Teaching and Learning Systems – The Role of AI in Past, Present, and Future

Workshop Proceedings

Edited by:
Alke Martens, University of Rostock, Germany
Andreas Harrer, University of Duisburg-Essen, Germany

Located at the
27th German Conference on Artificial Intelligence
September 20-24, Ulm, Germany

WS 4

15. Meeting of the GI Special Interest Group 'Intelligent Teaching and Learning Systems'

at the Konferenz KI 2004

KI 2004

Teaching and Learning Systems - The Role of Artificial Intelligence in Past, Present and Future.

Alke Martens, Andreas Harrer (Herausgeber)

Ulm, 24. September 2004

The Engineering and Evaluation of an Intelligent Problem-Oriented Learning Environment

Heiko Seebold, Andreas Lüdtke, Claus Möbus

Learning Environments and Knowledge Based Systems
Department of Computing Science, University of Oldenburg
D-26111 Oldenburg, Germany

heiko.seebold@offis.de
andreas.lüdtke@offis.de
claus.moebus@uni-oldenburg.de

Abstract: Bayesian belief networks (BBN) have become the representation of choice for building decision-making systems in domains characterized by uncertainty (e.g. medicine). For various reasons they are relevant for the success of intelligent systems in diagnostics, therapy planning and e-Learning.

BBN-models consist of a qualitative and a quantitative part. If objective probability data are unavailable subjective data (dependence and independence judgments, subjective probabilities) have to be acquired from domain experts instead. This is time consuming and puts much effort on experts, especially, when they are novices in modelling BBNs. There is a knowledge acquisition bottleneck, which hampers the introduction of new and the revision of old models.

To ease the acquisition of the qualitative data required to model a BBN we developed a new knowledge acquisition procedure. In this paper we present a case study in which we developed together with two expert cardiologists a first prototype BBN with n = 39 variables in a two day crash modelling workshop using our new greedy algorithm. The developed BBN was validated and it showed that only about 17% of the network structure had to be corrected due to slips in the judgements given by our experts. The BBN was integrated into a problem-oriented learning environment which confronts the students with cases from their daily routine. In an evaluation the students expressed mixed expectations regarding the role of the learning environment as a means for exam preparation. However we show how it successfully supports the well known concept of “Evidence-Based Learning” by Florian Eitel.

1 Introduction

Bayesian belief networks (BBN) are the representation of choice for building decision-making systems in domains characterized by uncertainty [Pe86; LS88; Ne90; Pe98; Sh96; CGH97; CDL99; Je01; Pe01; RN03] with applications in robotics [RN03], machine learning [Fr98], natural language processing [Ch93; MS02], medicine [Ma97], marketing [NS01], and psychology [Gl01].

For various reasons BBNs are relevant for the success of intelligent systems in diagnostics, therapy planning and eLearning. They are used for the representation of uncertain causal knowledge [e.g. FMS96; SMF96], testing hypotheses about diseases and treatment [He91], being learning object per se [HJK96], and assessing or modelling student knowledge [GSM94; Mi95; MG96; He00; MA00; BC02; ZG03].

The classical procedure for the construction of BBNs under the knowledge based approach was published by Pearl as the *boundary strata method* [Pe88, Pe98, p.119]. Because of its cognitive demanding aspects it is unsuitable for domain experts without modelling experience.

We designed a new method [MSL04] which worked very well with cardiologists even under severe time pressure. In the first step of the procedure experts are asked to judge the causal precedence in pairs of stochastic variables.

A new greedy algorithm for the anytime determination of transitive closures controls the selection of pairs, guarantees that the data comprise a partial order relation (POR) and generates the Hasse diagram of the POR (Hasse model). In the best case the monitor acquires the Hasse model of the causal precedence relation in one pass. Then the savings in pair-comparisons are (1-2/n)*100%, the judgement complexity is O(n) and the computational complexity is $O(n^3)$. If the Hasse model also passes a Markov blanket independence test, it is without further modifications the DAG of the BBN. In the worst case the monitor needs the full number of n(n-1)/2 comparisons. The judgement complexity is $O(n^2)$ and the computational complexity stays $O(n^3)$. If the Hasse model does not pass the Markov blanket test, experts think that influences (or links) are missing. These have to be added back to the Hasse model. The modified DAG is then considered as the qualitative model of the BBN. Despite its flexibility the computational complexity of the greedy algorithm is only $O(n^3)$.

