An Intelligent Problem Solving Environment for Designing Explanation Models and for Diagnostic Reasoning in Probabilistic Domains

Jörg Folckers, Claus Möbus, Olaf Schröder, Heinz-Jürgen Thole[1]

OFFIS Institute, Escherweg 2, D - 26121 Oldenburg, Germany

E-Mail: {folckers, moebus, schroeder, thole}@informatik.uni-oldenburg.de

Abstract. MEDICUS[2] is an Intelligent Problem Solving Environment (IPSE) currently under development. It is designed to support i) the construction of explanation models, and ii) the training of diagnostic reasoning and hypotheses testing in domains of complex, fragile, and uncertain knowledge. MEDICUS is currently developed and applied in the epidemiological fields of environmentally caused diseases and human genetics. Uncertainty is handled by the Bayesian network approach. Thus the *modelling task* for the learner consists of creating a Bayesian network for the problem at hand. He / she may test hypotheses about the model, and the system provides help. This differs from existing reasoning systems based on Bayesian networks, i.e. in medical domains, which contain a built-in knowledge base that may be used but not created or modified by the learner. For supporting *diagnostic reasoning*, MEDICUS proposes diagnostic hypotheses and examinations. This will be extended to support learners' acquisition and training of diagnostic strategies.

1 Introduction

Diagnosis is a reasoning and problem solving task that can be quite difficult. This is especially true in medical domains (Barrows & Tamblyn, 1980; Boshuizen & Schmidt, 1992; Elstein et al., 1978; Elstein & Bordage, 1980; Patel & Groen, 1986) where the knowledge is particularly complex, interrelated, fragile, and uncertain. Two examples of such domains are the epidemiology of diseases caused by environmental influences, like pollution, and of diseases caused by genetic defects. In these domains, a clear-cut taxonomy of diseases, or syndromes, has not been developed yet. Still these domains are getting increasingly important. This is reflected by the fact that these fields receive increasing attention in medical science at university as well as in further education training courses and postqualification courses for physicians.

Computer-based support of medical reasoning started more than twenty years ago. From the beginning, the problem of uncertainty received central attention. Since there were no efficient algorithms for processing probabilities, early systems like MYCIN

---

[2] $\underline{M}$odelling, $\underline{e}$xplanation, and $\underline{di}$agnostic support for $\underline{c}$omplex, $\underline{u}$ncertain $\underline{s}$ubject matters

(Shortliffe, 1976), CASNET (Weiss et al., 1978), PIP (Szolovits & Pauker, 1978; 1993), or INTERNIST (Miller et al., 1982) used heuristic approaches. This situation changed in the 1980's, enabling the development of normative probability-based medical expert systems (i.e., NESTOR, Cooper, 1984; MUNIN, Andreassen et al., 1987; PATHFINDER, Heckerman, 1991). The main aim of these systems is to provide the user with diagnostic hypotheses, given the available evidence, and to suggest further diagnostic evidence gathering steps, for example, for differential diagnosis. Some systems, like CASNET, also generate therapeutic recommendations. But in spite of some capability to explain their reasoning steps, the reasoning and knowledge structures of these systems remain largely hidden to the user.

For the purpose of medical training, the recommendation of diagnostic hypotheses and investigations is important but not sufficient. Since hypothetical model building and diagnostic reasoning are problem solving tasks, there is a need for students of medicine to *train* these two skills (Barrows & Tamblyn, 1980):

Firstly, apart from diagnosis the learner should have an opportunity to *actively construct* models of diseases, their possible causes, and the symptoms associated with them, and to evaluate the consequences of these models. In this way the learner acquires and uses the knowledge necessary for diagnostic reasoning.

Secondly, the learner should be given the opportunity to *actively perform* diagnostic reasoning and to apply diagnostic strategies.

