Knowledge Acquisition for Decision-theoretic Expert Systems*

Peter Lucas
Department of Computer Science, Utrecht University
P.O. Box 80.089, 3508 TB Utrecht
The Netherlands
E-mail: lucas@cs.uu.nl

Abstract

In this paper, the construction of decision-theoretic expert systems in collaboration with domain experts is discussed. In particular, the role of domain models in guiding the knowledge-acquisition process is reviewed, and various techniques that may help in the design of a decision-theoretic expert system are presented. Treatment planning in patients with a congenital heart disease is described as an example domain. The development of a decision-theoretic expert system for this domain is taken as a running example.

1 Introduction

Decision-theoretic networks offer a mathematically sound collection of formalisms for building knowledge-based (expert) systems for domains in which uncertainty is of central concern. In this paper, such systems will be called *decision-theoretic expert systems*. The main applications of the formalisms are in classification, e.g. diagnosis (cf. [4]), and in decision-making under uncertainty, e.g. optimal treatment management of a patient (cf. [1]).

The decision-theoretic network formalisms originate from two different fields: knowledge-based systems [11], and statistical decision theory [16], which is also reflected in the various issues that arise when building decision-theoretic expert systems. As with any expert system, extracting knowledge of a specific domain from various sources, such as experts, literature and databases, is required in the process of building such systems. The knowledge-acquisition process is guided by specific, often extensive, models of the domain. Thus, the process of building decision-theoretic networks fits into the methodologies of knowledge engineering proposed in recent years (cf. [14]). These knowledge-engineering techniques are supplemented with modelling techniques from decision analysis and statistics. Hence, designing such systems requires drawing upon techniques from various fields. It is, however, not apparent how these techniques must be combined; detailed guidelines for building decision-theoretic expert systems are currently lacking.

In this paper, the process of building decision-theoretic networks in collaboration with a domain expert is addressed. To illustrate the approaches discussed, treatment selection for congenital heart disease is adopted as an example domain. The results derive from a project

*Published in: *AISB Quaterly*, **94**, 23–33, 1996.

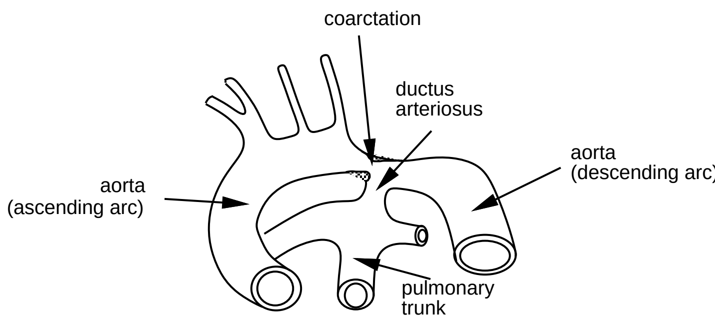

Figure 1: Anatomy of the aorta in aortic coarctation.

in which the suitability of decision-theoretic networks for treatment selection under timing constraints are investigated.

The structure of the paper is as follows. Since some basic medical knowledge is required for understanding the examples in this paper, in the next section, the medical problem of diagnosis and treatment of congenital heart disease is sketched. In Section 3, the theory of decision-theoretic networks is reviewed. Section 4 addresses the issue of creating domain models that can be used as a basis for the construction of decision-theoretic networks. A methodology for the design of decision-theoretic networks is proposed. Next, Section 5 deals with methods for the refinement of decision-theoretic networks. Finally, in Section 6, some limitations and future directions of research are identified.

2 Treatment selection in congenital heart disease

Of every thousand children born, eight have a cardiac anomaly. Of these, *aortic coarctation* is the anomaly that occurs most frequently, often in combination with other anomalies of the cardiovascular system. In Figure 1, the anatomy of aortic coarctation is depicted. In aortic coarctation there is a constriction of the aorta, nearby the ductus arteriosus. The ductus arteriosus is part of the fetal circulation; normally, it closes soon after birth. An open ductus arteriosus is encountered frequently in connection with cardiovascular anomalies. Blood flow and pressure are maintained by the left ventricle, which, upon contraction (systole), ejects blood into the aorta. The aorta transfers oxygen-rich blood to the body. Due to the constriction of the aorta, the left ventricle is overloaded; it is not longer capable of meeting the requirements of the body for blood supply. This condition is known as *heart failure*. Problems due to heart failure arise early after birth in patients with aortic coarctation.