In the first part of the paper we present a medical case study which shows how the new method was successfully used to develop a BBN for the disease of aortic stenosis. The knowledge acquisition for the complete model of the first prototype with 39 nodes (pair-comparisons, Markov blanket tests and estimation of conditional probability tables) could be accomplished in a two day crash workshop. In the second part we show how the BBN was integrated into an e-Learning system for problem oriented diagnostics in aortic stenosis – we call this system Kardiobayes. The paper closes with a presentation of our validation results where we argue that Kardiobayes supports the concept of "Evidence-based Learning" [Ei99].

1.1 Bayesian Networks

Bayesian networks capture independence and conditional independence where they exist. Among variables where dependencies exist, they encode the relevant portion of the full joint distribution. BBNs use a graphical representation, making it easier to investigate complexity and study inference algorithms. The formal definition of a BBN is [CGH97, p.248]:

"A Bayesian network model, or simply a Bayesian network, is a pair (D, P), where D is a DAG, $P = \{p(x_1 \mid \pi_1), \ldots, p(x_n \mid \pi_n)\}$ is a set of n CPDs (conditional probability distribution), one for each variable, and Π_i is the set of parents of node X_i in D. The set P defines the associated JPD (joint probability distribution) as

$$p(x) = \Pi_{i=1}^{n} p(x_i \mid \pi_i).$$

The DAG D is a minimal directed I-map of p(x)" in the sense, that no edge can be deleted without destroying its I-map character.

2 A Medical Case Study

2.1 The Domain

The new method was used to build a BBN model for the aortic stenosis disease. Aortic stenosis is the narrowing or obstruction of the heart's aortic valve, which prevents it from opening properly and blocking the flow of blood from the left ventricle to the aorta. It can either be congenital or acquired (our study). The BBN model was embedded into a problem based learning environment for students of cardiology.

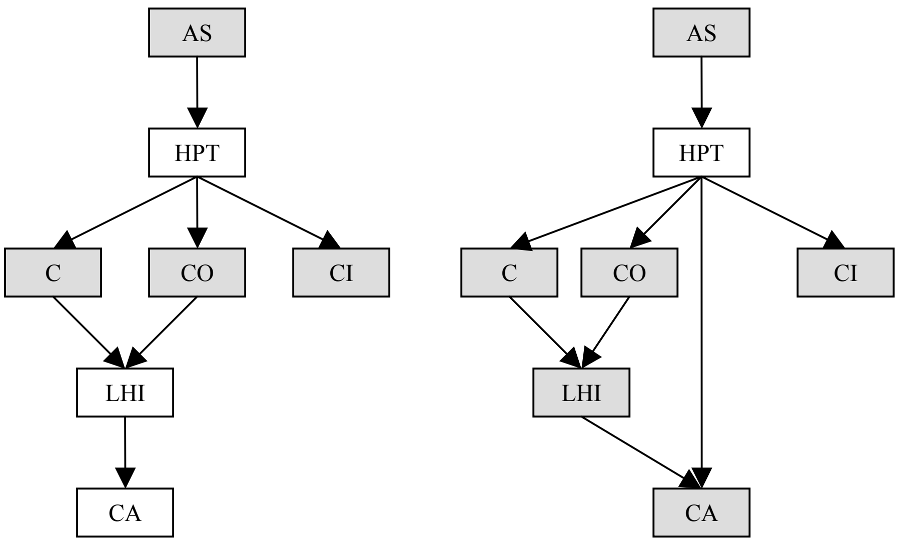
2.2 Construction of the Qualitative Model

Modelling of the qualitative part of the BBN consisted of the following steps:

