Optimal sequential decisions in liver transplantation based on a POMDP model

Guenter Tusch[1]

Abstract. The present investigation aims at the construction of a sequential decision procedure derived from a partially observable Markov decision process (POMDP) model. An optimal clinical management strategy based on a risk assessment of patients is to be found. A decision theoretic cost model different from the general approach has been selected for this clinical management (and classification) task: costs were determined by specifying a minimum acceptable sensitivity and specificity of the overall procedure. The aim is to find the earliest possible decision epoch where a final decision can be made under these quality restrictions. Solution method is non-linear optimisation combined with a robust partial classification method. The probabilities necessary for the model are estimated from data of a clinical study in liver transplantation patients. Decision epochs were at donor organ assessment, immediately before surgery and postoperatively in the intensive care unit. Parameters obtained within decision epochs were combined to scores by artificial neural networks (ANNs). The encouraging results show the applicability of the model in the clinical setting.

1 INTRODUCTION

Our aim is to improve cost-effectiveness in medical supply, especially in diagnostics and therapy. One example is the demand for cutting costs in clinical and laboratory tests, while maintaining the present high diagnostic level. A common problem for patient management in intensive care is the prognosis of the patient. While for some patients it is necessary to follow a substantial part of the clinical course to adequately assess the final prognosis, for others this assessment can easily be made very early because all clinical parameters point to the same direction. This situation is found in liver transplantation patients, a field where the high risk decision can be modelled in a decision-theoretic framework. The aim was to construct a decision procedure to determine if a patient is a high risk patient or not and to estimate the probabilities used in the model from the data. Thus it is shown that the model can be applied in the clinical setting.

The data for the investigated decision problem come from a prospective longitudinal cohort study in liver transplant patients (Donor) covering a five year observation period. 24 clinical tests including 9 subjective assessments were included in the study. A database of 257 organs (cases) was available for analysis. Risk was defined as a transplant survival of less than 30 days. The sequential decision problem was to determine the risk for the organ (patient) as early as possible, (or at the lowest cost), and includes three steps: 1. selection of the donor, 2. after organ harvesting, and 3. three days after the recipient operation. Possible therapeutic consequences are at the first step, rejection of the donor, at the second, rejection of the organ for transplantation and, at the third step, optimisation of intensive care management. All donor organs included in the study were assessed as principally suited for transplantation. General criteria were: donor age under 60 and a stay in the intensive care unit less of than one week [3].

[1] Department of Visceral and Transplantation Surgery, Medical School Hanover, 30627 Hannover, Germany, email: Tusch.Guenter@MH-Hannover.de

2 THE PARTIALLY OBSERVABLE MARKOV DECISION PROCESS

In the following, first the framework of the partially observable Markov decision process (POMDP) will be briefly described. Then the focus will be on the application domain, the management of patients after liver transplantation, and a model will be described that was built to represent the problem. Some issues of problem solving will be discussed after this. Finally, it is shown how the problem-solving procedure for POMDPs could be adapted to deal with additional structure.

A POMDP is a sequential decision model describing a stochastic control process with partially observable states and has the following key ingredients: a finite or infinite set of decision epochs $E := \{1,2,3,...\}$, a finite set of states Θ, a finite set of actions A independent of Θ, a finite set of observations X, a function $T: \Theta \times A \times \Theta \to [0,1]$ representing state and action dependent transition probabilities that describe the dynamic behaviour of the modelled environment, a function $O: \Theta \times A \times X \to [0,1]$ representing observation probabilities that model the relationship among observations, states and actions, and a function $L: \Theta \times A \times \Theta \to \Re$ denotes a loss (cost) model assigning immediate losses to state transitions and models payoffs associated with such transitions. The Markov property simplifies the model in a way, that losses and transition probabilities depend only on the current state and action and not on the past.

In POMDPs process states are hidden to the doctor and decisions can be based only on observations and past actions. This is important when the optimal policy for all possible situations a doctor may encounter should be found. At each decision epoch k, the doctor selects - based on previous observations resulting in a belief state - an applicable action $a \in A$ that again results in an observation $x(a)$ and a new belief state. This defines a *decision rule* δ_k. A policy π provides the doctor with a prescription of choosing an action for every future state and can be considered as a

sequence of decision rules. For a given POMDP the objective is to construct a policy that minimises the expected cumulative risk over some horizon of interest. POMDPs are, in essence, Bayesian decision processes [4].

