ELSEVIER

Artificial Intelligence in Medicine 8 (1996) 147–166

Artificial Intelligence in Medicine

Synthesis of an anaesthetic agent administration system using fuzzy inductive reasoning

Angela Nebot [a,1], François E. Cellier [b,*], Derek A. Linkens [c,2]

[a] *Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Diagonal 647, 8na. planta, Barcelona 08028, Spain*

[b] *Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ 85721, USA*

[c] *Department of Automatic Control and Systems Engineering, University of Sheffield, P.O. Box 600, Sheffield S1 4DU, UK*

Received 1 March 1995; revised 1 July 1995; accepted 16 July 1995

Abstract

Control of the depth of anaesthesia is a difficult undertaking. Progress has been made during recent years by use of different methodologies and monitoring systems that suggest the safe amount of an anaesthetic drug, considering the condition of an individual patient. Despite these improvements, anaesthetists still rely heavily on personal experience when suggesting the anaesthetic dosage during surgical operations. The purposes of this paper are twofold. One is a description of the design of an anaesthetic agent control system using a qualitative modelling and simulation methodology called *Fuzzy Inductive Reasoning* (*FIR*). A comparison with a system developed for the same application using a neural network approach is also presented. The second purpose is a discussion of the problem of separating system-generic from patient-specific behaviour in the context of inductive modelling using the FIR methodology. In order to be useful, the model generated by FIR should reflect upon system-generic behavioural characteristics exclusively, while suppressing patient-specific behavioural patterns. A technique based on combining knowledge obtained from different patients is designed that makes it possible to derive a single model characterizing a specific class of similar patients undergoing similar operations, preserving the common characteristics of all these patients while filtering out the specific behavioural patterns of any one of the individual patients from whom the data were obtained.

Keywords: Model-based reasoning; Inductive reasoning; Fuzzy systems; Anaesthetic depth; Anaesthetic control

* Corresponding author. Tel.: (1-520) 621-6192; Fax: (1-520) 621-8076; e-mail: cellier@ece.arizona.edu

[1] Tel.: (34-3) 401-6076; Fax: (34-3) 401-6000; e-mail: angela@lsi.upc.es

[2] Tel.: (44-1742) 825-133; Fax: (44-1742) 731-729; e-mail: d.linkens@sheffield.ac.uk

SSDI 0933-3657(95)00030-5

1. Introduction

Both sleep and general anaesthesia are states of unresponsiveness which vary in depth. While sleep is healthy, natural and repeats itself rhythmically once every 24 h, anaesthesia is an artificial state maintained by the continuing presence of chemical agents in the brain.

Anaesthetic agents affect the respiratory system, the cardiovascular system, the central nervous system, and the muscles. The use of anaesthetic agents can produce severe complications and side effects, which, under extreme conditions, may even cause the death of the patient. It is therefore essential that the dose of anaesthetic agents is limited to the minimum amount necessary for proper anaesthesia thereby reducing undesired side effects and minimizing the risk to the patient.

Monitoring devices can be used to record the values of indicator variables, to reason about the consistency of these values, and suggest to the anaesthetist an appropriate dose of anaesthetic agent. Research results have recently been reported in the area of monitor integration that enhance the clinical robustness of such monitoring devices by improving their reasoning capabilities through the detection of critical events and by means of enhancing their alarm accuracy [11].

Several new results have been reported in the past few years relating to the control of the depth of anaesthesia. Both open-loop and closed-loop techniques have been explored [5,8,12,16].

One of these studies resulted in the development of a computer-based on-line expert system called RESAC (Real-time Expert System for Advice and Control) [9]. RESAC comprises a rule-based backward chaining inference engine with about 400 rules and makes use of fuzzy logic and Bayesian reasoning. The rule-base was obtained through knowledge acquisition in consultation with expert anaesthetists [5]. The major problem of this approach is the formidable size of the resulting rule-base. Obviously, this has to be a real-time expert system in order to be of any practical use.

Triggered by the aforementioned difficulties, another study was carried out by the same group that promised to enhance the run-time efficiency of the monitoring system. The new system, ANNAD, involves an Artificial Neural Network for Anaesthetic Dose determination [10]. ANNAD is a feedforward neural network trained through back-propagation. This work is reviewed in the next section.

