Comparison of Rule-Based and Bayesian Network Approaches in Medical Diagnostic Systems*

Agnieszka Oniśko[1], Peter Lucas[2], and Marek J. Druzdzel[3]

[1] Białystok University of Technology, Institute of Computer Science, Białystok, 15–351, Poland, aonisko@ii.pb.bialystok.pl
[2] Department of Computing Science, University of Aberdeen, Aberdeen AB24 3UE, Scotland, UK, plucas@csd.abdn.ac.uk
[3] Decision Systems Laboratory, School of Information Sciences, Intelligent Systems Program, and Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A., marek@sis.pitt.edu

Abstract. Almost two decades after the introduction of probabilistic expert systems, their theoretical status, practical use, and experiences are matching those of rule-based expert systems. Since both types of systems are in wide use, it is more than ever important to understand their advantages and drawbacks. We describe a study in which we compare rule-based systems to systems based on Bayesian networks. We present two expert systems for diagnosis of liver disorders that served as the inspiration and vehicle of our study and discuss problems related to knowledge engineering using the two approaches. We finally present the results of a simple experiment comparing the diagnostic performance of each of the systems on a subset of their domain.

1 Introduction

Two major classes of expert systems are those based on rules, known as *rule-based expert systems*, and those based on probabilistic graphical models, often referred to as *probabilistic expert systems* or *normative systems*. Rule-based expert systems, originating from the pioneering work of Buchanan and Shortliffe on the Mycin system [1], aim at capturing human expertise in terms of rules of the form **if** *condition* **then** *action*. There is overwhelming psychological evidence (e.g., [9]) that such rules are capable of modelling the human thought process. A set of rules can capture a human expert's relevant knowledge of a domain and can be subsequently used to reproduce the expert's problem solving in that domain. Probabilistic expert systems originate from research at the intersection of statistics and artificial intelligence. This research focuses on the concepts of relevance and probabilistic independence and has led to the development of intuitive and efficient graphical tools for knowledge representation. A prominent tool for

* The following grants supported our work: KBN 8T11E02917, W/II/1/00, AFOSR F49620–00–1–0112, NSF IRI–9624629, NATO PST.CLG.976167.

S. Quaglini, P. Barahona, and S. Andreassen (Eds.): AIME 2001, LNAI 2101, pp. 283–292, 2001.

capturing expert knowledge in this approach are *Bayesian networks* [11], often referred to, somewhat imprecisely, as *causal networks*, because of their ability to capture causal relations. Bayesian networks, while also aim at capturing expert knowledge, are based on the mathematical foundations of probability theory. When used in reasoning, they apply mathematical formalism and make no claim about reproducing the expert's thought process.

Several authors have studied theoretical differences between rule-based expert systems and normative systems (e.g., [2,6,13]), in particular with respect to handling uncertainty. Much less work, however, has been done on studying the implications that choosing one approach over the other has on the knowledge engineering effort and overall system performance. Today, theoretical developments and practical experiences with the probabilistic systems are matching those of rule-based expert systems. Both rule-based and probabilistic systems are in wide use and it is more than ever important to understand the advantages and drawbacks of each of the approaches.

Our paper focuses on comparing the two approaches in the context of a challenging practical problem that we worked on independently, using both rule-based and probabilistic approaches: diagnosis of liver disorders. Expert systems that we have developed are of considerable size and have taken several years to build. Hepatology, the study of diseases of the liver and biliary tract, is an excellent domain for such comparison, as it is complex, contains both rare and frequently occurring disorders, disorders for which both much biomedical knowledge is available and which are described only in terms of symptoms and signs.

The remainder of this paper is structured as follows. Section 2 summarises the main principles of the rule-based and probabilistic expert systems, and introduces the systems on which our comparison is based. Section 3 focuses on the qualitative comparison of the two systems and, in particular, on the knowledge engineering aspects of the two approaches. Section 4 describes the results of a study that aimed at evaluating the diagnostic performance of the two systems. Finally, Section 5 summarises our most important findings.