Physical examination of a child with aortic coarctation and associated heart failure reveals the following symptoms and signs: cardiomegaly (enlarged heart), hepatomegaly (enlarged liver), abnormal breath sounds on auscultation (pulmonary crepitations), and various signs of breathing difficulty (dyspnoea): tachypnoea (increased respiratory rate) and tachycardia (increased heart rate). The constriction of the aorta causes the blood pressure in the upper half of the body to increase; it is decreased in the lower half.

Surgical treatment aims at redressing the constriction. The choice is between the following types of surgical treatment:

- surgical correction of the narrowed aorta, and

- balloon angioplasty, i.e. the aorta is dilated by means of an intravascular balloon.

In addition to surgical therapy, there is a place for medical management to suppress the symptoms associated with the disease. Among others, the substance Prostaglandin E is administered. It supresses the closure of the ductus arteriosus, which has an important function in by-passing the constriction. If a patient is too ill to undergo surgical treatment, medical management is the only available therapy.

The choice of the appropriate therapy depends on the patient's condition as well as on expected complications of a treatment. Of special interest in this respect is the appropriate timing of operation [12]. The risks are extremely high when a patient is operated on at an early age. However, waiting too long with operating is dangerous too, because after some time a patient develops persistent hypertension, and late complications which diminish life expectancy.

3 Decision-theoretic networks

We briefly review the theory of decision-theoretic networks so far as is required for the reading of this paper. For recent overviews on the subject, the reader is referred to [2] and [7].

Often, a distinction is made between two types of decision-theoretic network:

- belief networks, also called probabilistic networks, and
- influence diagrams.

A *belief network* is an acyclic directed graph $G = (N, A)$, with a set of nodes N representing random variable $V \in N$, and a set of arcs A representing causal or influential relationships between random variables. The presence of an arc between two variables denotes the existence of a direct influential relationship; absence of an arc means that the variables do not influence each other directly. Associated with a belief network G is a joint probability distribution Pr, defined in terms of conditional probability tables according to the topology of the graph. By means of special algorithms for probabilistic inference (cf. [13, 9]), a once constructed belief network can be employed to process evidence, yielding, for example, a diagnosis.

The belief-network formalism provides only for probabilistic inference. For making decisions, certain extensions to the belief-network formalism are required, as offered, among others, by the influence-diagram formalism. Like a belief network, an *influence diagram* is an acyclic directed graph $G = (N, A)$, except that in addition to probabilistic nodes $\mathcal{O} \subseteq N$, two additional node types are distinguished: decision nodes and a value node. Figure 2 shows a simplified influence diagram for the treatment of aortic coarctation. A probabilistic node is indicated by means of an ellipse. As holds for belief networks, arcs between probabilistic nodes denote direct causal or influential relationships. An example of such an uncertain causal relationship in Figure 2 is the arc between 'heart failure' and 'dyspnoea'. *Decision nodes* $\mathcal{D} \subseteq N$, denote decision variables; a decision node is indicated by means of a box. The decision node with name 'treatment' in Figure 2 stands for the form of surgical or medical treatment to be selected. Incoming arcs to decision nodes indicate information that is required for the purpose of making the decision. The order of decision nodes in an influence diagram expresses the order in which the decisions are made in the domain. In the present diagram, there is only one decision node. In order to select the most optimal sequence of

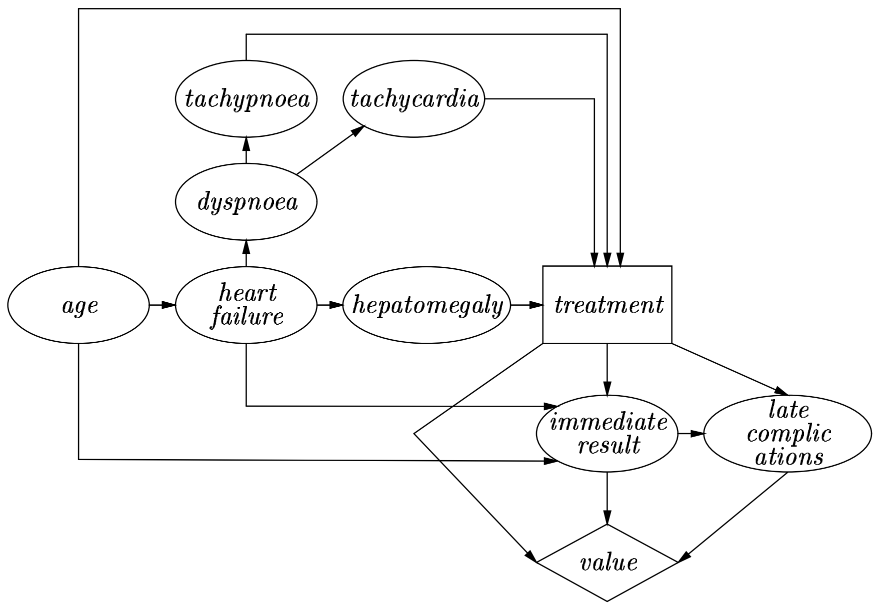

Figure 2: Simplified influence diagram for the treatment of aortic coarctation.