Currently we develop MEDICUS, an Intelligent Problem Solving Environment (IPSE, Möbus, 1995), in collaboration with several medical institutions (Health Authority of Oldenburg, Documentation and Information Center for Environmental Issues, Osnabrück, Medical Institute for Environmental Hygiene, Düsseldorf, Robert-Koch Institute, Berlin). MEDICUS differs from existing medical expert systems by being designed to support the two activities mentioned:

• *Model construction* is supported by a model editor based on a simplified natural language. After creating an initial model of the domain of interest, the learner may further specify, evaluate, and revise the model at a qualitative and quantitative level.

• *Training of diagnostic strategies* will also be supported qualitatively (i.e., what information is necessary in order to support or differentiate between what hypotheses?) and quantitatively (i.e., how does information gathered affect my diagnostic hypotheses? What is the most important information to acquire next?)

Though MEDICUS is developed for fields of medicine, our intention is that it will be applicable to diagnosing and modelling problems in other domains of uncertain and complex knowledge as well. The next section gives an overview of the design decisions for the system. Then the implementation state is described. Conclusions and directions of further work will be sketched in the closing section.

2 Design Decisions for MEDICUS

In order to create a system designed to support problem solving in a knowledge domain, a set of design principles is required that is based on a theory of problem solving and knowledge acquisition. We call our approach an *Intelligent Problem Solving Environment* (IPSE, Möbus, 1995): The learner acquires knowledge by actively *testing hypotheses*. This means that the learner creates solution proposals for

problems, tests hypotheses about their correctness, and the system analyzes the proposals making use of an oracle or an expert knowledge base, and provides help and explanations. The psychological foundation of our IPSE approach is the ISP-DL Theory of knowledge acquisition and problem solving (Möbus, 1995) which is influenced by theoretical positions of van Lehn (1988), Newell (1990), and Anderson (1993), and Gollwitzer (1990). Briefly, it states that new knowledge is acquired as a result of problem solving: applying weak heuristics in response to impasses. In contrast, existing knowledge is optimized if applied successfully. Furthermore there are four distinct problem solving phases: deliberating, resulting in setting a goal, planning how to reach the goal, executing the plan, and evaluating the result. The ISP-DL Theory leads to several design principles for IPSE's. Some of them are (Möbus, 1995; Möbus et al., 1992):

• According to the theory the learner will look for and appreciate help at an impasse. So the system should not interrupt the learner but offer help on demand.

• Feedback and help information should be available any time, aiming at the actual problem solving phase of the learner.

• At an impasse, the learner should be prevented from trapping into follow-up impasses. Thus help should refer to the learner's pre-knowledge as much as possible.

MEDICUS is designed according to these criteria. Help information is or will be always available on demand. Planning a model is facilitated by the simplified-natural-language model editor which allows the learner to state her or his ideas in an informal way. The evaluation of models is supported qualitatively and quantitatively. Close correspondence to the learner's knowledge will be achieved by giving help that changes the learner's proposal as little as possible (minimal modifications).

Besides general design criteria for IPSE's, each domain has special features. In this case, it is the uncertainty of knowledge, which we chose to handle by the Bayesian network approach. A Bayesian network (e.g., Neapolitan, 1990; Pearl, 1988) represents knowledge as a set of propositional variables and probabilistic interrelationships between them by a directed acyclic graph. The variables are represented by the nodes of the graph, and the relations by directed arcs. The relations are conditional probabilities (each variable conditioned on its parents in the network). They define a joint probability distribution that can be used to compute desired aposteriori distributions. Figure 1 shows a simple Bayesian network for the joint distribution p(injury, space requirement, haemodynamic irritation, vomiting, dislocation of vessels, infection of meninges, permanent headache). The conditional distributions needed to compute this joint distribution can be simplified by independence assumptions. In the graph, independencies between variables are represented by omitting arcs. For example, in Figure 1 the variables "vomiting" and "dislocation of vessels" are independent given knowledge about "space requirement". This means that information about "vomiting" is not relevant for the hypothesis "dislocation of vessels" (and vice versa) if it is already known whether the patient suffers from a space requirement process: p(dislocation of vessels | space requirement, vomiting) = p(dislocation of vessels | space requirement).