In a (perfectly observable) Markov decision process (MDP) (E, Θ, A, X, T, L) (see [4]) with a finite number of states that are always known, the optimal policy can be found efficiently using dynamic programming techniques. However, in a POMDP underlying states are not known with certainty and the POMDP must first be transformed into an equivalent belief-state MDP (see [4]). A belief state assigns a probability to every possible state $\vartheta \in \Theta$ and there is an infinite number of belief states one may encounter. The optimisation task predominantly is complicated by the infinite number of possible belief states.

A new belief state b' can be calculated from the current belief state b by Bayes rule when action a is performed and observation x is observed as follows: $b'(\vartheta) = \frac{O(\vartheta, a, x) \sum_{\vartheta' \in \Theta} T(\vartheta', a, \vartheta) \cdot b(\vartheta')}{P(x \mid a, b)}$, where b, b' denote belief states and $b(\vartheta)$, $b'(\vartheta)$ the probability of the state $\vartheta \in \Theta$ in respect to that belief state. The denominator $P(x \mid a, b)$ can be treated as a normalising factor causing the components of b' to sum up to 1.

In the present paper a finite horizon problem is considered. It includes a finite number K of possible decision epochs where a policy is adopted which minimises the expected risk $E\left(\sum_{k=0}^{k=K} \rho_k\right)$, such that ρ_k is the risk obtained at time k. It is also very common to define a reward model where rewards are viewed as negative losses. Then expected rewards are maximised instead of risks being minimised. In the medical domain it seems to be more appropriate to model the problem in terms of risks than in terms of rewards.

3 THE DECISION PROCESS

The sequential decision problem mentioned in the clinical data description is a special formulation of a more general sequential decision problem. It consists of several (K) decision epochs. Additional information is assumed to be supplied by clinical tests or by clinical scores. They will be comprised under 'actions' resulting in observations. Now the decision process runs as follows:

A doctor is supposed to decide a patient's further therapy based on the actual (risk) state being unknown to him/her. The clinical state of the patient is described by the set Θ of discrete, stochastic variables. However, the doctor only can observe the belief state at decision epoch k, and can use this observation to optimise subsequent decision making. (Note, that the observation function O is independent of time.) On each decision epoch (but the last) the doctor has to decide on three possible actions, first if the decision process should continue by an test action a and observation of the results $x(a)$, or second he or she can act as if the patient is to be classified as under risk, or third not under risk. Assume a prior distribution on Θ at the initial decision epoch k = 0, reflecting the doctor's prior beliefs on the clinical state of the patient. Given a sequence of action choices for all decision epochs, and a probability distribution over the set of possible state sequences, the expected loss of the sequence of action choices can be computed. If the minimum expected risk (including expected costs for future observations) is for acting as if the patient is at risk (not at risk), then the appropriate therapy will be maintained over all subsequent decision epochs and no future observations will be considered. If the physician decides on observation, he or she performs an action a_1 (clinical test) and makes an observation $x(a_1)$.In this case he has to decide again for finally selecting the appropriate therapy (and classifying the patient) or performing another action a_2 with observation. At the last decision epoch no observation is possible. A final decision for a therapy based on risk or non-risk is to be made.

The objective is to find an optimal decision-theoretic plan or a policy π, i.e. to select actions during the decision process in a way such that the expected risk is minimised. Prior to the first action choice, a policy π is composed, which prescribes an action choice for each decision epoch k < K, given the history of past actions and observations. At each decision epoch the doctor also incurs a loss; the losses associated with subsequent epochs in a realisation of the decision process are combined by a risk function. When the decision criterion is not the minimum expected risk, but the earliest possible decision on average, this can be achieved by setting all test costs equal to 1. This results in a decision criterion which simply minimises the average number of decision epochs and is the approach chosen for this paper.