1. Identification of the relevant medical concepts such as sequelas, causes, symptoms and examination methods: e.g. endocarditis as a cause, hypertrophy as a sequela and ascites as a symptom. The final model had 39 variables.
2. Partial order judgments concerning "precedes", "follows" or "none" (e.g.: hypertrophy precedes contractility, contractility precedes left-sided heart failure, which itself precedes cardiac arrhythmias) according a schedule generated by our greedy algorithm [MSL04].
3. Creating the Hasse diagram of the influence structure using a lattice drawing applet [La04].
4. Testing the Hasse diagram for missing direct influences using the Markov blanket test for each variable.

The Hasse Diagram constructed in the steps 1 to 3 represents the minimal data compliant D-Map [MSL04]. In step 4 the Markov blanket test is used to infer additional direct influences. The Markov blanket of a variable *v* is the set consisting of its direct predecessors, direct successors and the variables sharing a direct successor with *v*. A variable in a Bayesian network can by definition only be influenced by variables in its Markov Blanket. By checking for influences from variables outside the Markov blanket violations of the Bayesian network structure and thus additional necessary direct influences can be detected. It showed that only about 13 % of the overall network structure had to be modified due to the Markov blanket test.

For example using the Markov blanket for hypertrophy (see Figure 1a) it was identified that the variable cardiac arrhythmia (CA) should be influenced by hypertrophy (HPT), since hypertrophy could be a cause for cardiac arrhythmia. But as cardiac arrhythmia was outside of the Markov blanket (the shaded variables) this direct influence was not modelled in the network, a link between hypertrophy and cardiac arrhythmia had to be added (see Figure 1b). Checking for missing influences again showed that no other concept in the network outside the blanket influenced or was influenced by hypertrophy.



(a) (b)

Figure 1 - Markov Blanket (a) for Hypertrophy and “lost” link added (b)

2.3 Construction of the Quantitative Model

The quantification of the existing qualitative model was done in interviews with the medical experts according to existing methods [GR99; RW99; MH93]. To ease the specification of probabilities, we used a scale on which verbal probabilistic statements were correlated with numerical expressions. It turned out that our expert used the scale only in cases where they were not quite sure about the exact probabilities. Probabilities, which were certain, were specified directly without any help.

Different kinds of probabilities were specified: a-priori probabilities, probability matrices and restrictions. Since not every probability could be specified by the experts, missing probabilities were estimated using the principle of maximum entropy [MR96]. To this end, the existing information was transferred to SPIRIT [Sp04], which computed a minimal assumption model with the expert probabilities as hard constraints.

2.4 Validity of the Model

In telephone conferences the validity of the first prototyped BBN was tested. Evidences were entered into the network and the resulting probability changes were observed and evaluated. Discrepancies between the actual values and the expected values had two causes: (1) either incorrectly specified probabilities or (2) an error in the network structure.

(1) In the case of incorrect specified probabilities the medical expert had no knowledge of the exact probabilities or of the effect by combining evidence. These improper probabilities had to be specified anew.

(2) In case of a faulty network structure influences between the variables were modelled incorrectly. In discussions with the experts three different faults could be identified. The network structure were

- too specific: We solved the problem by aggregating several variables into a more general variable.
- too general: Too general influences were specialized by introducing more specific variables.
- wrong at all: Wrong influences were corrected by adding or removing links between the variables.

These errors could not be prevented by our new method, but the validation of the 39-variables-BBN showed that only a few corrections had to be done: only

- one generalization: The concepts of rheumatic endocarditis and bacterial endocarditis were generalized into the more general concept: endocarditis (Fig. 2).