An important reason for choosing the Bayesian network approach is that the system is designed to support qualitative reasoning. A physician engaged in medical diagnosis proceeds in a highly selective manner (i.e., Elstein et al., 1978). We pursue

the hypothesis that this selectivity corresponds to the kind of (in)dependencies present in Bayesian networks. Reviews of published case studies in the domain of environmental medicine support this hypothesis. More generally, there is empirical evidence that qualitative reasoning by (in)dependencies as supported by Bayesian networks (like for example "explaining away") corresponds closely to human reasoning patterns (Henrion, 1987; Pearl, 1993; Waldmann & Holyoak, 1992).
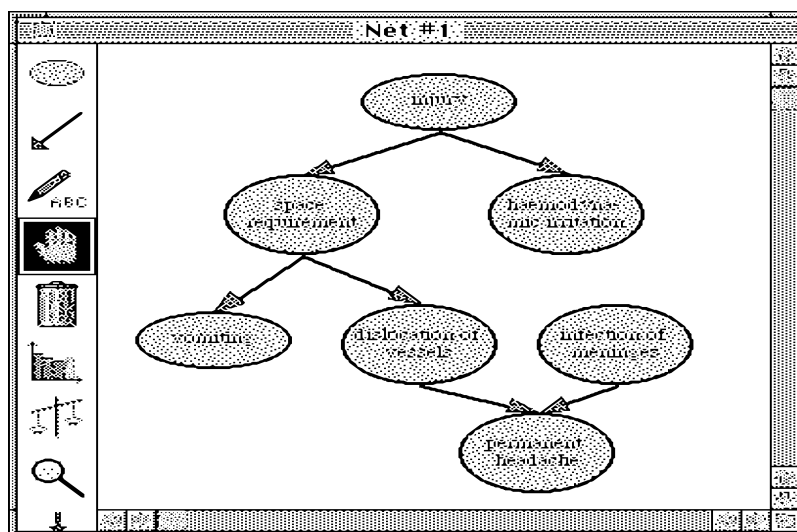

Fig. 1. A simple Bayesian network for a medical problem

Alternative approaches. The Dempster-Shafer Theory of evidential reasoning is an extension of probability theory. Uncertainty is represented by belief intervals limited by the positive belief in an event and its plausibility. This requires additional knowledge acquisition and explanation facilities. So at the current stage we consider the probability-based approach sufficient for our aims, which is not meant to preclude other methods at later stages. Fuzzy-Set Theory reasons with propositions that have vague meaning. In our application domain, there are many vaguely specified concepts, like "severe headache" or "typical symptom" (see also the examples below). Therefore we work on integrating fuzzy concept descriptions into the Bayesian network approach, that is, to generate conditional distributions for fuzzy relations.

3 The Implementation State of MEDICUS

This section is organized according to the two functions of MEDICUS: supporting the construction of explanation models, and training diagnostic reasoning.

3.1 Supporting Model Construction

Model construction is supported in three steps:

Initial Model Formulation

One of the main goals of MEDICUS is the assistance of the learner in developing a model of perceived causes, effects, and other relationships in a domain of interest with a formal tool, Bayesian networks. The reason to use a formal tool is to be able to derive consequences (dependencies, independencies, aposteriori distributions) that can be used for proposing recommendations and modifications. But at the same time it is necessary that a mathematically untrained learner is able to state his ideas in an informal way which he is used to. Therefore, we developed a simplified-natural-language model editor. After stating his model in this editor, the system can generate an initial graph automatically.

