables. This representation leads to efficient propagation algorithms that are used to draw conclusions. An Example of such an expert system shell is the HUGIN expert system [4]. The HUGIN system is a tool enabling the construction of model-based decision support systems in domains characterized by inherent uncertainty. The models supported are DAGs and their extension, influence diagrams. The HUGIN system allows us to define both discrete nodes and to some extent continuous nodes in our models [8]. The HUGIN system can be used to construct models as components in an application in the area of decision support and expert systems. When we have constructed a network, we can use it for entering evidence in some of the nodes where the state is known, and then retrieve the new probabilities corresponding to this evidence calculated in other nodes.

In recent years, diagnostic assistants have been built around Bayesian networks. These networks are a form of graphical probabilistic model that explicates independencies between system components and diagnostic observations in a directed graph. The structure of the graph allows the joint probability distribution over the system components and diagnostic observations to be expressed in a compact form [19]. The use of such a model along with graph-theoretic algorithms for probabilistic inference makes it possible to compute the probability of a component defect given the outcomes of diagnostic observations. There are several commercial and research tools designed for BN model authoring and testing. Among the most popular of these tools is the HUGIN package.

After constructing a Bayesian network that models, as in the example presented in the figure, the states of affairs and their probabilistic causal relationships, one would want to be able to determine, given observed values for any number of nodes in the network, the conditional probabilities of the remaining, unknown nodes. The utility of Bayesian networks lies in being able to make this calculation, which is called *evaluating* or *solving* the network. An algorithm for evaluating Bayesian networks can determine probabilities of causes given observed effects (e.g., the probability that a dam has failed given the observation that there is flooding) or probabilities of effects given observed causes (e.g., the probability that there is flooding given a low barometer reading). In order to maximize efficiency and minimize execution time, algorithms that give exact solutions of Bayesian networks must first simplify the network itself before proceeding with the evaluation process. There is no one algorithm for obtaining exact solutions that is efficient for all Bayesian networks; the choice of an exact algorithm depends on the topological characteristics of the particular Bayesian network that is to be evaluated. There are, however, several approximation schemes which yield reasonably accurate solutions and require less execution time than the exact algorithms. HUGIN is a software package that implements algorithms for evaluating Bayesian networks [3]. Algorithms that achieve exact solutions are derived from Bayes's theorem. Bayes's theorem can be used to make a simple calculation of the conditional probability of a hypothesis given its evidence.

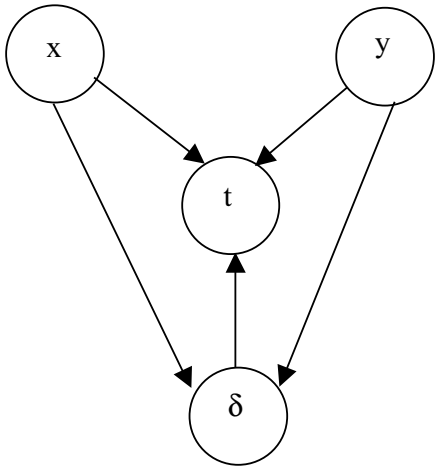
3. A Menopause Example

In this example, the patients who applied to Gazi University Gynecology and Obstetrics Menopause Clinic during the period August–October 1998 are studied [1]. A patient consults with a specialist who is going to start a search to discover whether the patient has the postmenopausal condition, D, or its absence, D'. The physician observes an indicant (E^+= normal bone density or E^-= abnormal bone density), which is new evidence associated with the patient. In a search for information about this new indicant of the postmenopausal condition D, doctors in a certain clinic select 100 patients known to be in postmenopause and another 100 patients known to be in premenopause. Here D is the event that a patient has the postmenopausal condition, while D' is the event

that a patient has the premenopausal condition D'. To each patient they applied a bone mineral densitometry (BMD) test, obtaining a response E^+ for evidence of normal bone density, or E^- for evidence of abnormal bone density [14, 17, 18].

In constructing the graph for the Bayesian network, human experts mostly use "causal" relationships between variables as a guideline. The situation can be modelled by the Bayesian network in Figure 1. In Figure 1, we have the graphical representation of the Bayesian network. However, this is only what we call the qualitative representation of the Bayesian network. We need to specify the quantitative representation. The quantitative representation of a Bayesian network is the set of conditional probability tables of the nodes.