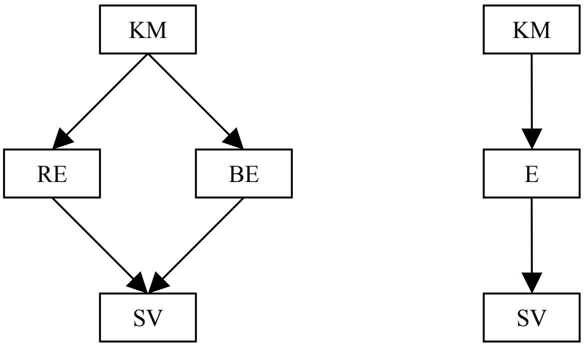


Figure 2 – Generalization of endocarditis

- one specialization: Testing the network an error in the relations between right-sided heart failure (RHI), ascites (A), jugular venous distension (HVS) and peripheral edema (POE) were uncovered. A new concept between right-sided heart failure and ascites had to be introduced: stasis liver (L) (Fig. 3).

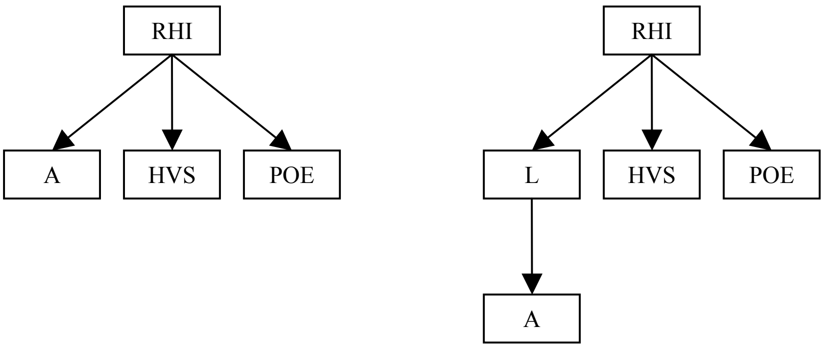


Figure 3 – Specialization of Ascites

- one missing link: A missing link representing the influence between aortic stenosis (AS) and contractility (C) was added (Fig. 4).

About 17% of the model had to be revised due to slips in the pair comparisons. Hence it seems that this new procedure is excellently suited for constructing the DAG of a BBN.

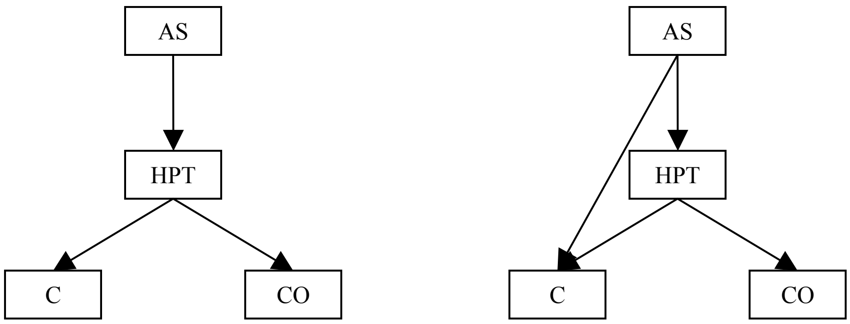


Figure 4 - Correction of a missing link

3 BBN Based Diagnostics in a Problem-oriented Learning Environment

The BBN was integrated into a problem-oriented learning environment, which we call Kardiobayes (see figure 5). The two main components of the learning environment are the BBN about aortic stenosis and a case-oriented task formulation.

The task specifies a situation, which could also occur during the daily routine of the learner. It includes information concerning medical history, results of examinations and symptoms (see figure 5 part 1). So the learner is confronted with a problem, which he should solve using the BBN.

The BBN, the second component, is available to the learner all the time while solving the task and supports the solution-finding-process (see figure 5 part 2). The learner has the possibility to enter the present information from the case and the problem context as evidence into the network. Regarding the changed probabilities of the network the learner is able to test various solution hypotheses and to choose the most appropriate. In addition the BBN can be used to freely explore the prevalent relations between the medical concepts. This way the learner is able to gain new insights about the modelled disease, in our case: aortic stenosis.

In combination the two components allow the student greatest possible freedom while using the learning environment. On the one hand the student is able to freely explore the BBN and learn the prevalent relations. On the other hand he can be directed by the case-oriented tasks.