Figure 2 shows a small example with four sentences. Each sentence is placed in a sentence field. In order to create sentences, the learner may use variable categories (like state, event, action, object, substance), relations, modifier, and logical junctions. After placing a variable category in a sentence field, it can be named.
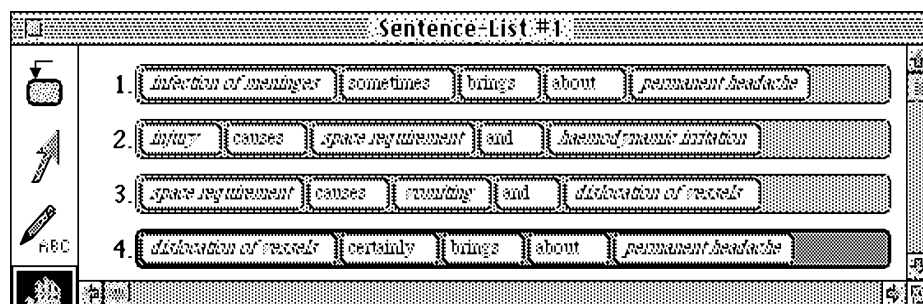

Fig. 2. Four sentences created in the simplified-natural-language model editor

For the relations there is a taxonomy based on i) probabilistic concepts of causality (Schurz, 1988; Suppes, 1970) organized according to the two dimensions "kind of influence" (positive / negative) and "direction of influence" (forward, backward, or undirected), and ii) has-part / is-a hierarchies. Table 1 shows the taxonomy. Relations currently available in the model editor are marked by asterisks. They were selected initially as a result of two sessions of discussing a topic of environmental medicine with an expert, but the user of MEDICUS may specify additional relations.

The sentences created by the learner are checked by a definite clause grammar. Besides syntactical correctness, semantic restrictions between variable categories and relations are checked. (For example, a state cannot cause a substance.) The learner receives feedback if the grammar finds errors.

The learner may ask the system to create a graph representation for the model specified. For the sentences in Figure 2, MEDICUS proposes the graph shown in Figure 1. The graph is an initial heuristic proposal that may have to be refined by the learner qualitatively, getting assistance from the system (see below). The graph is created in the following way:

• *Generating nodes.* Nouns, that is, variable categories named by the user, are represented by nodes (propositional variables). A proposition is assigned to each possible node instantiation. For example, if "injury" in Figure 2 has been categorized

as an "event" by the user, then the node "injury" is created (Figure 1) as a binary-valued node (unless specified otherwise by the user), and the propositions "Person experiences the event 'injury'" and "Person does not experience the event 'injury'" are assigned to it. The propositions can be inspected by the user.

• *Generating edges*. The relations between nouns are represented by links as depicted in the rightmost column of Table 1. For relations describing undirected relations (like "corresponds to"), a dialog is called where the learner is asked to specify the direction of influence, or to specify a third variable as the common cause or effect of the two variables in question.

In order to facilitate understanding of the graph representation further, natural-language expressions of the conditional distributions for each node (resp. apriori distributions for root nodes) are generated. Furthermore, the learner may ask for an explanation of the relationship between the sentences in the model editor and the net. Alternatively, a net can be created directly by editing and naming nodes and edges.
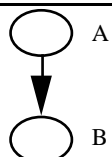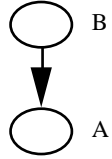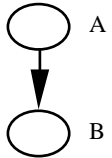
		Directed and undirected relations:		Representation in the graph:
		kind of influence		
		positive	negative	
direction of influence	forward $tA \leq tB$	A causes B* A brings about B* A triggers B* $p(B \mid A) > p(B)$	A counteracts B* A prevents B $p(B \mid A) < p(B)$	A B
	backward $tA \geq tB$	A follows B A is consequence of B ... $p(A \mid B) > p(A)$	A does not follow B ... $p(A \mid B) < p(A)$	B A
	undirected	A corresponds to B* A occurs with B*	A and B are mutually exclusive	Dialog required
		Is-a and Part-of hierarchies:		Representation in the graph:
		A is example for B* A contains B* $p(B \mid A) = 1$ A is exemplified by B A is part of B $p(B \mid \neg A) = 0$		A B