Figure 1. Bayesian network for the menopause example.



The Bayesian network consist of four nodes: x, y, t and δ which can all be in one of two states. Node x can be is the state corresponding to "normal bone density" or "abnormal bone density" as a result of a BMD test among all former patients with D and node y can be in the state corresponding to "normal bone density" or "abnormal bone density" as a result of a BMD test among all former patients with D'. The state of a new patient is

$$\delta = \begin{cases} 1 & \text{if the patient has postmenopausal condition } D \\ 0 & \text{if the patient has the premenopausal condition } D'. \end{cases}$$

The result of the test for a new patient is

$$t = \begin{cases} 1 & \text{if the BMD test gives normal bone density, i.e. } E^+ \\ 0 & \text{if the BMD test gives abnormal bone density, i.e. } E^-. \end{cases}$$

Here, the conditional probabilities are $\Pr(x), \Pr(y), \Pr(t \backslash x, y, \delta)$ and $\Pr(\delta \setminus x, y)$. Note that all four tables show the probability of a node being in a specific state depending on the states of its parent nodes, but x and y do not have any parent nodes.

The Bayesian network diagram that permits us to evaluate the diagnostic probabilities for all possible values of δ, x, y is presented in Figure 1.

The diagnostic probabilities, the object of the analysis, are $\Pr\{\delta = 1 \backslash t = 1\}$ and $\Pr\{\delta = 1 \backslash t = 2\}$. If a new woman patient's BMD test response is known to be "normal bone density" or "abnormal bone density", what is the probability that this woman is in

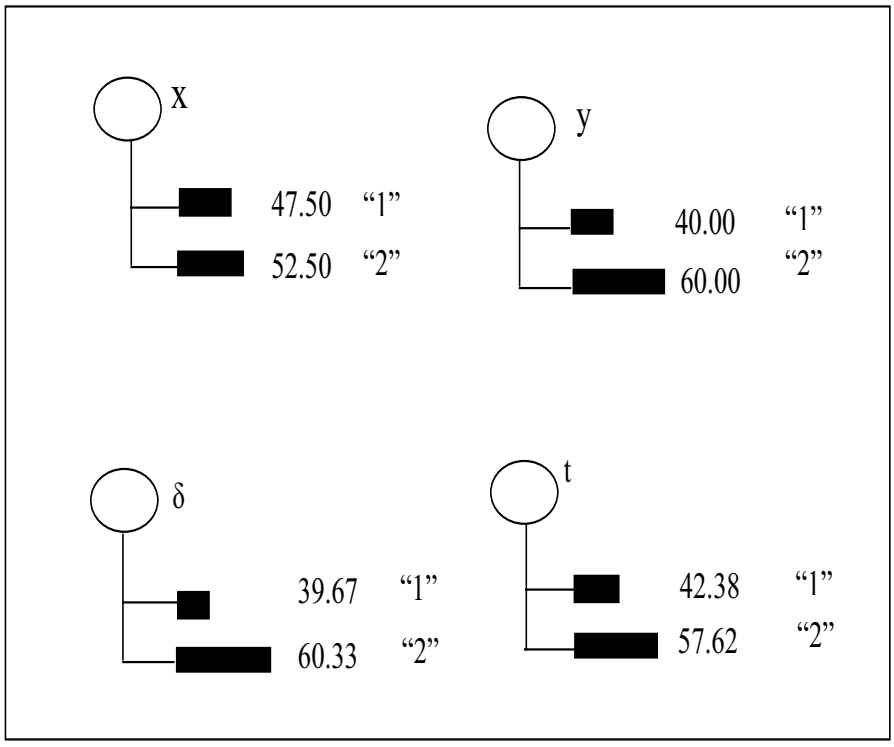
a postmenopausal or premenopausal condition? The answer to our problem is given by the probability functions attached to node t.

In this example, the model is defined using binary variables. In the following figures, empty boxes show observable variables whereas the values shown by full boxes are probability values. Also, the value "100" in the figures indicate that the selected level of variable is known. In the figures "1" indicates "normal bone density", whereas "2" indicates "abnormal bone density".

The menopause Bayesian network has been constructed using the HUGIN software. Here the probability that $\delta = 1$ is the prior probability. This prior probability was taken as 0.15 using expert belief. In other words, $\Pr(\delta = 1) = 0.15$.