In [5] a loss model has been proposed that departs from the general approach in decision theory. Application of decision theory is based on elicitation of losses/utilities or costs attached to the possible outcome by doctors or patients. For ease of calculation in applications of decision theory often the patient's loss and the costs of the tests will be assessed on the same scale, e.g. in cost-benefit-analysis. This may be unacceptable to some patients. To avoid this complication, in [5] and also in this paper a different approach to obtain the loss functions is taken. The losses are determined by a priori fixed sensitivity and specificity of the overall procedure. Instead of minimising the whole risk function, this function is split up into two parts: the loss of the patient and the test costs. Then first the patient's loss is limited by a fixed threshold, and then the costs of the tests are minimised under this restriction.

4 THE DECISION MODEL

It this application area 24 clinical tests (observations in POMDP terms) were available in the clinical data set. However, they cannot assumed to be independent. In the case of dependency the problem of determining an observation function O and the calculation of the new belief state is hardly feasible given the situation of 24 observations and available data of 257 patients. Therefore, it proved to be useful to restrict the complete model and to define a set of decision epochs where a decision was possible, i.e. not for each of the 24 observations (clinical tests) a decision epoch was assumed, but decision was delayed until all donor parameters, all preoperative or all postoperative parameters were measured or assessed.

Peek [2] modelled in a similar context the decision process by a causal probabilistic network. Here a different approach is selected possibly coping both for the dependencies in the data and because of the limited data base for the combinatorial complexity for possible series of actions: all observations done up to a decision point will be combined to scores based on three artificial neural networks (dotted circles in figure 1). Thus the probabilities needed for the POMDP model are implicitly estimated by neural networks.

Artificial neural networks (ANN) can be considered as a wide class of flexible non-linear regression and discriminant models. They consist of an often large number of "neurons", i.e. simple linear or non-linear computing elements, interconnected in often

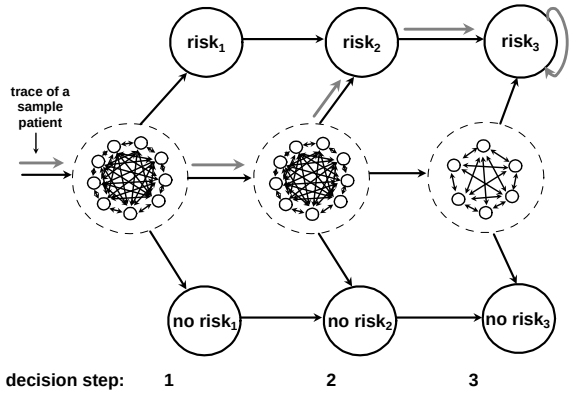

Figure 1. The complete POMDP model (simplified to a finite number of states). For actions within dotted circles all sequences are possible. Unlabeled states correspond to observations (test actions). Within dotted circles all sequences are possible. A trace of a sample patient with final risk decision after step 2 is indicated by grey arrows.

complex ways and organised into layers. (See e.g. Ripley [6]). Here the interest is in classification models (Feed-Forward-Neural-Networks) because the POMDP models a classification rule. The Gold standard is given by 30 days survival of the patient. The ANN is modelled with as many output units as are states in the POMDP. For each state a (ANN-)score will be determined by the output function of the output neuron corresponding to the state. The output neurons are modelled using a logistic activation function.

The underlying Markov process is as follows: Assume the patient is in a non-risk state. If a bad organ is transplanted, he might get into a very bad state. If, however, the doctor decides to discard this organ, the patient would remain in this state and the decision process is finished. If he gets a good organ he also remains in the good state, etc. Unfortunately this process cannot be observed, because the Gold standard (30 days survival) does not cope for short-term dynamics in the process. Thus an approximation has to be made by assuming that the patient does not switch states. Then the underlying process is no longer a real Markov process, but the belief MDP still is a Markov process.