decisions, information of the utility of these decisions with respect to other, partially uncertain, information must be available. This is represented by a singleton set of nodes $\mathcal{V} \subseteq N$, called the *value node*. A value node is indicated by a diamond. Incoming arcs of the value nodes indicate variables on which an assessment must be based. For example, in Figure 2 it is expressed that any decision concerning the best treatment for a patient will be based on the combined assessment of: treatment alternatives, immediate results (death or alive) and late complications of a selected treatment. A value node has no outgoing arcs. It holds that $N = \mathcal{O} \cup \mathcal{D} \cup \mathcal{V}$, where the sets $\mathcal{O}$, $\mathcal{D}$ and $\mathcal{V}$ are mutually disjoint.

The graph representation only denotes the qualitative relationships between variables; it must be supplemented with a quantitative representation to obtain an influence diagram. The quantitative representation is specified in the form of local conditional probability tables for each probabilistic variable $O \in \mathcal{O}$:

$$\Pr(O|\pi(O))$$

where $\pi(O)$ denotes the set of parents of node O. With a decision node D a function

$$d : \pi(D) \rightarrow \{d_1, \ldots, d_n\}$$

is associated, where for each collection of values for variables in $\pi(D)$, the optimal decision d_i after evaluation of the diagram is stored. With a value node $V \in \mathcal{V}$ a function

$$U : \pi(V) \rightarrow \mathbb{R}$$

is associated, i.e. with every instantiation of all variables in $\pi(V)$, a real number is supplied. This function is called a *utility function*. An influence diagram can be evaluated by means of various algorithms, e.g. the algorithm proposed by R. Shachter [15]; the result of evaluation is an optimal sequence of decisions (a decision strategy), obtained by computation of the optimal expected utility.

The reader is referred to [15] or [16] for more detailed treatments of influence diagrams.

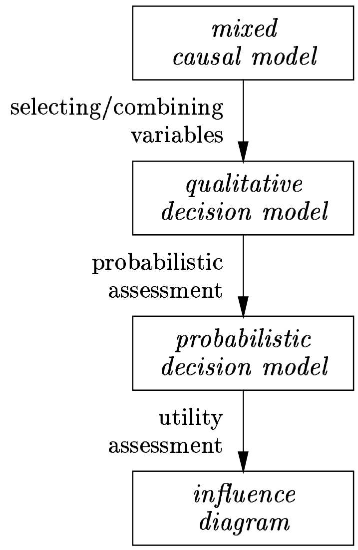

Figure 3: Phases in the modelling of a decision-theoretic expert system.

4 Modelling

Since decision-theoretic networks have been introduced fairly recently in the field of knowledge-based systems, no extensive development methodology is available for building such systems. However, there exists a large body of experience in the related field of decision analysis [16]. Most of that experience concerns the construction and evaluation of decision-tree models. One of the characteristics of the decision-analytic approach is the low level of detail on which the design activity focusses from the start. This contrasts with current knowledge-engineering practice, where the choice of appropriate levels of detail in the design process – initially high and low at the end – is emphasized.

4.1 A methodology for designing decision-theoretic expert systems

The construction of a decision-theoretic expert system can be best viewed as an iterative process, starting with a rough, qualitative model, which, after several refinement steps, yields a definite decision-theoretic network. The following sequence of steps have proven useful in a number of projects:

(1) design of a mixed causal model

(2) development of a qualitative decision model

(3) development of a probabilistic decision model

(4) development of an influence diagram

The last two steps deal with the quantitative representation of an influence diagram. The various stages in the development process are depicted in Figure 3.

Design of a decision-theoretic expert system requires close collaboration with one or more domain experts. The role of a domain expert is particularly important for:

- laying out the topology of the network at the various stages of the development;
- pointing out literature for probability assessment;
- the subjective assessment of probability and utility information;
- the selection of test cases in the process of refining a decision-theoretic expert system.

In the following sections, the various issues that play a role in designing a decision-theoretic expert system in collaboration with a domain expert are discussed. The construction of the example models was carried out in close collaboration with a paediatric cardiologist.