2 The Basics

The systems, Hepar and Hepar II (a successor of another Hepar project!), focus on the diagnosis of liver disorders. *Hepar* is Greek for *liver* and this explains similarity in names of the systems, which have been conceived independently. We realize that these names may lead to confusion on the part of the reader and, therefore, we will refer to them in this paper by Hepar-RB and Hepar-BN respectively. Hepar-RB is a rule-based system which was developed from 1984 to 1990 [8]. Its development was a joint work of Roelof Janssens (hepatologist, St Elisabeth Hospital, Leidschendam, The Netherlands) and Peter Lucas; the latter has since then also done work on probabilistic and decision-theoretic systems. Hepar-BN, still under development, is a probabilistic system and it is a collaborative effort of Agnieszka Oniśko, Marek Druzdzel and Hanna Wasyluk (hepatologist, Medical Centre for Postgraduate Education, Warsaw, Poland) [10].

2.1 Rule-Based Expert Systems and the HEPAR-RB Project

A rule-based expert system $\mathcal{S}$ can be defined as a triple $\mathcal{S} = (\Delta, \Phi, \mathcal{R})$, with a set of possible conclusions Δ, a set of observable findings Φ, and a set of generalised Horn clauses, or *rules*, $\mathcal{R}$ taking the form:

$$(e_1 \wedge \cdots \wedge e_p \wedge \sim f_1 \wedge \cdots \wedge \sim f_q) \rightarrow d_x \tag{1}$$

where $d \in \Delta$ and $e_i, f_j \in (\Delta \cup \Phi)$ (the subscript x of d is discussed below). The negation sign $\sim$ in a rule denotes a special type of closed-world assumption, called *negation by absence* [7]. Negation by absence was especially designed to accommodate the way medical doctors handle patient findings: usually only positive (present) findings are recorded, whereas only a small proportion of negative (absent) findings are written down. A finding may only be assumed to be absent when the corresponding test has been performed. For example, $\sim$(*jaundice* $\in$ Signs) is true when an attempt has been made to observe signs in the patient, and no jaundice was observed; formally: ((Signs $\neq$ *unknown*) $\wedge$ (*jaundice* $\notin$ Signs)).

If $H \subseteq \Delta$ is a set of possible diagnostic conclusions, called the *hypothesis set*, and $E \subseteq \Phi$ is the set of observed patient findings, then a *diagnosis* D is defined as follows [5]:

$$D = \{d \in H \mid \mathcal{R} \cup E \vdash_{\mathrm{NA}} d\}$$

where $\vdash_{\mathrm{NA}}$ denotes logical deduction using negation by absence. Note that this definition implies that rule-based systems of the Mycin type are deductively incomplete. However, this incompleteness has no significant drawback. On the contrary, it even has an advantage: it reduces the amount of information requested from the user [5].

A traditionally popular way to deal with uncertainty in rule-based expert systems has been the certainty-factor calculus as originally developed for the (E)Mycin system by Shortliffe and Buchanan [1]. The subscript x in rule (1) expresses uncertainty with respect to d given absolute certainty of the rule's conditions. Although the certainty-factor calculus has attracted a fair amount of criticism, the model is in fact related to encoding and processing uncertainty in Bayesian networks, as was only recently shown [6].

HEPAR-RB [7] is a rule-base expert system which is structured along the lines briefly discussed above. The system is able to differentiate among nearly 80 disorders of the liver and biliary tract. It uses a hierarchical reasoning strategy, as illustrated in Fig. 1. The system includes 118 different variables, some of which are multivalued. Only non-invasive tests are included; liver biopsy, for example, is not included, as one of the aims of the development of the system was to assist clinicians in the appropriate selection of patients who need to be submitted to such invasive tests.