As an alternative, a Fuzzy Inductive Reasoning model for Anaesthetic Dose, FIRAD, has been developed at the Technical University of Catalonia [13], which is discussed in this paper. The same data that were used to drive ANNAD have also been used with FIRAD in order to be able to compare the results from the two approaches.

2. Background: ANNAD

The artificial neural network approach was chosen due to its inherent ability to learn the input/output behaviour of a system in situations where it is possible to specify the inputs and outputs, but where it is difficult to define analytically a relationship between them. This is precisely the situation in biomedical applications, such as anaesthesia, since clinical signals are readily available through measurements, but no precise

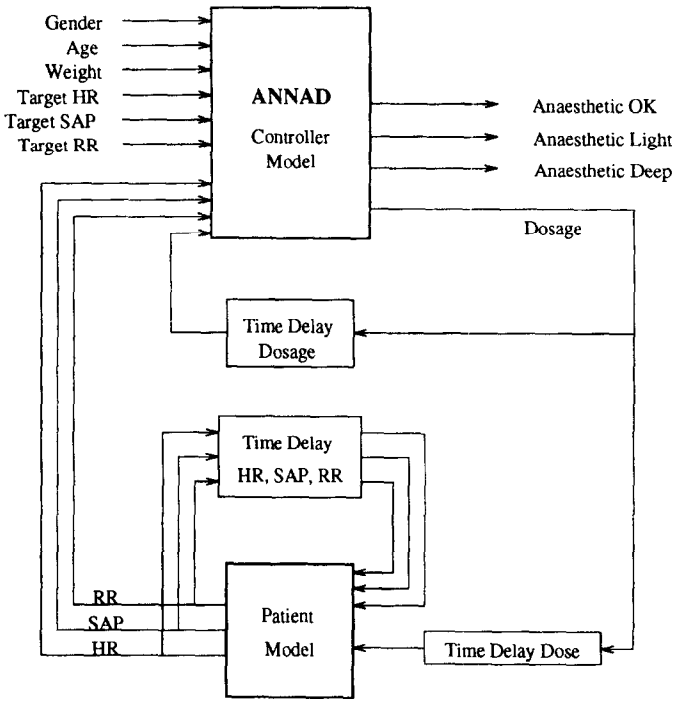


Fig. 1. Feedback loop involving patient simulator and drug controller.

analytical relationships are known between them, and variations between patients are large. Also, neural networks are inherently parallel in nature, and are therefore well suited for real-time environments.

The clinical variables comprising heart rate (HR), respiration rate (RR), systolic arterial pressure (SAP), gender, age and weight of the patient were selected as the key clinical indicator signals to be used for suggesting an anaesthetic dose (control signal).

A patient model and a controller model were independently synthesised by means of the neural network methodology. The control loop was then closed as shown in Fig. 1.

2.1. Artificial neural network patient model

An Artificial Neural Network (ANN) patient model was obtained using a back-propagation algorithm applied to a set of data measured on a patient during a surgical operation.

Three separate neural networks were trained, one for each output: HR, SAP, and RR. The inputs for the training networks were the Dose, older (delayed) values of the Dose, as well as delayed values of HR, SAP, and RR. Each neural network employed two hidden layers.

2.2. Artificial neural network controller

An ANN controller model (ANNAD) was obtained using a back-propagation algorithm applied to another set of data collected from a second patient during a similar

surgical operation as for the patient model. In this case, a neural network with three hidden layers was found to be optimal (showing the smallest deviation from the measured data after training). The inputs for this neural network were Gender, Age, Weight, RR, SAP, HR and the desired values of the latter three variables, while the output was the anaesthetic agent, Dose.

2.3. Closed-loop control

As shown in Fig. 1, the control loop was then closed by connecting ANNAD with the ANN patient model. The results of this experiment demonstrate the stability of the control loop. ANNAD was able to replicate satisfactorily the advice that was obtainable from RESAC. ANNAD also produced good control performance when coupled to a patient simulator. Contrary to RESAC, which was actually used during surgical operations, ANNAD has not yet undergone real-life testing.

For a deeper insight into this work, the reader is referred to [5,16].

3. FIRAD

The motivation for the research described in this paper was to investigate how the *fuzzy inductive reasoning* (*FIR*) methodology performed in comparison with the neural network approach when applied to the identification of dynamic processes from the soft sciences. To this end, we first tried to develop a FIR model for the patient, and then to find a FIR model for the controller. The controller model is called FIRAD (Fuzzy Inductive Reasoning for Anaesthetic Dose). The insights gained during this research effort are detailed in the following subsections.