The learning environment supports two different problem-types: problems which could be solved

- by multiple-choice: A set of solutions is presented to the learner, from which he has to select the right answer (see figure 5 part 3). E.g. the learner has to choose from a list of five different symptoms the one, which mostly supports the diagnosis of aortic stenosis the most.
- using the Bayesian Network: The learner has to enter the right solution into the Bayesian Network. The solution of this kind of problem is a combination of different medical concepts, which are entered as evidence into the network. E.g. the learner has to find a combination of causes and symptoms, which result in an 80% likelihood for aortic stenosis.

After the learner found a solution, his answer is compared to the right solution stated by the experts. Depending on the given solution the learner gets a message about the grade of his success or failure.

The learning environment supports the training of two different knowledge types:

- Domain-dependent knowledge. The relevant medical concepts of the considered disease and their relations are learned. (E.g. symptoms, causes, etc.) The students learns which diagnostic information discriminates most between alternative diseases in different medical cases (e.g. how strong does the symptom contractility support the diagnosis of aortic stenosis, when hypertrophy is already known)

- Domain-independent knowledge. The strategic diagnostic skills of the students are trained. On the one hand this is done by presenting naturalistic diagnostic tasks which train the diagnostic skills directly, on the other hand by visualizing the interconnections and complex dependencies between the medical concepts by the BBN.

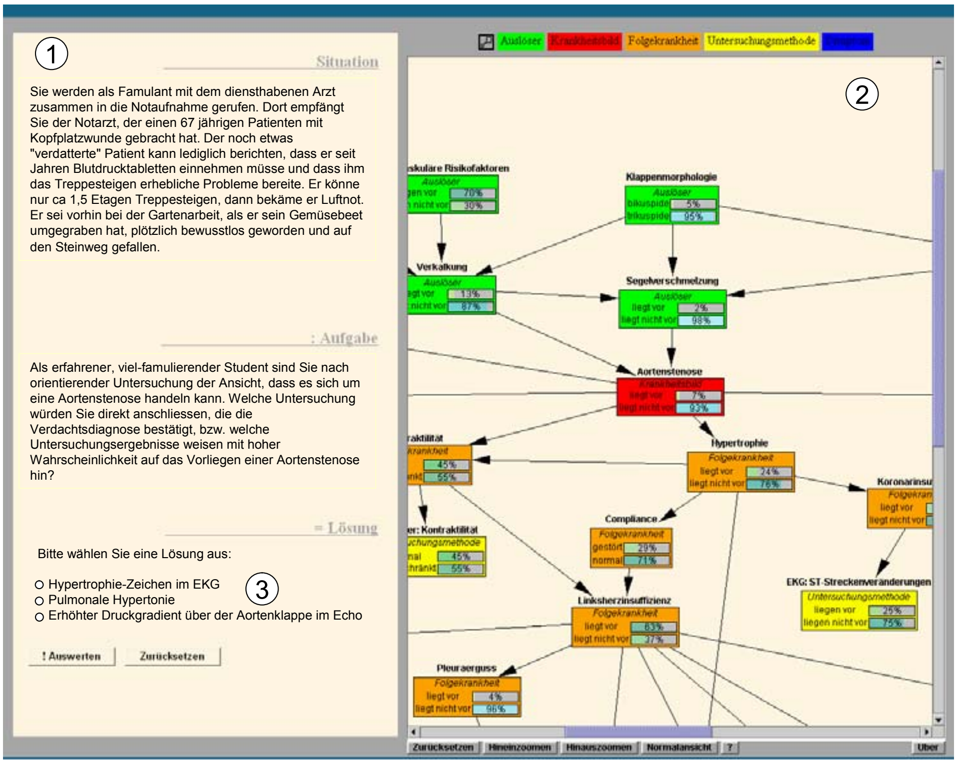


Figure 5 - Kardiobayes

4 Evaluation of the Learning Environment

The learning environment was evaluated by eight students of medicine at the RWTH Aachen and six at the Uniklinikum Münster. To evaluate the acceptance of the learning environment we chose a pre-/post test-design. By this method it was possible to identify significant changes in the acceptance.