4.1.1 Mixed causal model

The construction of a topology of a decision-theoretic network is arguably the most important step in the development of a decision-theoretic expert system. The topology can be designed from scratch, starting with an inventory of relevant variables with their domains, and relationships between them, supplied by the expert. However, in many domains there are already models available that may guide the design. The expert will be familiar with such models, which eases the design process and promotes the collaboration. A certain amount of work will be required to transform such diagrams to a decision-theoretic network [8]. Medicine is a field where such domain models are available in literature, often in graphical form. In the sequel, we shall focus on medical domain models, but similar principles apply to other fields.

In medicine, the various factors involved in the development of a disorder are often described in terms of normal or abnormal structure and behaviour of biological control systems. This behaviour is described in terms of effects on physical and chemical agents, for example blood pressure or level of glucose in the blood. Such models can be found in standard medical textbooks of physiology (e.g. [3] discusses several of such models in detail). Similar models for pathological behaviour are described in standard textbooks of internal medicine (e.g. [6]), and monographs on specific disorders. These models are frequently described as directed graphs, where arcs denote cause-effect relationships.

In Figure 4, the mixed causal model that resulted for aortic coarctation is shown. The model consists of three parts:

(1) A pathophysiological model of processes that cause the symptoms and signs (e.g. tachypnoea) of aortic coarctation (left part of the graph). Arcs between these nodes denote causal mechanisms underlying the disease.

(2) A decision model, consisting of treatment alternatives (balloon angioplasty, surgery or medical treatment); incoming arcs are physical signs influencing a decision, outgoing arcs point to effects of treatment.

(3) An effects model, representing the various effects of treatment.

Finally, the variables have been grouped as patient, test, therapy and result data, respectively. The paediatric cardiologist felt confident with the diagram, because it was quite similar to diagrams in medical textbooks.