In Figure 2, the model is shown with initial probabilities. For example, the "normal bone density" and "abnormal bone density" response probabilities for 100 postmenopausal women were 0.4750 and 0.5250, respectively. In other words, $\Pr(x = 1) = 0.4750$ and $\Pr(x = 2) = 0.5250$. On the other hand the "normal bone density" and "abnormal bone density" response probabilities for 100 premenopausal women were 0.40 and 0.60 respectively. In other words, $\Pr(y = 1) = 0.40$ and $\Pr(y = 2) = 0.60$. However, the probability of a new woman patient being postmenopausal is 0.3967, the probability of a woman not being postmenopausal is 0.6033. In other words, $\Pr(\delta = 1) = 0.3967$ and $\Pr(\delta = 2) = 0.6033$. The "normal bone density" and "abnormal bone density" response probability to the BMD test for a new patient are 0.4238 and 0.5762 respectively. In other words, $\Pr(t = 1) = 0.4238$ and $\Pr(t = 2) = 0.5762$.

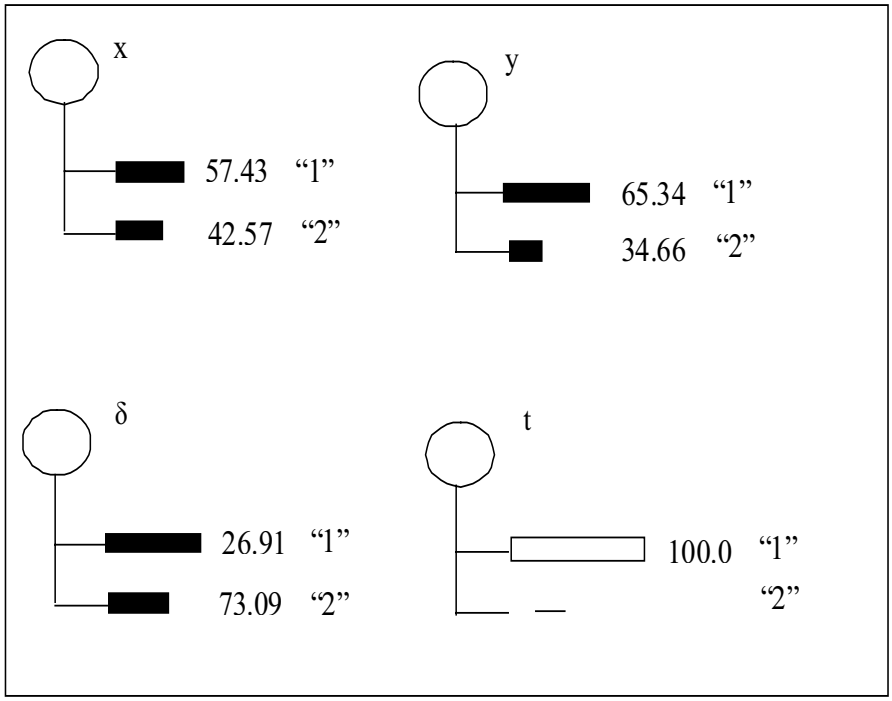
Figure 2. Marginal Probabilities



Now, one might want to know the probability of any other combination of states under the assumption that the evidence entered holds. Here, we want to calculate the probability of any other combination of states given the evidence provided by the result of the test for the new patient.