Table 1. Taxonomy of relations used in the model editor

Qualitative Model Revision

After the initial formulation of the model, it has to be analyzed and revised on a qualitative level. In particular, it has to be verified that the dependencies and independencies implied by the graph correspond to the assertions stated by the

modeller. For example, the graph in Figure 1 states that space requirement and haemodynamic irritation are independent, given injury, that is, p(space requirement | injury) = p(space requirement | injury, haemodynamic irritation). This means that knowledge about a haemodynamic irritation is not relevant for the hypothesis "space requirement", if it is known that an injury took place or not. Information about a haemodynamic irritation is useless for the hypothesis "space requirement", given knowledge about injury. In contrast, if nothing is known about injury, information about a haemodynamic irritation is useful for the hypothesis "space requirement": p(space requirement) ≠ p(space requirement | haemodynamic irritation). Similarly, vomiting and dislocation of vessels are independent, given space requirement. In contrast, dislocation of vessels and infection of meninges are dependent, given permanent headache. For example, if it is known that a patient suffers from permanent headache, then if new evidence arrives that weakens the hypothesis "dislocation of vessels", the hypothesis "infection of meninges" will be strengthened, and vice versa: Weakening one explanation for "permanent headache" strengthens the other one, and vice versa. Formally, conditional independence is described by the d-separation criterion (Pearl, 1988).

The knowledge of the modeller has to be acquired by the system in a way that is at the same time comfortable to the modeller and informative for generating independence assertions. Therefore, a second variant of knowledge acquisition (besides the model and net editor) is needed that can be used for model *construction* or for model *validation*, that is, for verifying or rejecting the independencies inherent in the graph. The system will offer a diagnostic dialog that proceeds in three steps. The current, yet to-be-improved dialog is described here:

1. For a case, the modeller specifies the initially known data and symptoms (left window in Figure 3: for example, "injury" and "dislocation of vessels"). Next, he specifies a diagnostic hypothesis (middle window in Figure 3, for example "space requirement"). Thirdly, he specifies what information he would look for next, that is, what information he considers relevant for his hypothesis (right window in Figure 3: for example "haemodynamic irritation" and "vomiting"). Independencies are constructed from this dialog in the following way: Information not considered relevant to the hypothesis by the modeller, given the case data and symptoms, is independent of the hypothesis. In Figure 3, "permanent headache" was *not* selected by the user in the right window, so "permanent headache" and "space requirement" are considered independent, given "injury" and "dislocation of vessels": p(space requirement | injury, dislocation of vessels, permanent headache) = p(space requirement | injury, dislocation of vessels).

2. The modeller states the hypothesis that the graph is consistent with the information specified by her or him in the diagnostic dialog. The system analyzes this hypothesis using the d-separation criterion. If differences are found, a graph is constructed internally (using the algorithm of Srinivas et al., 1990) from the dependence and independence assertions acquired in the diagnostic dialog (provided they are consistent). This internal graph is compared to the modeller's graph. This may lead to the result that i) graph and in-/ dependencies are consistent, or ii) edges have to be removed from the graph in order to be consistent with the in-/ dependencies, or iii) edges have to be added to the graph (this is the feedback for the

net in Figure 1 after the dialog in Figure 3 has taken place), or iv) edges have to be removed from as well as added to the graph in order to be consistent with the in-/ dependencies.

3. On further request, the modeller may ask the system for modification proposals and their explanation. For example, for the net in Figure 1 and the dialog in Figure 3, MEDICUS recommends to add an edge from "space requirement" to "haemodynamic irritation" because in the dialog the user considered "haemodynamic irritation" informative for "space requirement" given "injury" and "dislocation of vessels".

Fig. 3. Diagnostic dialog for the acquisition of information about independencies

Quantitative Model Specification

After the qualitative model revision, the modeller may quantify the net with apriori and conditional probabilities and let MEDICUS generate marginal distributions. Additionally, the user may enter evidences and let MEDICUS generate posterior distributions. Like for example in ERGO and HUGIN, evidence propagation is implemented according to the Lauritzen & Spiegelhalter (1988) algorithm.