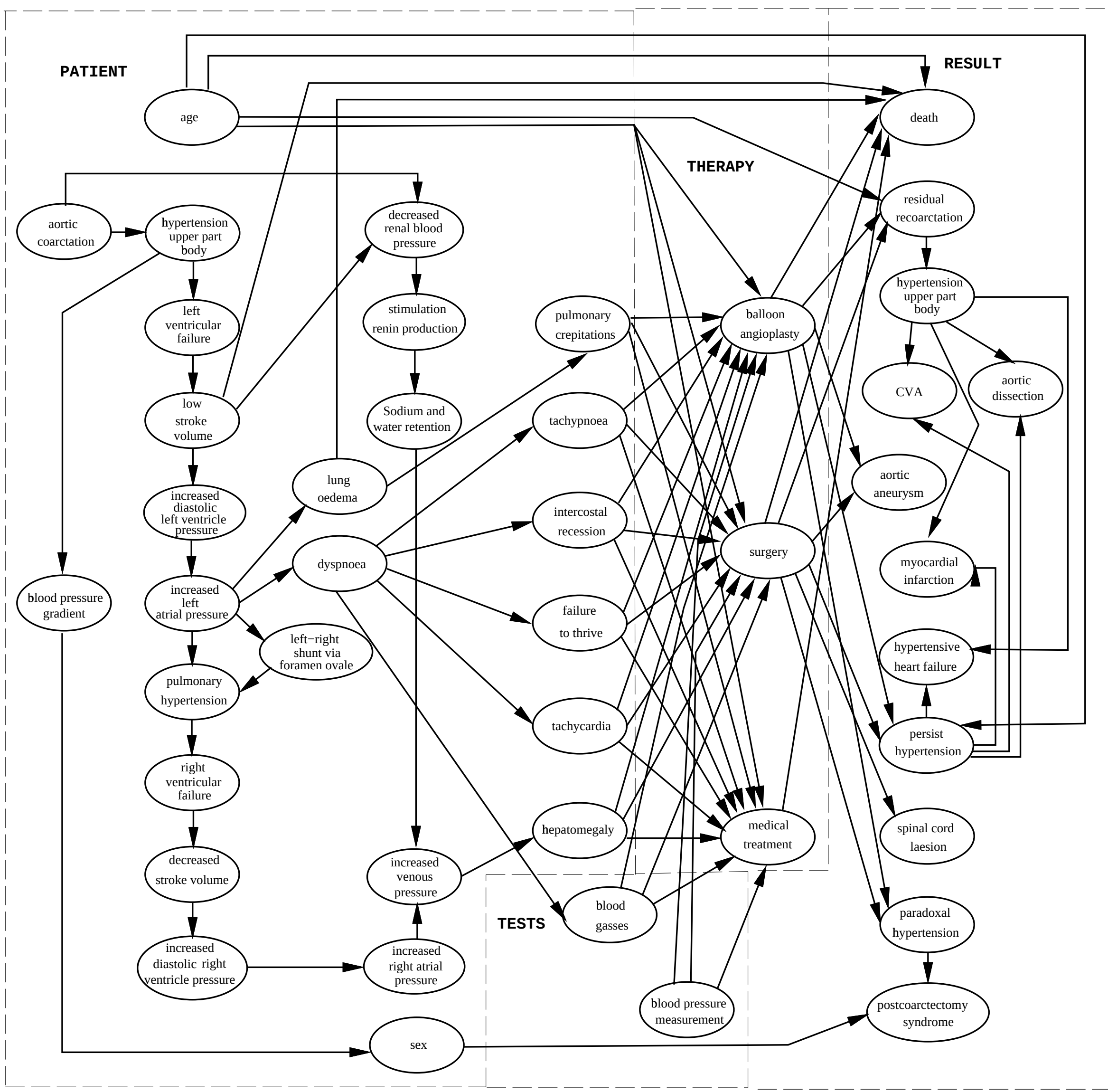

Figure 4: Mixed causal model of aortic coarctation.

4.1.2 Qualitative decision model

A mixed causal model can be transformed to a qualitative decision model by focussing on decision-making: only variables that are relevant in the light of the represented decisions are kept in the model. Furthermore, related binary variables that are mutually exclusive may be combined into an n-valued variable ($n > 2$), yielding a more abstract representation.

For the qualitative decision model of aortic coarctation, this process resulted in a single '*treatment*' variable, that combines the nodes '*balloon angioplasty*', '*surgery*', and '*medical treatment*' from the mixed causal model. To further reduce the complexity of the graph, variables that were strongly related to each other were also combined. Variables of which the role in decision making was not completely clear, were removed.

Timing of treatment was an important aspect of the problem. Since the paediatric cardiologist was not able to view treatment selection and timing as two separate entities, both timing and treatment modalities were combined in the domain of the '*treatment*' variable.

4.1.3 Probabilistic decision model

In the next stage of the development, the qualitative model is supplemented by probabilistic information. The assessment of the utility function is postponed until the following stage. Probabilistic information can be collected from the literature, databases or subjective assessment. Medical literature contains much probabilistic information, but it is usually incomplete, or conditioned on variables different from those required. However, by means of specific computation rules from probability theory, such as Bayes' theorem and marginalisation, it is sometimes possible to transform probabilistic information from the literature to those required.

Although a large database is being compiled by Dutch paediatric cardiologists for epidemiological research, this database contained insufficient data for learning probabilistic information. Hence, many probabilities had to be assessed by the domain expert. Therefore, to make probability assessment feasible, and to profit from probabilistic information available in the literature, the model was simplified. Note that probabilities are assessed as effects conditioned on causes, e.g.

$$\Pr(\mathit{dyspnoea} = \mathit{yes} \mid \mathit{heart\_failure} = \mathit{yes}) = 0.9$$

where '*dyspnoea*' is an effect, and '*heart_failure*' denotes a cause. It is well-known that this eases the subjective assessment of probabilistic information [13]. The subjective assessment of probabilities can be enhanced by:

- ordering the effects from most occurring to least occurring;
- combining subjective estimates with incomplete probabilistic information from the literature.

In transforming the qualitative decision model (which is not shown) to the probabilistic decision model, the number of nodes was reduced from thirty-one to twenty-one.

4.1.4 Influence diagram

Addition of a value node to a probabilistic decision model yields an influence diagram. Utilities are determined by assessing various outcomes from different perspectives, e.g. from objective

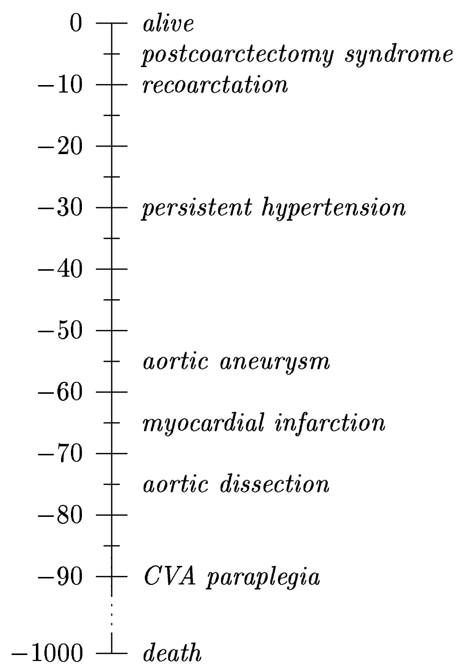

Figure 5: Utility scale for treatment of aortic coarctation.

and subjective (the patient's) perspective. Two methods for the assessment of utilities are distinguished:

- *global methods*, i.e. all outcomes and different perspectives are assessed together as a whole, and not explicitly distinguished;
- *decomposition methods*, i.e. the various outcomes and perspectives are assessed separately, and later combined. For example, life expectancy of a patient may be assessed separately, and later combined with losses due to complications of treatments.