States

The patient can be in two states: risk and non-risk. The organ is assumed not to switch states, i.e. will remain in one state for the whole decision process. This is due to the fact that risk is clinically defined as survival of less than 30 days. This definition does not allow for a determination of transitions. An alternative would be to judge rejected organs as risky, and non-rejected organs as safe and to allow for transitions after transplantation. But this was not agreed to be a Gold standard by the clinicians. Of course, the model itself is flexible enough to cope with transitions. However, here the clinical interest is not to model short term transitions in the intensive care unit, which can be seen as a system partially under control, but to get information on how to act early on long term events that can not be foreseen from clinical practice. Thus, it was not the aim to model short term dynamics but long term risk as an additional information to the clinician for strategic decision making in intensive care. Therefore, the *transition probabilities* are quite simple, because the organ (patient) always remains in the same state: $T(\vartheta, a, \vartheta') = 1$ if $\vartheta = \vartheta'$ and 0 else independent of the action.

Observations

There are a total of 24 possible clinical tests, grouped in three scores, resulting in three observations of the restricted model.

Actions

Two sets of actions are defined:

1. Therapeutic actions as surgery, actions defining intensive care management, etc. There may be several clinical actions necessary at one decision epoch, and therefore they consist of actions necessary for risk patients and those necessary for non-risk patients. The sets may be different after each decision epoch (donor organ, immediately preoperatively and postoperatively). Assuming that at least one observation has to be taken, there are six therapeutic actions forming the set $A_{treat} := \{\mathrm{risk}_1, \mathrm{no\_risk}_1, \mathrm{risk}_2, \mathrm{no\_risk}_2, \mathrm{risk}_3, \mathrm{no\_risk}_3\}$, each element representing a bundle of necessary therapeutic actions for the specific situation. When the physician once selects an action from A_{treat}, the remaining selections are determined and have to be from this set, for subsequent decision epochs and for the same risk, e.g. after selecting $\mathrm{no\_risk}_2$ only $\mathrm{no\_risk}_3$ is admissible.
2. Test actions represent both an action (ordering a test etc.) to observe a clinical parameter and the necessary therapeutic actions to maintain the present clinical status of the patient. A side effect is a delay in the therapeutic (risk) classification. In the liver transplantation domain, performing the same test twice provides no additional information. Therefore, each test is assumed to be applied at most once. Then three test actions are possible, each provided by an ANN score including the former tests.

Losses

Each test score (ANN score) has to be assigned a specific cost (e.g. monetary) which is assumed to be the loss of the corresponding action. Therapeutic actions being inappropriately assigned to the patient result in a loss the patient incurs when incorrectly treated. If the therapy is appropriate for the present state of the patient, the loss is 0. Because of the relation of long-term risk (30 days survival) to the short-term dynamics of the decision process it is reasonable to assume, that for all therapeutic actions for inappropriate treating the patient as risk patient ($A_{treat\_risk} := \{\mathrm{risk}_1, \mathrm{risk}_2, \mathrm{risk}_3\}$) the same losses apply. The same is true for $A_{treat\_no\_risk} := \{\mathrm{no\_risk}_1, \mathrm{no\_risk}_2, \mathrm{no\_risk}_3\}$, however, the losses may have a different value compared to the former. The losses for the therapeutic actions are implicitly determined by the given constraints in the loss model.

The *observation function O* must be estimated from the data (modelled by the ANNs) except for actions from A_{treat} where no observations are taken and the value of the function equals 1.

5 A SOLUTION METHOD FOR THE CONSTRAINED POMDP MODEL

Before the POMDP model can be applied in the clinic two problems have to be solved, the implementation of the proposed here loss model and the estimation of the observation function. The POMDP provides a means to help the physician to act in an optimal way and thus supports the clinical management. At the same

time in the model investigated here an adequate management assumes the classification of the patient– at least at the last decision epoch – into the correct risk class (risk or no risk). This fact can be validated by the Gold standard – 30 days survival. Thus the POMDP can be considered as a classification procedure, and for each state conditional probabilities $P(\neg s \mid s)$ can be determined. $P(\neg s \mid s)$ is the probability that the patient who is in state s according to the Gold standard was classified into a different state by the POMDP procedure. The loss model now can be implemented by constraining the conditional error probabilities, in this specific setting let $P(1 \mid 0, \pi)$ be the probability that a non-risk patient is classified as a risk patient by policy π, and $P(0 \mid 1, \pi)$ the probability that a risk patient is classified as non risk, respectively. Then the loss model leads to a *constrained POMDP* as a POMDP with the additional condition that given conditional error probabilities α and β for all policies π $P(1 \mid 0, \pi) \le \alpha$ and $P(0 \mid 1, \pi) \le \beta$ are true. Admissible policies for the constrained problem are identified by considering only those satisfying both the constraints. Then an optimal policy for the constrained problem can be defined in a similar manner as for the POMDP.