3.1. The methodology

As is the case of neural networks, the inductive reasoning methodology has the ability to describe (model) systems that are not well understood, that is, systems for which physical laws are only partially or not known. Contrary to the neural network approach, the inductive reasoners contain information about the likelihood of any particular state transition. This is important for model validation purposes. If the accumulated likelihood of a particular state drops below a level that can be specified by the user, forecasting will come to a halt. In this manner, the technique guarantees that the model will not forecast behaviour beyond a time for which the available data are insufficient to substantiate the prediction. Also, this technique is able to enumerate all possible system behaviours that are consistent with the available knowledge (data), and can assign a measure of likelihood of occurrence to each of them.

In the FIR approach [6], the qualitative systems are represented (modeled) by a special class of finite state machines called *optimal masks*, and their episodical behaviour is inferred (simulated) by a technique called *fuzzy forecasting*.

Since FIR, just like all other qualitative reasoning approaches, bases its decisions on discrete (qualitative) variables, it is necessary to discretise continuous variables by means of a technique called *fuzzy recoding* (a fuzzification technique), before the identification of a qualitative model can be attempted.

The fuzzy recoding technique converts quantitative values into qualitative triples. The

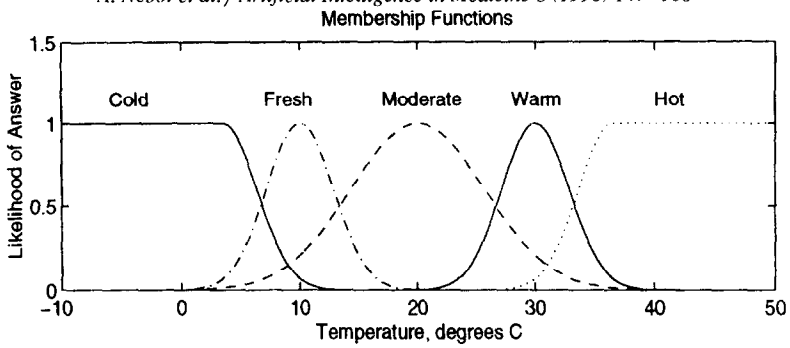


Fig. 2. Membership functions of the temperature.

first element of the triple is the class value, the second element is the fuzzy membership value, and the third element is the side value. The class value represents a coarse discretisation of the original real-valued variable. The fuzzy membership value denotes the level of confidence expressed in the class value chosen to represent a particular quantitative value. Finally, the side value tells us whether the quantitative value is to the left or to the right of the peak value of the associated membership function. The side value, which is a specialty of our methodology since it is not commonly introduced in fuzzy logic, is responsible for preserving the complete knowledge in the qualitative triple that had been contained in the original quantitative value. Fig. 2 shows the fuzzy recoding of a quantitative variable (the temperature) into the five classes 'Cold', 'Fresh', 'Moderate', 'Warm', and 'Hot', using, in the shown example, popular knowledge to determine the so-called *landmarks*, i.e., the borders between neighbouring classes. A quantitative value of *temperature* = 18.0 would in this case be recoded into a class value of 'Moderate', a membership value of 0.938, and a side value of 'left'. Evidently, the qualitative triple contains exactly the same information as the original quantitative value.

After the fuzzification process has been completed, the quantitative trajectory behaviour has been converted into a qualitative episodical behaviour. The episodical behaviour is stored in a *qualitative data model*. It consists of three matrices of identical size, one containing the class values, the second storing the membership information, and the third recording the side values. Each column represents one of the observed variables, and each row denotes one time point, i.e., one recording of all variables, or one recorded state.

In the process of modelling, it is desired to discover finite automata relations among the fuzzified variables that make the resulting state transition matrices as deterministic as possible. If such a relationship is found for every output variable, the behaviour of the system can be forecast by iterating through the state transition matrices. The more deterministic the state transition matrices are, the higher is the likelihood that the future system behaviour will be predicted correctly.