A performance evaluation has been carried out twice, using data from the Rotterdam University Hospital (Dijkzigt Hospital) [8]. The second of these datasets, which includes 181 patient cases after removal of patients of whom the diagnosis was unclear, will be used in this paper as one of the datasets for a comparative quantitative analysis of HEPAR-RB and HEPAR-BN. This dataset will be referred to in the sequel as the *DH dataset*.

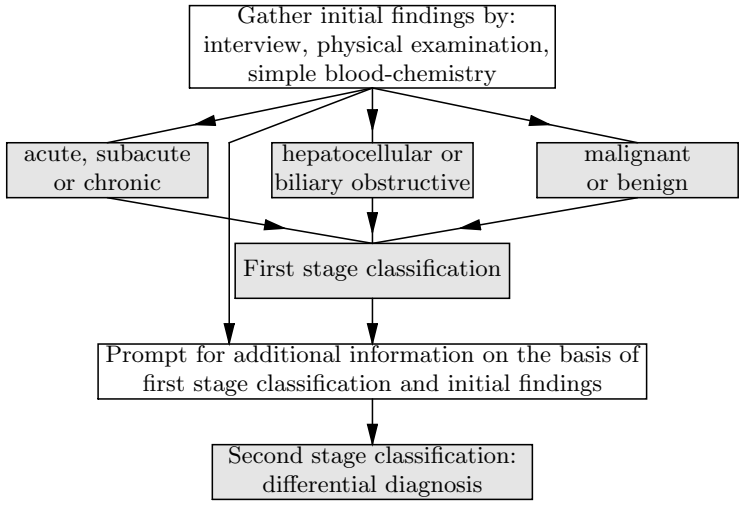

Fig. 1. Diagnostic strategy in HEPAR-RB.

2.2 Probabilistic Expert Systems and the HEPAR-BN Project

A probabilistic expert system consists of a Bayesian network with a set of algorithms to manipulate its incorporated probabilistic information. More formally, a Bayesian network is defined as a pair $\mathcal{B} = (G, \Pr)$, where G is an acyclic directed graph, modelling probabilistic (in)dependencies among variables, and Pr is a set of local conditional probability distributions, which together define a joint probability distribution on the variables. The graphical part of a Bayesian network normally reflects the causal structure of a problem. Fig. 2 shows a simplified fragment of the HEPAR-BN Bayesian network model. The network models 18 variables related to diagnosis of a small set of hepatic disorders: three risk factors, 12 symptoms and test results, and three disorder nodes.

Given a patient's case E, i.e., values of some of the modelled variables, such as risk factors, symptoms, and test results, a probabilistic system derives the posterior probability distribution over the possible disorders D: $\Pr(d \mid E)$, for each $d \in D$. This probability distribution can be directly used in diagnostic decision support.

To give the reader an idea of the number of numerical parameters needed to quantify a Bayesian network, let us assume for simplicity that each variable in the model in Fig. 2 is binary. The complete joint probability distribution for 18 binary variables would contain $2^{17} = 131{,}072$ independent parameters (we take here into account that for every propositional variable x, $\Pr(\overline{x}) = 1 - \Pr(x)$). Explicit information about independencies included in the model allows for specifying the joint probability distribution by means of a series of conditional probability distribution tables (CPTs) of individual nodes conditional on their direct predecessors. A CPT for a binary variable with n binary predecessors requires specification of 2^n independent parameters. A popular approximation of the in-

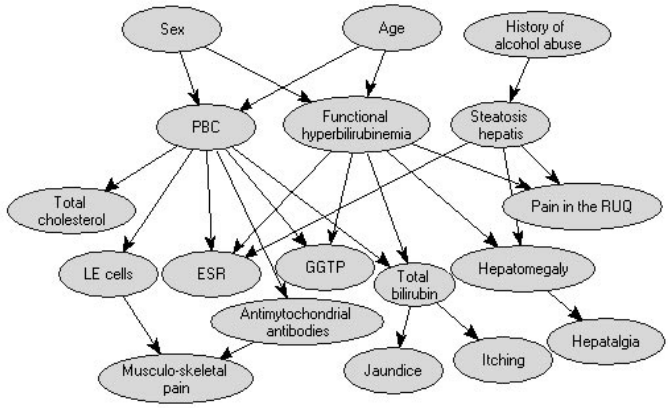

Fig. 2. A Bayesian network (a simplified fragment of the HEPAR-BN model).