First the local behaviour of the POMDP is considered. For the rest of the paper, at least one observation is assumed to be taken. At epoch *k* the doctor decides for three actions, risk_k, $\mathsf{no\_risk}_k$, or test_k, and divides all possible belief states into three disjoint sets. Let $\delta_k(b)$ denote the action resulting in the application of decision rule δ_k to the actual belief state *b*, $\Omega_0^{(k)} := \{ b \mid \delta_k(\mathrm{b}) = \mathsf{no\_risk}_k \}$ and $\Omega_1^{(k)} := \{ b \mid \delta_k(\mathrm{b}) = \mathsf{risk}_k \}$. Furthermore, the expected risk r_i, $i = 0,1$, attributed to one of the states risk and no risk ($\vartheta_0, \vartheta_1 \in \Theta$) for an admissible policy π can be written as follows:

$$r_i(\pi) = \sum_{k=1}^{K} \int_{\Omega_j^{(k)}} b(\vartheta_i) \cdot L(j \mid i)\, d\mathbf{x}^{(k)} := p_i \cdot P(j \mid i, \pi) \cdot L(j \mid i);$$

$i \neq j; i, j = 0,1$, for the priors $\mathbf{p} = (p_0, p_1)$, where $L(1 \mid 0)$ and $L(0 \mid 1)$ are the losses for the non-risk and risk group. Given the assumptions of the final model $b'(\vartheta)$ can be simplified to $b'(\vartheta) = \dfrac{O(\vartheta, a, x(a)) \cdot b(\vartheta)}{P(x \mid a, b)}$.

The risk function of an admissible policy π is the expected loss

$$R(\mathbf{p}, \pi) = \sum_{k=1}^{K} \sum_{i=1}^{0} r_i(\pi) + \sum_{k=1}^{K} \left(\sum_{j=1}^{k} c_j \right) P(\Omega_i^{(k)}). \tag{1}$$

under the constraints $P(1 \mid 0, \pi) \le \alpha$ and $P(0 \mid 1, \pi) \le \beta$.

It can be shown [5] that an equivalent problem is a non-linear optimisation problem with $2 \cdot K - 1$ dimensions. Let $\mathrm{F}_k := \left\{ b \mid S_0^{(k)} \le \eta_0^{(k)^*} \text{ and } \quad S_1^{(k)} \le \eta_1^{(k)^*} \right\}$. Then the problem can be formulated as a constrained minimisation problem:

$$\left\{ \eta_i^{(k)} \mid i = 0,1; 1 \le k \le K \right\} \min \left\{ c_1 + \sum_{k=2}^{K} c_k \cdot P(\mathrm{F}_{k-1}) \right\} \tag{2}$$

s. t. $\sum \alpha_k = \alpha, \quad \sum \beta_k = \beta, \quad$ and $P(\mathrm{F}_K) = 0,$

where $\alpha_k := P(\Omega_0^{(k)}) \ge 0, \quad \beta_k := P(\Omega_1^{(k)}) \ge 0.$

The cut-off points of the ANN scores are determined by a rank procedure (see Tusch [5]). The optimal α_k^* and β_k^* can be determined by a non-linear optimisation procedure. The search for optimal cut-offs follows $2 \cdot K - 3$ dimensions, because at stage K the cut-offs are equal.

It should be mentioned that a solution for the constrained decision problem can not always be found. It depends highly on the quality (information content) of the tests scores at the different stages, if there is a solution for given α and β.