A possible relation among the qualitative variables for a five-variables system example could be of the form:

$$y_1(t) = \tilde{f}(y_3(t-2\delta t), u_2(t-\delta t), y_1(t-\delta t), u_1(t)) \quad (1)$$

where $\tilde{f}$ denotes a qualitative relationship. Notice that $\tilde{f}$ does not stand for any (known or unknown) explicit formula relating the input arguments to the output argument, but

only represents a generic causality relationship that, in the case of the FIR methodology, will be encoded in the form of a tabulation of likely input/output patterns, i.e., a state transition table. In SAPS-II (our implementation of the FIR methodology), Eq. (1) is represented by the following matrix:

$$
\begin{array}{c|ccccc}
t\backslash x & u_1 & u_2 & y_1 & y_2 & y_3 \\
t-2\delta t & 0 & 0 & 0 & 0 & -1 \\
t-\delta t & 0 & -2 & -3 & 0 & 0 \\
t & -4 & 0 & +1 & 0 & 0
\end{array}
\tag{2}
$$

The negative elements in this matrix are referred to as *m*-inputs. *m*-inputs denote input arguments of the qualitative functional relationship. They can be either inputs or outputs of the subsystem to be modeled, and they can have different time stamps. The above example contains four *m*-inputs. The sequence in which they are enumerated is immaterial. They are usually enumerated from left to right and top to bottom. The single positive value denotes the *m*-output. The terms *m*-input and *m*-output are used in order to avoid a potential confusion with the inputs and outputs of the plant. In the above example, the first *m*-input corresponds to the output variable y_3 two sampling intervals back, $y_3(t-2\delta t)$, whereas the second *m*-input refers to the input variable u_2 one sampling interval into the past, $u_2(t-\delta t)$, etc.

In the FIR methodology such a representation is called a *mask*. A mask denotes a dynamic relationship among qualitative variables. A mask has the same number of columns as the episodical behaviour to which it should be applied, and it has a certain number of rows, the *depth* of the mask.

How is a mask found that, within the framework of all allowable masks, represents the most deterministic state transition matrix? This mask will optimise the predictiveness of the model. In SAPS-II, the concept of a *mask candidate matrix* has been introduced. A mask candidate matrix is an ensemble of all possible masks from which the best is chosen by a mechanism of exhaustive search. The mask candidate matrix contains '−1' elements where the mask has a potential *m*-input, a '+1' element where the mask has its *m*-output, and '0' elements to denote forbidden connections. Thus, a good mask candidate matrix to determine a predictive model for variable y_1 in a five-variable system example might be:

$$
\begin{array}{c|ccccc}
t\backslash x & u_1 & u_2 & y_1 & y_2 & y_3 \\
t-2\delta t & -1 & -1 & -1 & -1 & -1 \\
t-\delta t & -1 & -1 & -1 & -1 & -1 \\
t & -1 & -1 & +1 & 0 & 0
\end{array}
\tag{3}
$$

Corresponding mask candidate matrices are used to find predictive models for y_2 and y_3.

Each of the possible masks is compared to the others with respect to its potential merit. The optimality of the mask is evaluated with respect to the maximisation of its forecasting power using the Shannon entropy measure.

Once the optimal mask is found, it is possible to derive a state transition matrix from the optimal mask and the available data. The state transition matrix is a finite-state machine that lists, for each input state (i.e., each combination of input values), all possible output states together with an assessment of the likelihood of their occurrences.

Once the state-transition matrix has been found, a qualitative simulation can be performed by applying the *forecasting* function of the inductive reasoning methodology.

In the process of forecasting the mask is simply shifted further down beyond the end of the raw data matrix (that is the matrix that contains the original continuous data), the values of the *m*-inputs are read out from the mask, and the behaviour matrix is used to determine the future value of the *m*-output, which can then be copied back into the raw data matrix. The membership and side functions of the new input state are compared with those of all previous recordings of the same input state contained in the behaviour matrix. The one input state with the most similar membership and side functions is identified.

The *regenerate* function of SAPS-II is responsible for converting the qualitative predicted episode back into a quantitative output trajectory. It is the inverse operation of the previously described *recode* function. It performs the *defuzzification* operation.

For a deeper and more detailed insight into this methodology, the reader is referred to Refs. [1–4,7,17].

3.2. SAPS patient model

The patient model should be determined by the qualitative relationship between its input variable, the administered Dose and its output variables, the clinical signals of the patient that reflect his or her body reaction to the amount of agent applied (SAP, HR and RR).