teraction between a node and its direct predecessors in a Bayesian network is the Noisy–OR gate [11]. In Noisy–OR gates, each of the arcs is described by a single number expressing the causal strength of the interaction between the parent and the child. If there are other, unmodelled causes of a, we need one additional number, known as the *leak probability* [3], denoting the cumulative causal strength of all unmodelled causes of a. If each of the interactions in our model is approximated by a leaky Noisy–OR gate, only 43 numbers suffice to specify the entire joint probability distribution. It is apparent from the above example that Bayesian networks offer a compact representation of joint probability distributions and are capable of practical representation of large models.

The HEPAR-BN model [10] is a causal Bayesian network involving a subset of the domain of hepatology: 11 liver diseases (described by 9 disorder nodes), 18 risk factors, and 44 symptoms and laboratory tests results. It was designed for gathering and processing clinical data of patients with liver disorders and, through its diagnostic capabilities, reducing the need for liver biopsy. An integral part of the HEPAR system is its dataset, created in 1990 and thoroughly maintained since then at the Gastroentorogical Clinic of the Institute of Food and Feeding in Warsaw (we will refer to this dataset as the *IFF dataset*). The current IFF dataset contains over 800 patient records and its size is still growing. Each hepatological case is described by over 200 different medical findings.

3 Qualitative Comparison

There are numerous structural, qualitative differences between rule-based and Bayesian network systems which determine the way these systems diagnose disorders in patients. These differences also give rise to different methodologies for the development of such systems. Due to space constraints, we will compare HEPAR-RB and HEPAR-BN only with regard to multiple-disorder diagnosis and knowledge modelling.

3.1 Single versus Multiple Disorder Diagnosis

The *single disorder assumption* takes that a patient only suffers from one disorder at the same time, i.e., disorders are assumed to be *mutually exclusive*. This is often unnecessarily restrictive, as it happens fairly often that a patient suffers from *multiple disorders* and a single disorder may not account for all observed findings. This is certainly the case for liver diseases, where an initial disease process may give rise to multiple disorders.

As diagnostic problem solving in rule-based systems is taken as logical deduction, the capability of producing multiple, alternative diagnostic solutions is dependent on the freedom allowed in the syntax of rules. Multiple disorder diagnosis is only possible when disjunctions are allowed in the rules, i.e., when rules are non-Horn clauses. As many rule-based systems, including HEPAR-RB, restrict rules to the (generalised) Horn-clause format, multiple disorder diagnoses such as $d_1 \vee (d_2 \wedge d_3)$, the patient has either d_1 or both d_2 and d_3, is not possible. Instead, diagnostic conclusions have the form $D = \{d^1_{x_1}, \ldots, d^k_{x_k}\}$, to be interpreted as: $d^1_{x_1} \wedge \cdots \wedge d^k_{x_k}$. This conjunction consists of diagnostic conclusions $d^j_{x_j}$ with a measure of uncertainty x_j. HEPAR-RB includes a multi-valued variable 'diagnosis' which can take on disorders as values. A simple type of multiple disorder diagnosis is therefore obtained in HEPAR-RB, as values are not assumed to be mutually exclusive.

In the case of a Bayesian network, the single disorder diagnosis assumption is modelled by using a single diagnostic variable D, which takes disorders as its possible values. Due to the axiom $\sum_{d \in \rho(D)} \Pr(D = d) = 1$, where $\rho(D)$ is the domain of the variable D, the disorders are *mutually exclusive*. This was one of the underlying assumptions of an initial version of HEPAR-BN, influenced by the IFF dataset that was available to us that had every patient case ultimately diagnosed with only one liver disorder.