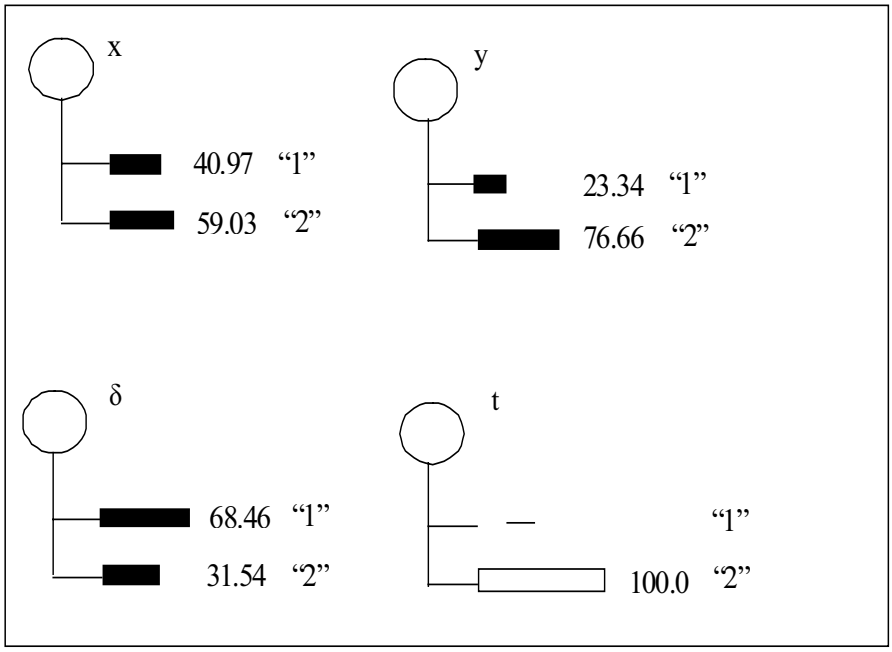
If the result of the test for the new patient is "normal bone density", then the evidence is entered and a sum-propagation is performed. In other words, this gives probabilities $\Pr(\delta = \text{“1”} \setminus t = \text{“1”})$, $\Pr(\delta = \text{“2”} \setminus t = \text{“1‘”})$. The result is shown in Figure 3. For example, if the response to the BMD test for new for woman patient is "normal bone density", the "normal bone density" and "abnormal bone density" response probabilities are 0.5743 and 0.4257 for the 100 postmenopausal women respectively. In other words, $\Pr(x = \text{“1”} \setminus t = \text{“1”}) = 0.5743$ and $\Pr(x = \text{“2”} \setminus t = \text{“1”}) = 0.4257$. Similarly if the new woman patient's BMD test response is known as to be "normal bone density", the "normal bone density" and "abnormal bone density" response probabilities are 0.6534 and 0.3466 for the 100 premenopausal women respectively. In other words, $\Pr(y = \text{“1”} \setminus t = \text{“1”}) = 0.6534$ and $\Pr(x = \text{“2”} \setminus t = \text{“1”}) = 0.3466$. Conversely if the new woman patient's BMD test result is known to be negative, the probability of being postmenopausal or premenopausal for this patient are 0.7309 and 0.2691 respectively. In other words, $\Pr(\delta = \text{“1”} \setminus t = \text{“1”}) = 0.2691$ and $\Pr(\delta = \text{“2”} \setminus t = \text{“1”}) = 0.7309$.

Figure 3. The conditional probabilities of other nodes if the new patient is known to have "normal bone density" as a result of the BMD test.



If the result of the test for the new patient is "abnormal bone density", then the evidence is entered and sum-propagation is performed. The result is shown in Figure 4. In other words, this produces the probabilities $\Pr(\delta = \text{“1”} \setminus t = \text{“2”})$ and $\Pr(\delta = \text{“2”} \setminus t = \text{“2”})$. The result is shown in Figure 4.

Figure 4. The conditional probabilities of other nodes if the new patient is known to have "abnormal bone density" as a result of the BMD test.



For example, if a new woman patient's BMD test response is known to be "abnormal bone density", the "normal bone density" and "abnormal bone density" response probabilities are 0.4097 and 0.5903 for the 100 postmenopausal women respectively. In other words, $\Pr(x = \text{“1”} \setminus t = \text{“2”}) = 0.4097$ and $\Pr(x = \text{“2”} \setminus t = \text{“2”}) = 0.5903$. Conversely if a new woman patient's BMD test response is known to be "abnormal bone density", the probability of being postmenopausal or premenopausal for this patient are 0.4756 and 0.5254 respectively. In other words, $\Pr(\delta = \text{“1”} \setminus t = \text{“1”}) = 0.6846$ and $\Pr(\delta = \text{“2”} \setminus t = \text{“1”}) = 0.3154$.

4. Conclusion

Bayesian networks are becoming an increasingly important area in applications to medical diagnosis. Here, the cause-effect relation among variables is explained and thus the relations between the variables are modelled.

The analysis of this medical problem has several important applications, including updating the probabilities for data in expert systems. In this study, depending on the results of a clinical test, the probability of a new woman patient being in menopause or not is examined. Also, the conditional probabilities of other nodes are obtained if the new patient is known to have "normal bone density" or "abnormal bone density" as a result of a BMD test.