3.2 Training Diagnostic Reasoning

The second purpose of MEDICUS is to assist in training diagnostic reasoning and strategies. Training of diagnostic strategies will also be supported qualitatively (i.e.: What information is necessary in order to support or differentiate between what hypotheses? How does information gathered affect my diagnostic hypotheses? What do I have to do in order to diagnose for example "space requirement"? What is the most important information to acquire next? When is it appropriate to consider an environmental etiology of the patient's complaints, and when is it not?) and quantitatively (i.e.: How much does a certain piece of evidence contribute to the currently best hypothesis? What evidence would be necessary in order to achieve a certain probability level for a diagnostic hypothesis?).

Currently MEDICUS is able to generate qualitative recommendations. For this purpose, the user has to specify whether propositional variables refer to syndromes, symptoms, external influences, or the general living and working environment. The

most probable syndromes in the light of the available evidence are presented to the learner. For example, in Figure 1, "injury" might have been specified as available evidence. After this evidence, the most probable syndrome hypotheses (i.e. the syndromes with the highest a posteriori probabilities) are "space requirement" and "infection of meninges". The system now recommends to examine those symptoms that are yet unknown and that depend on these syndromes directly, namely, "vomiting" and "permanent headache". These qualitative recommendations have been successfully demonstrated with a multiply connected net containing about fifty variables to a community of environmental medicinal professionals.

4 Conclusions and Further Work

The intelligent problem solving environment, MEDICUS is designed to support the construction of explanatory models in complex, uncertain domains, and to support diagnostic reasoning. Our example application domains stem from medicine. Unlike most existing systems based on the Bayesian network approach, MEDICUS is designed as a problem solving tool. The learner may construct explanatory models, evaluate their consequences qualitatively and quantitatively, and revise the models, getting support from the system. Secondly, the learner may state diagnostic hypotheses and receive feedback about the usefulness of diagnostic investigations.

Concerning the *modelling component*, our near future plans are i) to develop a component for the acquisition of conceptual domain knowledge that can be used for generating explanations for relations between variables, and ii) to be able to acquire distributions from verbal descriptions of relations between variables (phrases like e.g. "sometimes brings about" in Figure 2). We are pursuing a Bayesian and a fuzzy approach in this respect. Concerning the *diagnostic support component*, we plan to implement a dialog that supports the learner's generation and revision of diagnostic hypotheses as well as diagnostic reasoning steps, and that provides feedback and explanations.

Concerning *applications* of MEDICUS, one of our cooperations is aimed at generating realistic models with the help of case data. These models will serve for health consulting as well as for diagnostic training. Some of the other cooperations are aimed at applying MEDICUS to assist in planning, executing, and evaluating environmental exposition analyses ("environmental monitoring"). MEDICUS is seen as a promising contribution to quality ensurance and improvement in this subdomain.

5 References

Anderson, J.R. (1993). *Rules of Mind*. Hillsdale: Erlbaum.

Andreassen, S., Woldbye, M., Falck, B., Andersen, S.K. (1987). *MUNIN - A Causal Probabilistic Network for Interpretation of Electromyographic Findings*. Proceedings 10th IJCAI 87, 366-372.

Barrows, H.S., Tamblyn, R.M. (1980). *Problem-Based Learning: An Approach to Medical Education*. New York: Springer.

Boshuizen, H.P.A., Schmidt, H.G. (1992). On the Role of Biochemical Knowledge in Clinical Reasoning by Experts, Intermediates, and Novices, *Cog. Science*, 16, 153-184.

Cooper, G.F. (1984). *NESTOR: A Computer-Based Medical Diagnostic Aid that Integrates Causal and Probabilistic Knowledge*, PhD Thesis, Medical Computer Science Group, Stanford University, Stanford, CA (Report HPP-84-48).

Elstein, A.S., Bordage, G. (1980). Psychology of Clinical Reasoning., in G.C. Stone, F. Cohen, N.E. Adler, and Associates (eds), *Health Psychology - A Handbook*. San Francisco: Jossey-Bass Publ., 333-367.

Elstein, A.L., Shulman, L.S., Sprafka, S.A. (1978). *Medical Problem Solving - An Analysis of Clinical Reasoning*, Cambridge: Harvard University Press.