The *standard reference gamble method* can be used to assess utilities in terms of the best and worst outcomes. An easier approach is offered by the *utility scale*, which means that the various possible outcomes are placed by the expert on a numerical scale.

The utility scale for aortic coarctation, designed under the guidance of the expert, is shown in Figure 5. The utility function associated with the value node is a function of instantiated parent variables of the value node. Here, the utility function is assumed to be (almost) additive, i.e. the utility of a combination of instantiated variables is assumed equal to the sum of the utilities associated with the values of individual variables. However, when the value '*death*' was included, the value of the utility was set to -1000. The resulting influence diagram is shown in Figure 6.

The system was implemented using the *Ideal* system, an expert-system shell for the construction and consultation of decision-theoretic expert systems [18].

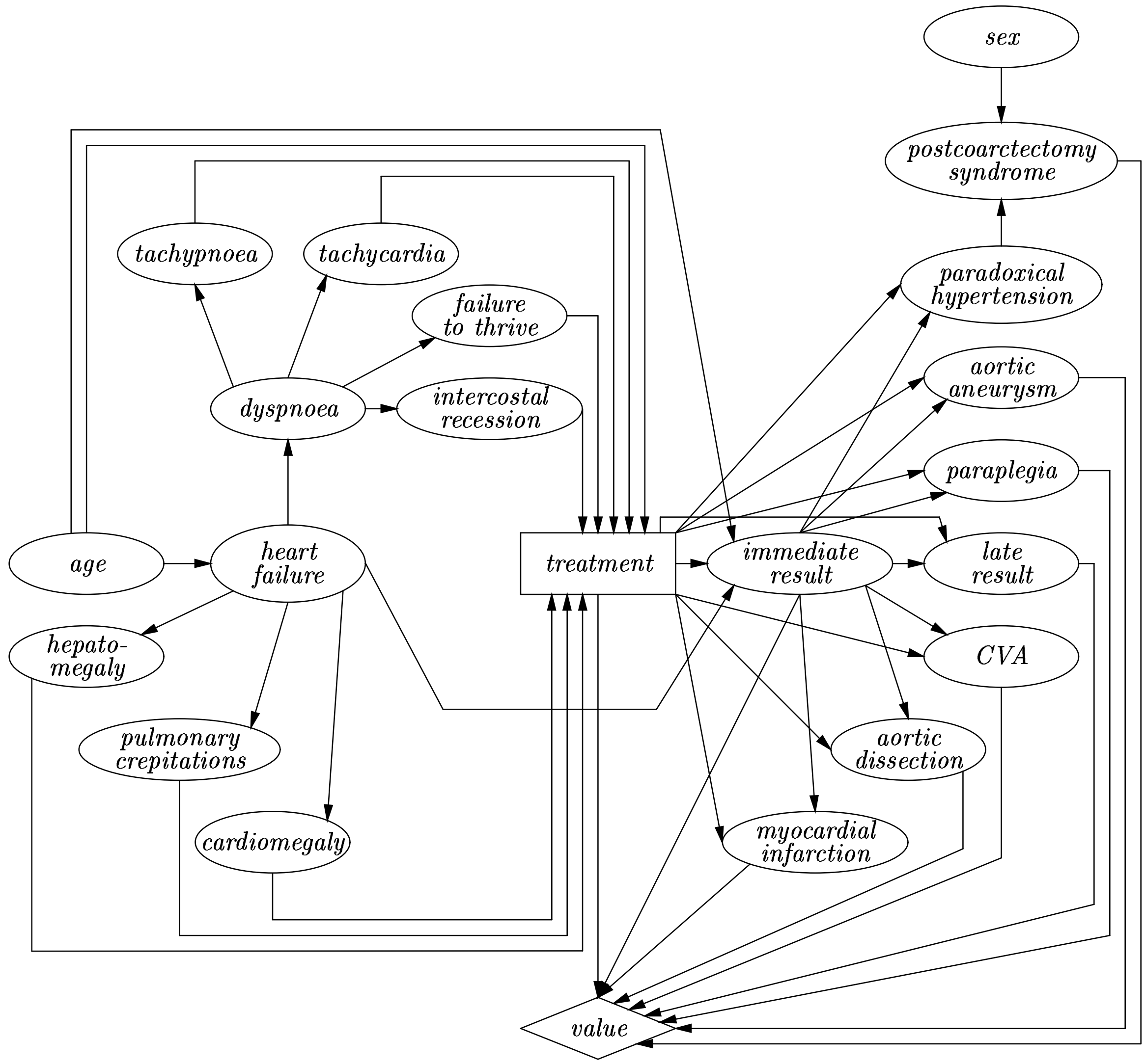

Figure 6: Influence diagram for aortic coarctation.

5 Refinement of a decision-theoretic expert system

After completion of the initial design of a decision-theoretic network, it is necessary to investigate the validity of the assumptions underlying its construction. The results of such investigations will act as a basis for discussion with a domain expert, and may be used for the refinement of the network (cf. [10] for a discussion of related issues for the refinement of rule-based expert systems).