The same questionnaire with questions about the acceptance and expectation concerning different kinds of learning media (books, teachers, Kardiobayes) were handed to the students before and after the training with Kardiobayes. Similar questions were asked for the different kinds of media. For example the students had to answer the three questions “A tutor/book/Kardiobayes is able to support in me preparing for my exams.” The students could rate this statement from 1 (I agree) to 5 (I disagree). Using statistical methods (t-tests) the results were examined for significant changes with regard to the acceptances and expectations.

A second questionnaire was handed to the students only after the test. This questionnaire contained general questions about the learning environment. The students were able to rate the learning environment with respect to its usability, the comprehensibility of the task formulations and its documentation.

The evaluation of the questionnaires showed that almost all students (92%) considered the uncertain knowledge represented by the BBN to be important for their profession and agreed on the fact that BBN represented a novel point of view.

Although the students stated the novelty and importance of the learning environment, they expressed mixed expectations regarding the role of the BBN as an exam preparation. The number of students who felt the learning environment as unsuitable for exam preparation increased from 7% to 20% after training with the learning environment.

From the evaluation of the questionnaires and the discussion with the students after the training two main reasons for the scepticism could be identified. One reason is difficulty to understand the complex representation of the BBN. Thus an initial training by the students is required to comprehend every aspect of the representation.

Another reason is the nature of the exams in medicine, which is mainly a multiple-choice-questionnaire to test the knowledge of the students. However Kardiobayes supports evidence based learning (EBL) [Ei99], which aims among others at finding an evidence based guide to enhance the students' performance in medical practice. EBL consist of several interdependent learning-steps performed by a small group of students:

- A group of students is confronted with a diagnostic problem. Each student finds an individual solution for the problem, the so called 'individual standard'. This is done in Kardiobayes by presenting the task to the students and giving him the opportunity to specify the solution in the BBN. No feedback is given to the student at this time.

- The different individual standards are brought together to form a 'group standard'. Every student presents his BBN and in a moderated discussion a consensus is developed, the 'group standard'.

- The 'group standard' is validated against best evidence resulting in the 'evidence based standard'. Using Kardiobayes this could be achieved by entering the 'group standard' into the BBN. Since our BBN represents the knowledge of the experts it can serve as the best evidence, expert evidence in this case. Kardiobayes checks if the group standard forms a sound solution, the 'evidence based solution'.

Thus Kardiobayes supports these three very important steps of the EBL in serving as a basis to form the individual, group and evidence based standards.

5 Summary

Under the supervision of two cardiologists we succeeded in developing a BBN prototype model with 39 nodes in a two day crash-workshop. The BBN prototype was developed using our new greedy algorithm. Validation of the BBN showed that only 18% of the network had to be corrected due to slips in the pair comparisons. We embedded the BBN model in a problem based learning environment, named Kardiobayes, so students have the opportunity to improve their diagnostic skills in an authentic goal based learning scenario. The students have the opportunity to freely explore the BBN or to be guided by our learning environment. In an evaluation the importance of the presented knowledge was confirmed by the students. Nevertheless the learning environment appeared to the students as unsuitable for their exam preparations. However have shown in the paper how Kardiobayes is able to support the well-known concept of "Evidence-Based Learning".

Acknowledgements

We would like to thank the medical experts Dr. Dierk Rulands and Florian Reinke. Also we are indebted to Heinz-Jürgen Thole for supporting us in the conception as well as implementation of Karidobayes.