Gollwitzer, P.M. (1990). Action Phases and Mind-Sets. In E.T. Higgins, R.M. Sorrentino (eds): *Handbook of Motivation and Cognition: Foundations of Social Behavior*, 2, 53-92.

Heckerman, D.E. (1991). *Probabilistic Similarity Networks*, Cambridge: MIT Press.

Henrion, M. (1987). Uncertainty in Artificial Intelligence: Is Probability Epistemologically and Heuristically Adequate? In J.L. Mumpower, L.D. Philipps, O. Renn, V.R.R. Uppuluri (eds), *Expert Judgements and Expert Systems*. Berlin: Springer (NATO ASI Series F: Computer and Systems Science), 106-129.

Lauritzen, S.L., Spiegelhalter, D.J. (1988). Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems, *Journal of the Royal Statistical Society*, B50(2), 157-224.

Miller, R.A., Pople, H.E., Myers, J.D. (1982). INTERNIST-I, an Experimental Computer-Based Diagnostic Consultant for General Internal Medicine, *The New England Journal of Medicine*, 307(8), 468-476.

Möbus, C. (1995). Towards an Epistemology of Intelligent Problem Solving Environments: The Hypothesis Testing Approach, in J. Greer (ed), *Proc. of the World Conf. on Artificial Intelligence and Education AI-ED 95*, 138-145.

Möbus, C., Pitschke, K., Schröder, O. (1992). Towards the Theory-Guided Design of Help Systems for Programming and Modelling Tasks. In C. Frasson, G. Gauthier, G.I. McCalla (eds): *Intelligent Tutoring Systems, Proc. ITS 92*. Berlin: Springer, LNCS 608, 294-301.

Neapolitan, R.E. (1990). *Probabilistic Reasoning in Expert Systems*, New York: Wiley.

Newell, A. (1990). *Unified theories of cognition*. Cambridge: Harward University Press.

Patel, V.L., Groen, G.J. (1986). Knowledge Based Solution Strategies in Medical Reasoning, *Cognitive Science*, 10, 91-116.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo: Morgan Kaufman (2nd ed.).

Pearl, J. (1993). Belief Networks Revisited. *Artificial Intelligence*, 59, 49-56.

Schurz, G. (1988). *Erklären und Verstehen in der Wissenschaft (Explanation and Understanding in Science)*, München: Oldenbourg.

Shortliffe, E.H. (1976). *Computer-Based Medical Consultations: MYCIN*, New York: North-Holland.

Srinivas, S., Russell, S., Agogino, A. (1990). Automated Construction of Sparse Bayesian Networks from Unstructured Probabilistic Models and Domain Information. In M. Henrion, R.D. Shachter, L.N. Kanal, J.F. Lemmer (eds): *Uncertainty in Artificial Intelligence 5*, Amsterdam: North-Holland, 295-307.

Suppes, P. (1970). *A Probabilistic Theory of Causality*, Amsterdam: North-Holland.

Szolovits, P., Pauker, S.G. (1978). Categorical and Probabilistic Reasoning in Medical Diagnosis, *Artificial Intelligence*, 11, 115-144.

Szolovits, P., Pauker, S.G. (1993). Categorical and Probabilistic Reasoning in Medicine Revisited, *Artificial Intelligence*, 59, 167-180.

Van Lehn, K. (1988). Toward a Theory of Impasse-Driven Learning. In H. Mandl, A. Lesgold (eds): *Learning Issues for Intelligent Tutoring Systems*. Springer, 19-41.

Waldmann, M.R., Holyoak, K.J. (1992). Predictive and Diagnostic Learning Within Causal Models: Asymmetries of Cue Competition, *Journal of Experimental Psychology: General*, 121, 222-236.

Weiss, S.M., Kulikowski, C.A., Amarel, S. (1978). A Model-Based Method for Computer-Aided Medical Decision Making, *Artificial Intelligence*, 11, 145-172.