Comparison of an expert system with some other system (computer system or human expert) may yield valuable insight into the capabilities and limitations of the expert system. The system that is used as a reference, is often called the *gold standard*. Sensitivity analysis provides insight into the sensitivity of a model to changes in probability and utility information [16]. A good model should not be very sensitive to small variations in this information. Prediction analysis is especially suited for evaluation of a belief network. Usually, a belief network is applied for diagnosis of a disorder given certain evidence. By using a belief net-

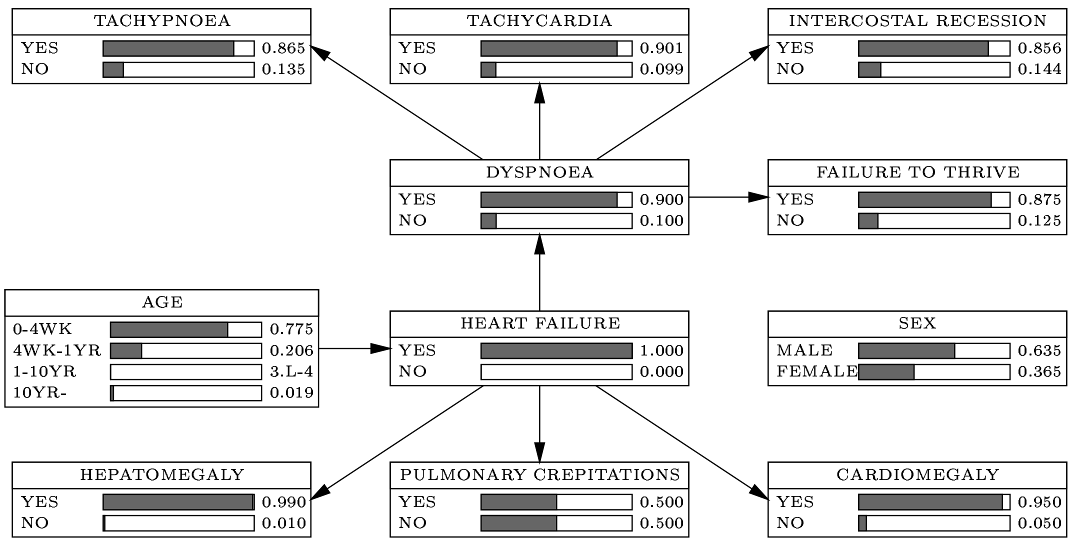

Figure 7: Belief network with evidence for heart failure.

work in the reverse manner, i.e. entering a diagnosis and observing the predicted evidence, valuable information is obtained for a comparison with data available in the literature. Typically, the medical literature contains information about the relative frequency of occurrence of symptoms and signs in connection with the presence of certain disorders.

No similar expert system for aortic coactation was available, so we had to resort to a comparison with the paediatric cardiologists at the hospital. For this purpose, data of seven patients were entered into the system. The results of the system were in correspondence with those of the paediatric cardiologists. The time intervals of the advice (e.g. operating between four weeks and one year) were considered too broad, and were considered to require further refinement.

In order to make sensitivity analysis feasible with the tools we had available, the influence diagram was simplified by summarising the pathophysiological states by the node '*heart failure*'. Similarly, the complications of treatment were summarised by the node '*late complications*'. It appeared that the probabilistic information could be varied to a significant extent, before changes in the advice did occur.