References

[BC02] Bunt, A., Conati, Ch.: Assessing Effective Exploration in Open Learning Environments Using Bayesian Networks, in: St. A. Cerri, G. Gouardères, F. Paraguacu (Eds.), Intelligent Tutoring Systems, Berlin: Springer, 698 – 707, 2002

[CDL99] Cowell, R.G., Dawid, A.P., Lauritzen, St.L., Spiegelhalter, D.J.: Probabilistic Networks and Expert Systems, Berlin: Springer, 1999

[CGH97] Castillo, E., Gutiérrez, J.M., Hadi, A.S.: Expert Systems and Probabilistic Network Models, Berlin: Springer, 1997

[Ch93] Charniak, E.: Statistical Language Learning, Cambridge, Massachusetts: MIT Press, 1993

[Ei99] Eitel, F.; Steiner S.: Evidence-based Learning, in: Medical Teacher 21 (5), 506 – 512, 1999

[FMS96] Folckers, J., Möbus, C., Schröder, O. & Thole, H.J.: An Intelligent Problem Solving Environment for Designing Explanation Models and for Diagnostic Reasoning in Probabilistic Domains, in: C. FRASSON, G. GAUTHIER, A. LESGOLD (eds), Intelligent Tutoring Systems, ITS 96, Montreal, Canada, Proceedings, p. 353-362, Berlin: Springer (LNCS 1086), 1996

[Fr98] Frey, B.J.: Graphical Models for Machine Learning and Digital Communication, Cambridge, Massachusetts: MIT Press, 1998

[GC93] Goldman, R.P. & Charniak, E., A language for construction of belief networks, IEEE Transactions of Pattern Analysis and Machine Intelligence, 15 (3), 1993, 196 – 208

[Gl01] Glymour, C.: The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology, Cambridge, Massachusetts: MIT Press, 2001

[GR99] Van der Gaag, L.C., Renooij, S., Witteman, C., Aleman, B.M.P., Taal, B.G.: How to elicit many probabilities, in: Proceedings of the 15th Conference of Uncertainty in AI (UAI-99), 1999, pp. 647-654

[GSM94] Gitomer, D. H., Steinberg, L. S., Mislevy, R.J.: Diagnostic Assessment of Troubleshooting Skill in an Intelligent Tutoring System, Princeton, N.J.: Educational Testing Service, RR-94-21-ONR, April 1994

[HJK96] Haddaway, P., Jacobson, J., Kahn, Ch. E.: BANTER: A Bayesian Network Tutoring Shell, Lab Report, September 15, 1996, Decision Systems and Artificial Intelligence Lab, Department of EE & CS, University of Wisconsin-Milwaukee

[He89] Henrion, M.: Some Practical Issues in Constructing Belief Networks, in: J.F. Lemmer & L.N. Kanal (Eds), Uncertainty in Artificial Intelligence, 3, 161 – 173, Amsterdam: Elsevier, 1989

[He91] Heckerman, D., Probabilistic Similarity Networks, Cambridge, Mass.: MIT Press, 1991

[He00] Henze, N., Adaptive Hyperbooks: Adaption for Project-Based Learning Resources, University of Hannover, Germany, Ph. D. Thesis, 2000

[Je01] Jensen, F.V.: Bayesian Networks and Decision Graphs, Statistics for Engineering and Information Science, Berlin: Springer, 2001

[La04] Lattice Drawing, 2004, URL: http://www.math.hawaii.edu/~ralph/LatDraw/ (March 18, 2004)

[LM97] Laskey, K.B. and Mahoney, S.M.: Network fragments: Representing knowledge for constructing probabilistic models, in D. Geiger and P. P. Shenoy, (eds.), Uncertainty in Artificial Intelligence: Proceedings of the Thirteenth Conference, Palo Alto, CA: Morgan Kaufmann, 1997

[LS88] Lauritzen, S.L.; Spiegelhalter, D.J.: Local computations with probabilities on graphical structures and their application to expert systems, Journal of Royal Statistical Society Series B 50 (2), 1988, 157-224

[MA00] Mislevy, R.J., Almond, R. G., Yan, D., Steinberg, L.S.: Bayes Nets in Educational Assessment: Where Do the Numbers Come From ?, Center for the Study of Evaluation, University of California, Los Angeles, CSE Technical Report 518, March 2000