In this application, the marginal probability of a new patient, who comes to the clinic, being in menopause is 0.3967. If she is known to have "normal bone density", an increase is not observed in the probability of this woman being in menopause. But if she is known to have "abnormal bone density", an increase is observed in the probability of this woman being in menopause.

Osteoporosis is one at the diseases causing bone resorption after the menopause. But it can be seen at an earlier age in young persons infected by the bone disease, and other metabolic disfunctions beside osteoporosis and the menopause. Osteoporosis may not be seen in every women during the menopause. As a conclusion of this study, menopause probability was seen to be higher than normal for those who were diagnosed as having bone resorption by the BMD test. However, it is not possible to correlate bone resorption with the menopause alone. Bone resorption can result from pregnancy, smoking, using alcohol, malnutrition and some hormonal and genetic disturbances. In this study, by using the HUGIN software, the correlation between the menopause and osteoporosis was evaluated by neglecting all other parameters.

References

[1] Ataoğlu, F. *Menapoz dönemindeki kadınların hormon replasman tedavisine (HRT) başlanmadan önceki ve tedavi sonrasındaki serum lipid, alkalen fosfataz ve kalsiyum değerleri*, (Gazi Üniversitesi Sağlık Bilimleri Enstitüsü Biyokimya Anabilim Dalı, Yayınlanmamış Yüksek Lisans Tezi, Ankara, 1999).

[2] Buntine, W. *A guide to the literature on learning probabilistic networks from data*, IEEE Transactions on Knowledge and Data Engineering **8** (2), 195–210, 1996.

[3] Charniak, E. *Bayesian networks without tears*, AI Magazine **12** (4), 50–63, 1991.

[4] Castillo, E., Gutierrez J. M. and Hadi A. S. *Expert systems and probabilistic network models*, (Springer-Verlag, New York, 1997).

[5] Cowell, R. G. *Intrduction to inference in Bayesian networks*, in Learning in Graphical Models, 9–26, 1999.

[6] Dawid, A. P. *Properties of diagnostic data distributions*, Biometrics **32**, 647–658, 1976.

[7] Heckerman, D. *A tutorial on learning with Bayesian networks*, Technical Report MSR-TR-95-06i, Microsoft RESEARCH, 301–354, 1995.

[8] HUGIN EXPERT A/S. *Hugin Help Pages*, URL: [http://www.hugin.dk] Date: 23.05.2000, 1998.

[9] Jensen, F. V. *An introduction to Bayesian Networks*, (UCL Press Ltd., London, 1996).

[10] Jensen, F. V., Olesen, K. G. and Andersen, S. K. *An algebra of Bayesian belief universes for knowledge-based systems*, Networks **20**, 637–659, 1990.

[11] Liarokapis, D. *An introduction to belief networks*, htttp://www.cs.umb.edu/ dimitris, 1999.

[12] Lucas, P. *DAGs in Medicine: a Model-based Approach to Medical Decision Graphs*, (Department of Computing Science, University of Aberdeen, Scotland, UK, 1999).

[13] Niedermayer, D. *An introduction to bayesian networks and their contemporary applications*, (http://www.gpfn.sk.ca/ daryle/papers/bayesian
_networks/bayes html, 1998).

[14] Oliver, R. M. and Smith, J. Q. *Influence Diagrams, Belief Nets and Decision Analysis*, (John Wiley & Sons, New York, 1990).

[15] Pearl, J. *Causality: Models, Reasoning and Inference*, (Cambridge University Press, England, 2000).

[16] Pearl, J. *Bayesian Networks*, (UCLA Cognitive Systems Laboratory, Technical Report (R-246), revision I, 1997).

[17] Pereira, C. A de B. *Medical diagnosis using influence diagrams*, Networks **20**, 565–577, 1990.

[18] Pereira, C. A de B. and Pericchi L. R. *Analysis of diagnosability*, Applied Statistics **39**, No 2, 189–204, 1990.

[19] Przytula, K. W., Dash, D. and Thompson, D. *Evaluation of Bayesian Networks Used for Diagnostics*, (IEEE Aerospace Conference, Big Sky, Montana, 2003).

[20] Spiegelhalter, D. J. and Lauritzen, S. L. *Sequetial updating of conditional probabilities on directed graphical structures*, Networks **20**, 579–605, 1990.

[21] Stephenson, T. A. *An introduction to Bayesian network theory and usage*, (IDIAP Research Report 00-03, 2000).