Instead of representing all disorders by a single variable, every disorder can also be represented by a separate variable. This allows for exploiting the multiple disorder assumption in Bayesian networks. Normally, a Bayesian network algorithm is used to compute the posterior probability:

$$\Pr(d \mid E) \tag{2}$$

for each $d \in \Delta$, where Δ is the set of disorder variable, and E is the set of patient findings. However, the probability of this co-occurrence is actually:

$$\Pr(D \mid E) \tag{3}$$

where $D \subseteq \Delta$; for example $D = \{d_1, d_2\}$. Bayesian networks incorporate the necessary probabilistic information to compute probability (3). Computing the probability distribution over all possible combinations of diseases is for a sufficiently large set of diseases infeasible. Many Bayesian-network systems including HEPAR-BN, therefore, compute the probability distributions (2), and do not attempt to do full multiple disorder diagnosis.

We may conclude that although the foundations of HEPAR-RB and HEPAR-BN are quite different, the systems implement essentially the same restricted

type of multiple disorder diagnosis. This facilitates the comparison of the systems' diagnostic performances.

3.2 Knowledge Modelling

Development of a rule-based expert system amounts to eliciting heuristic knowledge of the type Features ⇒ Class, where 'Features' is a Boolean expression involving features relevant with respect to the class, and ⇒ a classification relation. For example, one of classification relations underlying one of the logical rules in HEPAR-RB is the following:

$$(\text{Chronic\_disorder} \wedge \text{Female\_gender} \wedge (age > 40) \wedge (\text{Raynaud's\_phenomenon} \vee \{burning\_eyes, dry\_mouth\} \subseteq \text{Signs}) \wedge \text{Biliary-obstructive\_type}) \Rightarrow \text{PBC}$$

Development of a Bayesian network usually involves causal modelling, where one attempts to acquire knowledge of the form $\text{Cause}_1, \ldots, \text{Cause}_n \rightarrow \text{Effect}$. For example, in HEPAR-BN one of the underlying causal relations is:

$$\text{Cirrhosis}, \text{Total\_proteins\_low} \rightarrow \text{Ascites}$$

For rule-based systems, the essential modelling problem is to decide which features to include in the classification relation, and which to leave out, as this will have a significant bearing on the quality of the system. The decision which variables to include in a causal relation when developing a Bayesian network is less hard, as experts are usually confident about the factors influencing another factor. Moreover, acquiring information about the actual interaction among the factors involved in a causal relation is dealt with separately, when acquiring probabilistic information. This seems to suggest that the development of a Bayesian network may require less time than the development of a corresponding rule-based system. On the other hand, the developer of a rule-based system has a more direct control over the diagnostic behaviour of the system, whereas in the case of a Bayesian network, insight into possible changes in behaviour is mainly obtained by examining the model's result for test cases.

HEPAR-RB was developed over a period of four years, during regular meetings with Dr. Roelof Janssens, taking a total of approximately 150 hours. Certainty-factors were gathered by using the direct scaling method, i.e., a scale from -100 to 100 was used, and the hepatologist was asked to indicate his belief concerning certain statements on this scale. Subsequently, considerable time has been invested in refining the rule base using patient data [4].

In case of the HEPAR-BN system, we elicited the structure of the model based on medical literature and conversations with our domain expert, Dr. Hanna Wasyluk. We estimate that elicitation of the structure took approximately 50 hours with the experts (in addition to Dr. Wasyluk, we verified the parts of the model with Drs. Daniel Schwartz and John Dowling of the University of Pittsburgh). Access to the IFF dataset allowed us to learn all the numerical parameters of the model from data rather than eliciting them from experts.

4 Quantitative Comparison

In this section, we compare the two systems Hepar-RB and Hepar-BN in quantitative terms, using data from two different datasets mentioned earlier: the DH dataset from Rotterdam and the IFF dataset from Warsaw.