For the purpose of a prediction analysis, the probabilistic part of the influence diagram was selected. The resulting belief network is shown in Figure 7. After entering '*heart failure*' as evidence, the system predicted that in 86% of the patients tachycardia will be observed, and in 88% of the patients patients fail to thrive. These probabilities were higher than those occurring in the medical literature. This indicates that some further refinement of the probabilistic information in the diagram may be required if the diagram is modified for the purpose of more detailed timing advice.

6 Conclusions

In this paper, we have discussed the process of the development of decision-theoretic expert systems with a domain expert. In particular, the application of models of a problem domain in the design of a decision-theoretic expert system has been discussed. We have focussed on using decision-theoretic networks for decision-making. As a typical example, the treatment of patients with aortic coarctation, a congenital heart disease, was discussed. It was argued that the design of a decision-theoretic network requires the integration of techniques from knowledge engineering and decision analysis. In the course of the development of a decision-theoretic network the emphasis shifts from high-level modelling to applying decision-analytic techniques.

Above, we have only dealt with the construction of an influence diagram, not with many other aspects that play a role in developing expert systems. An architecture in which properties of the domain may be defined, such as observable variables, different uses of a network, e.g. for diagnosis, treatment planning and prognostic assessment, seems desirable, but does not yet exist. An initial attempt to define such an architecture is described in [17]. In addition to research on integrating knowledge-engineering methods and decision analysis, those architectural issues need also be addressed in the future.

Acknowledgement. The medical models in this paper were developed in a joint project together with Fergus Macartney, Frea Kruisinga and the author. Linda van der Gaag offered valuable comments on an earlier draft of the paper.

References

[1] Andreassen, S, Benn, J, Hovorka, R, Olesen, KG, Carson, ER, 1994. A probabilistic approach to glucose prediction and insulin dose adjustment: description of metabolic model and pilot evaluation study, *Computer Methods and Programs in Biomedicine* 41, 153–165.

[2] Gaag, LC van der, 1996. Bayesian belief networks: odds and ends, *The Computer Journal* 39, 74–113.

[3] Ganong, WF, 1995. *Review of Medical Physiology*, Prentice-Hall, London.

[4] Heckerman, DE, Horvitz, EJ, Nathwani, BN, 1992. Towards normative expert systems: part I – The Pathfinder project, *Methods of Information in Medicine* 31, 90–105.

[5] Heckerman, DE, Nathwani, BN, 1992. Towards normative expert systems: part II – probability-based representations for efficient knowledge acquisition and inference, *Methods of Information in Medicine*, 31, 106–116.

[6] Isselbacher, KJ, Harrison, TR, eds., 1994. *Harrison's Principles of Internal Medicine*, McGraw-Hill, New York.

[7] Jensen, FV, 1996. Bayesian networks basics, *AISB Quaterly*, this issue.

[8] Korver, M, Lucas, PJF, 1993. Converting a rule-based expert system into a belief network, *Medical Informatics*, 18(3), 219–241.

[9] Lauritzen, SL, Spiegelhalter, DJ, 1987. Local computations with probabilities on graphical structures and their application to expert systems, *Journal of the Royal Statistical Society (Series B)*, 50(2), 157–224.

[10] Lucas, PJF, 1994. Refinement of the HEPAR expert system: tools and techniques, *Artificial Intelligence in Medicine*, 6(2), 175–188.

[11] Lucas, PJF, Gaag, LC van der, 1991. *Principles of Expert Systems*, Addison-Wesley, Wokingham.

[12] Macartney, FJ, Spiegelhalter, DJ, 1987. Medical Management in: Anderson, RH, Macartney FJ, Shinebourne, EA, Tynan, M, eds., *Paediatric Cardiology*, Churchill Livingstone, Edinburgh, 1, 421–442.

[13] Pearl, J, 1988. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, California.

[14] Schreiber, AT, Wielinga, B, Hoog, R de, Akkermans, H, Velde, W van der, 1994. CommonKADS: a comprehensive methodology for KBS development. *IEEE Expert*, 9(6), 28–37.

[15] Shachter, RD, 1986. Evaluating influence diagrams, *Operations Research*, 34(6), 871-882.

[16] Smith, JQ, 1988. *Decision Analysis*, Chapman and Hall, London.

[17] Sonneberg, FA, Hagerty, CG, Kulikowski, CA, 1994. An architecture for knowledge-based construction of decision models, *Medical Decision Making*, 14, 27–39.

[18] Srinivas, S, Breese, J, 1993. *IDEAL: Influence Diagram Evaluation and Analysis in Lisp Documentation and Users Guide*, Rockwell International Science Center, Palo Alto.