Finally, some technical aspects of artificial neural networks and linear discriminant analysis to build scores are considered. One important problem in practical applications of neural networks is to avoid overfitting, i.e. to determine the appropriate complexity of the network. Model selection and regularisation are the two main approaches to controlling the complexity of neural networks. For model selection the model with the minimum generalisation error estimate is selected, which best can be done by bootstrapping for non-linear models, but Schwarz's Bayesian information criterion (BIC) [7] is reasonable effective for larger samples and much cheaper than bootstrapping [6 p.61]. One way of regularisation is weight decay which can be viewed as penalised ridging. Regularisation reduces to ridge regression in the case of linear models. For reasons of easier interpretation for doctors [5] and reasons of adequate comparison to LDA only *augmented* multilayer perceptrons (AMLP) and radial basis function models (ARBF) were considered. Linear discriminant analysis (LDA) can be considered as a very simple ANN without hidden units and therefore is suited as a means for comparison. LDA is a well known method [9]. To cope for overfitting, variable selection for the *additional* variables at each step and the ridge method were applied.

6 RESULTS

The data was processed in the same way for the augmented multilayer peceptron (AMLP), the augmented RBF (ARBF), and the linear discriminant (LDA) model except for variable selection applied only to LDA. Discrete variables were dichotomised. As far as possible from a clinical point of view, missing values were eliminated by imputation, i.e. for continuous variables the mean value and for dichotomous ('yes/no') the most frequent value was imputed.

The ANNs (AMLP and ARBF) were modelled using SAS and the TNN3 macros by W.S. Sarle (SAS Institute, Cary NC) with modifications from the author to cope for pruning of the networks. (See [8].) The final networks had 4 input neurons for epoch 1, 4+2 for epoch 2, 4+2+3 for epoch 3. Input neurons of previous epochs were included to cope for dependencies in the clinical tests. Weight decay by MAP Bayesian training with hyperparameters estimated by the program was used (see e.g. Ripley [6]). The importance of connections to the hidden units of the ANNs was investigated by model selection [8]. All models with and without connections were compared according to three measures: 1. BIC, 2. 10-fold cross validation error rate, and 3. randomly selection of a training and a validation set (1:1) and using the validation error rate. The results were very similar. The first method (BIC) resulted in a model without hidden units which is equivalent to the linear discriminant function. When using the methods in the above sequence, an increasing number of hidden units (up to two) was obtained. For figure 2 a number of hidden units corresponding to measure 2 was chosen. Then ANNs were obtained with one hidden layer on average (for the three ANN scores). The results are given in figure 2.

For every epoch of the procedure a linear discriminant function was constructed using stepwise variable selection with the standard

criteria of the SPSS statistics program (SPSS Inc. Chicago, Il.). Nine variables were selected for the final procedure. Discriminant analyses were performed using SAS statistical procedures.

Figure 2 displays the results for the liver transplantation data set in terms of misclassification error and average step count for the entire sequential decision rules based on AMLP, ARBF, and LDA with constraints $\alpha=\beta$ and different values.

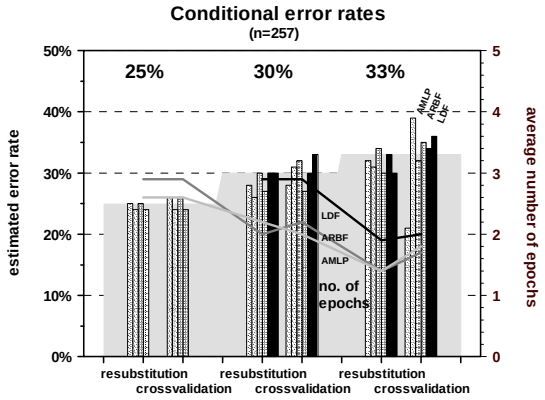

Figure 2. Comparison of the procedures under constraints $\alpha=\beta=25$, 30, or 33%. Probabilities for the POMDP were estimated by
AMLP = augmented multilayer perceptron,
ARBF = augmented RBF network,
LDF = linear discriminant function (not admissible for $\alpha=\beta=25\%$).