4.1 Patient Data

As the diagnostic categories distinguished in the systems Hepar-RB and Hepar-BN are different, so are the patient findings used in order to diagnose disorders. Both the IFF and DH datasets had to be converted to make comparison between the systems possible, which is not a straightforward process. First, we focused on identifying those medical findings that are common for them. Those attributes whose values were not mutually exclusive were broken down into several attributes. This is a simple consequence of the probabilistic constraint that the outcomes of a random variable must be mutually exclusive. We also assumed that missing values of attributes correspond to values *absent*, which is a popular assumption in case of medical datasets [12].

We identified 46 findings that were common for both systems. Due to different disorder mapping in both data sets, we found only two common disorders: *primary biliary cirrhosis* and *steatosis hepatis*. Of these two, primary biliary cirrhosis (PBC) was the only one with a reasonable number of patient cases in both data sets and, therefore, we focused on PBC in our experiment.

4.2 Results

The results of our experiment are presented in Tables 1(a) through 1(f). Our first observation was that in a significant portion of the cases Hepar-RB was not able to reach a conclusion. If the most likely disease is the one that the system recommends, Hepar-BN will always make a diagnosis. To make the comparison fair, we assumed that there is a probability threshold that has to be crossed in order for the system to make a diagnosis with a reasonable confidence. We would like to point out that in practice this threshold depends on the utility of correct diagnosis and misdiagnosis (which are disease-dependent). The threshold can be naturally introduced into a probabilistic system using decision-theoretic methods. In our experiments, we chose a fixed threshold of 50%, which is rather conservative towards Hepar-BN. Hepar-BN was not able to cross this threshold in a number of cases that is comparable to Hepar-RB.

Tables 1(a) and 1(b) summarise the results of both systems when tested with the 699 IFF patients. Similarly, Tables 1(c) and 1(d) present the results for Hepar-RB and Hepar-BN, respectively, for the 181 DH patients. In both cases, the performance of Hepar-BN was determined by cross-validation using the leave-one-out method. Table 1(e) gives the results for Hepar-BN trained on the DH dataset and tested with the IFF dataset, whereas the results of Table 1(f) were obtained the other way around. Note that the results of Table 1(e) are inferior to those given in Table 1(f).

Table 1. Comparison of diagnostic accuracy of HEPAR-RB (**H-RB**) and HEPAR-BN (**H-BN**); PBC+ and PBC− stand for PBC present and absent, respectively; UC: unclassified.

Results (a) and (b) for 699 IFF patients.

	Patients				
H-RB	PBC+	(%)	PBC−	(%)	Total
PBC+	174	(62)	15	(4)	189
PBC−	11	(4)	85	(20)	96
UC	95	(34)	319	(76)	414
Total	280	(100)	419	(100)	699

(a)

	Patients				
H-BN	PBC+	(%)	PBC−	(%)	Total
PBC+	263	(94)	61	(15)	324
PBC−	13	(5)	190	(45)	203
UC	4	(1)	168	(40)	172
Total	280	(100)	419	(100)	699

(b)

Results (c) and (d) for 181 DH patients.

	Patients				
H-RB	PBC+	(%)	PBC−	(%)	Total
PBC+	11	(73.3)	1	(1)	12
PBC−	2	(13.3)	143	(86)	145
UC	2	(13.3)	22	(13)	24
Total	15	(100)	166	(100)	181

(c)

	Patients				
H-BN	PBC+	(%)	PBC−	(%)	Total
PBC+	12	(80)	1	(1)	13
PBC−	3	(20)	65	(39)	68
UC	0	(0)	100	(60)	100
Total	15	(100)	166	(100)	181

(d)

Results of HEPAR-BN for (e) 699 IFF and (f) 181 DH patients.