[Ma97] Mannebach, H.: Die Struktur des ärztlichen Denkens und Handelns: Ein Beitrag zur Qualitätssicherung in der Medizin, Weinheim: Chapmann & Hall, 1997

[MG96] Mislevy, R.J., Gitomer, D. H.: The Role of Probability-Based Inference in an Intelligent Tutoring System, User-Modeling and User-Adapted Interaction, 1996, 5, 253-282

[MH93] Morgan, M. G. & Henrion, M.: Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis, - Repr. - Cambridge [u.a.]: Cambridge University Press, 1993

[Mi95] Mislevy, R.J.: Probability-based inference in cognitive diagnosis, in P. Nichols, S. Chipman & R. Brennan (Eds.), Cognitively diagnostic assessment, 43 – 71, Hillsdale, NJ: Lawrence Erlbaum Associates, 1995

[MR96] Meyer, C.-H., & Rödder, W.: Probabilistic Knowledge Representation and Reasoning at Maximum Entropy by Spirit, in: KI-96: Advances in Artificial Intelligence Proceedings of the 20th Annual German Conference on Artificial Intelligence (Ed.: S. Hölldobler), Berlin: Springer, 1996

[MS02] Manning, Ch.D., Schütze, H.: Foundations of Statistical Natural Language Processing, Cambridge, Massachusetts: MIT Press, 2002

[MSL04] Möbus, C.; Seebold, H.; Lüdtke, A.; Thole, H.-J.: Constructing the DAG of Bayesian Belief Networks on the Basis of Simple Pairwise Causal Precedence Judgments, 2004, (submitted to Artificial Intelligence in Medicine)

[Ne90] Neapolitan, R.E.: Probabilistic Reasoning in Expert Systems: Theory and Algorithms, New York: Wiley, 1990

[NS01] Nadkarni, S. & Shenoy, P.P.: A Bayesian Network Approach to Making Inferences in Causal Maps, European Journal of Operational Research, 128, 479 – 498, 2001

[NS04] Nadkarni, S. & Shenoy, P.P., A Causal Mapping Approach to Constructing Bayesian Networks, http://lark.cc.ku.edu/~pshenoy/papers/DSS03b.pdf (March 03, 2004), Decision Support Systems, 2004 (in press)

[Pe86] Pearl, J.: Fusion, propagation, and structuring in belief networks, Artificial Intelligence, 29, 1986, 241-288

[Pe98] Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference (Revised Second Printing), Morgan Kaufman Publishers, San Mateo, CA., 1998

[Pe01] Pearl, J.: Causality: Models, Reasoning, and Inference, Cambridge, UK: Cambridge University Press, 2001

[RW99] Renooij, S., Witteman, C.: Talking probabilities: communicating probabilistic information with words and numbers, in: International Journal of Approximate Reasoning, vol. 22, 1999, pp. 195-215.

[RN03] Russell, St., Norvig, P.: Artificial Intelligence. A Modern Approach, Upper Saddle River, N.J.: Pearson Education, Inc., 2003

[Sh96] Shafer, G.: The Art of Causal Conjecture, Cambridge, Massachusetts: MIT Press, 1996

[SMF96] Schröder, O., Möbus, C., Folckers, J. & Thole, H.J., Supporting the Construction of Explanation Models and Diagnostic Reasoning in Probabilistic Domains, in D.C. EDELSON, E.A. DOMESHEK (eds): International Conference on the Learning Sciences ICLS 96, Northwestern University, Evanston, IL, USA, Proceedings of the ICLS 96, p. 60-67, Charlottesville, VA: Association for the Advancement of Computing in Education (AACE), 1996

[Sp04] Spirit, 2004, URL: http://www.fernuni-hagen.de/BWLOR/spirit/ (March 18, 2004)

[ZG03] Zapata-Rivera, J.D., Greer, J.E., Student Model Accuracy using Inspectable Bayesian Student Models, in: U. Hoppe, F. Verdejo, J. Kay (Eds.), Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies, Amsterdam: IOS Press, 65 – 72, 2003