7 DISCUSSION

POMDPs provide a powerful modelling framework for decision-theoretic planning, with promising applications to multi-stage clinical decision problems. Recently, several models have been proposed [1,2]. A discussion of these developments is given in [2]. The POMDP model provides a considerable flexibility compared to other approaches. (See e.g. [10].) The restriction imposed by the Markov property, however, might be critical in the investigated application domain. To cope for dependencies in the clinical data input neurons of previous decision epochs were included into the ANNs in a way sacrificing the Markov property. This method is feasible (and necessary) only for the smaller number of decision epochs that was achieved by restricting the original model. The restriction was useful, because the generality of the standard POMDP model limits practical application of the framework due to the computational complexity of associated solution methods. The specialised POMDP form and an algorithm to support a frequently encountered type of clinical management problem assumes several restrictions on the effects of actions on state development, and on the structure of admissible solutions. These restrictions jointly reduce the number of action-sequence classes inducing a different probability distribution on state sequences. The proposed algorithm exploits this property by basing the decisions on clinical scores or test results. In this paper, the problem of constructing a sequential management decision process is embedded into a classification task because therapeutic decisions are based on the underlying classification. The classification into two (risk) groups at an average minimum cost or at the earliest possible time has been investigated by use of a POMDP model and implicitly estimation the probabilities by ANNs. The quality of the procedure was maintained by specified upper bounds for the conditional errors of the entire procedure α and β. Both α and β had to be feasible for the procedure, i.e. their possible range depends on the (statistical) information provided by the given (ANN) test scores. The aim was to construct an optimal (Bayes) policy and to investigate its properties on a clinical data set of liver transplantation patients. However, the (Bayes) policy is optimal when the distributions are completely *known*. This is a general problem, common to most classification procedures. Here a robust data-driven optimisation procedure was selected exhibiting a reasonable good cross validation error. (10-fold cross validation was chosen because it is more effective than a split into test and learning set, especially in relatively small data sets [6]).

In this paper an approach to supervised learning was explored. This led to a problem of determining the gold standard for every decision epoch. Therefore, it was assumed that donor organs (patients) did not switch states. Alternatively, a reinforcement learning approach as already anticipated by using ARBF networks could circumvent this problem by considering the decision problem as an optimal stopping problem. Approximation methods as Q-learning or TD(λ)-learning might then be feasible [11].

The model, as it has been developed in the paper, has been restricted to a very common clinical situation. However, extensions are possible. Two possible extensions will be discussed. Firstly, the model was restricted to two states. The methodology can easily be extended to more states [5]. Secondly, sensitivity and specificity were used as constraints. The positive and negative predicted value, i.e. the posterior probabilities $b(\vartheta)$, may also be used, instead of the conditional probabilities (see Tusch [5]).

ACKNOWLEDGEMENTS

I thank Prof. B. Schneider for his helpful advice, and two anonymous reviewers for their comments on the manuscript. I also thank Prof. J. Klempnauer and Prof. G. Gubernatis for making the clinical data set available to me.

REFERENCES

[1] M. Hauskrecht and H. Fraser, 'Planning medical therapy using partially observable Markov decision processes', in: *Proc. 9th Int. WS Principles of Diagnosis (DX-98)*, Cape Cod, MA, 1998, 182-189.

[2] N.B. Peek, 'Explicit temporal models for decision-theoretic planning of clinical management', *Artif. Intell. Med.* **15(2),** 135-154, (1999).

[3] G.Gubernatis, G.Tusch, K.Oldhafer and R.Pichlmayr, 'Subjective assessment of donor liver quality', *Transplant. Proc.*, **27(3),** 2191-4, (1995).

[4] D.J. White, *Markov Decision Processes.* Wiley, Chichester, 1993.

[5] G. Tusch, 'An optimization model for sequential decision-making - applied to risk prediction after liver resection and transplantation', *Proc. AMIA Symposium* (1999) 425-9.

[6] B.D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 1996.

[7] G. Schwarz, 'Estimating the Dimension of a Model', *The Annals of Statistics* **6,** 461-464, (1978).

[8] U. Anders and O. Korn, 'Model Selection in Neural Networks', *Neural Networks* **12,** 309-323, (1998).

[9] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition* (John Wiley & Sons, New York, 1992).

[10] N.B. Peek and P.J.F. Luckas, 'Trade-offs in decision-theoretic planning', in: J. Hunt and R. Miles, eds., *Research and Development in Expert Systems XIV*, SGES Publications, 1997, 147-157.

[11] D.P. Bertsekas and J.N. Tsitsiklis, *Neuro-dynamic programming*, Athena Scientific, Belmont, Mass., 1996.