	Patients				
H-BN	PBC+	(%)	PBC−	(%)	Total
PBC+	170	(60.7)	12	(2.9)	182
PBC−	9	(3.2)	76	(18.1)	85
UC	101	(36.1)	331	(79)	432
Total	280	(100)	419	(100)	699

(e)

	Patients				
H-BN	PBC+	(%)	PBC−	(%)	Total
PBC+	14	(93)	57	(34.3)	71
PBC−	1	(7)	71	(42.8)	72
UC	0	(0)	38	(22.9)	38
Total	15	(100)	166	(100)	181

(f)

5 Discussion

It seems that building the models in each of the two approaches has its advantages and disadvantages. One feature of the rule-based approach that we found particularly useful is that it allows testing models by following the trace of the system's reasoning. A valuable property of Bayesian network-based systems is that models can be trained on existing data sets. Exploiting available statistics and patient data in a Bayesian network is fairly straightforward. Fine-tuning a rule-based system to a given dataset is much more elaborate.

Rule-based systems capture heuristic knowledge from the experts and allow for a direct construction of a classification relation, while probabilistic systems capture causal dependencies, based on knowledge of pathophysiology, and enhance them with statistical relations. Hence, the modelling is more indirect, although in domains where capturing causal knowledge is easy, the resulting diagnostic performance may be good. Rule-based systems may be expected to perform well for problems that cannot be modelled using causality as a guid-

ing principle, or when a problem is too complicated to be modelled as a causal graph.

Our experiments have confirmed that a rule-based system can have difficulty with dealing with missing values: around 35% of the IFF patients remained unclassified by HEPAR-RB, while in HEPAR-BN only 2% of IFF patients remained unclassified. Note that this behaviour is due to the semantics of negation by absence, and in fact a deliberate design choice in rule-based systems. Refraining from classifying is better than classifying incorrectly, although it will be at the cost of leaving certain cases unclassified. In all cases, the true positive rate for HEPAR-BN was higher than for HEPAR-RB, although sometimes combined with a lower true negative rate.

Both systems were in general more accurate when dealing with their original datasets. The reason is that the systems were using then all available data, not only the common variables. We have noticed some indications of overfitting in case of HEPAR-BN, visible especially in those results, where the system was trained and tested on different data sets.

References

1. B.G. Buchanan and E.H. Shortliffe (Eds.). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, Reading, MA, 1984.
2. D. Heckerman. Probabilistic interpretations for Mycin's certainty factors. In: L.N. Kanal and J.F. Lemmer (Eds.). *UAI 1* (Elsevier, NY, 1986) 167–196.
3. M. Henrion. Some practical issues in constructing belief networks. In: L.N. Kanal, T.S. Levitt, and J.F. Lemmer (Eds.). *UAI 3* (Elsevier, NY, 1989) 161–173.
4. P.J.F. Lucas. Refinement of the HEPAR expert system: tools and techniques. *Artificial Intelligence in Medicine* 6(2) (1994) 175–188.
5. P.J.F. Lucas. Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review* 12(2) (1997) 109–146.
6. P.J.F. Lucas. Certainty-like structures in Bayesian belief networks. *Knowledge-based Systems* (2001).
7. P.J.F. Lucas and A.R. Janssens. Development and validation of HEPAR. *Medical Informatics* 16(3) (1991) 259–270.
8. P.J.F. Lucas, R.W. Segaar, and A.R. Janssens. HEPAR: an expert system for diagnosis of disorders of the liver and biliary tract. *Liver* 9 (1989) 266–275.
9. A. Newell and H.A. Simon. *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ, 1972.
10. A. Oniśko, M.J. Druzdzel, and H. Wasyluk. Extension of the Hepar II model to multiple-disorder diagnosis. In: S.T. Wierzchoń M. Kłopotek, M. Michalewicz (Eds.). *Intelligent Information Systems*. Advances in Soft Computing *Series* (Physica-Verlag Heidelberg, 2000) 303–313.
11. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA, 1988.
12. M. Peot and R. Shachter. Learning from what you don't observe. In: *UAI–98* (Morgan Kaufmann, San Francisco, CA, 1998) 439–446.
13. L.C. van der Gaag. *Probability-Based Models for Plausible Reasoning*. PhD thesis, University of Amsterdam, Amsterdam, The Netherlands, 1990.