In order to determine the patient model, we worked with the data of two different patients. The available measurement data are plotted in Fig. 3.

These plots reveal that the input variable, Dose, varies very little. It is 'high' in the beginning of the experiment, 'medium' for most part of the experiment, and goes 'low' only at the very end of the experiment. It is quite clear that, in the meantime, the output variables react in various ways that are obviously not driven by the input directly, since the input does not change at all. The changes in the output variables were caused by other extraneous factors that were not recorded, and therefore, the variations in the output variables look like *noise* to the inductive reasoner. In fact, the recorded data do indeed contain considerable *digitisation noise*, since all variables were recorded as integers only.

For this reason, the FIR methodology could not find a good mask that models the patient system. The best mask synthesised did not forecast correctly. This was not the case when using the neural network methodology. It turns out that, at least for one of the data files, the neural network gave reasonable responses for the patient model.

As for all inductive techniques, inductive reasoners need a lot of data to work with. It is not possible to generate meaningful and reliable inductive models without ample and

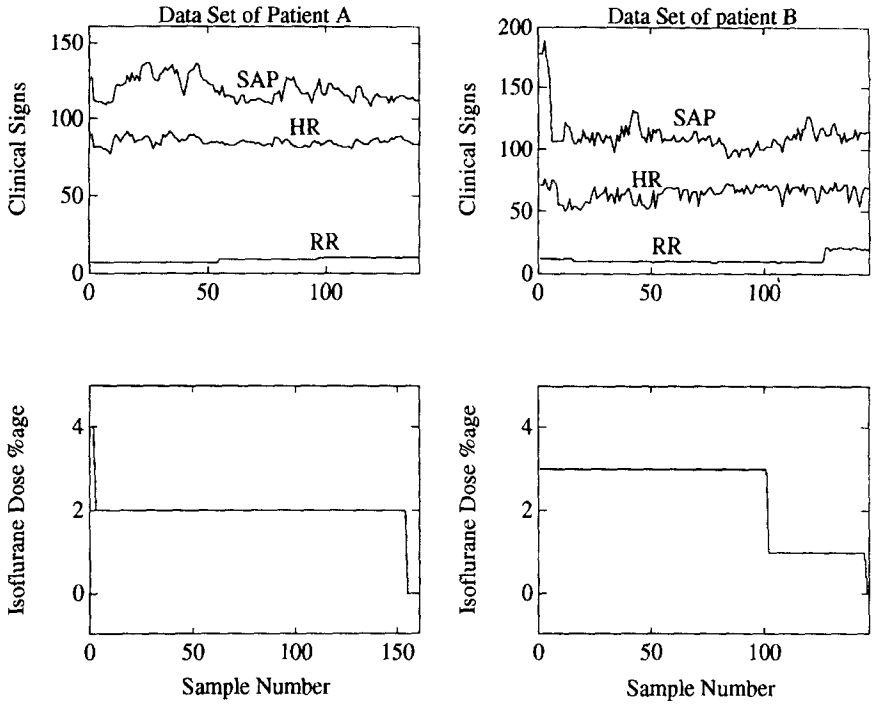


Fig. 3. Patient model measurement data.

rich data. This is equally true for the neural network approach (another inductive modelling technique). However, while the neural network will always predict something, the inductive reasoner will not predict anything that cannot be validated on the basis of the available data. SAPS, our inductive reasoner, simply declines to predict anything when confronted with the patient model data, since no prediction can truly be justified given the available facts.

Here, we observe one of the strengths of the FIR methodology. It will not generate models that are not justifiable from the given data. The neural network methodology generates models for any data, irrespective of whether they are justifiable or not. While SAPS contains an inherent model validation mechanism inside the modelling methodology, the neural network approach does not. The fact that the neural network was able to produce a reasonable response for one of the data sets does not mean that the model is validated. The fact that it was unable to produce a reasonable response for the other data set proves just the opposite. Since inductive models necessarily lack physical insight, we believe it to be absolutely essential for any inductive modeller to contain an intrinsic model validation mechanism that is inseparable from the modelling tool itself. Our fuzzy inductive reasoner, SAPS, does precisely that.

3.3. SAPS controller model

The controller model is determined by the qualitative relationship between its three input variables, SAP, HR and RR, and its single output variable, the administered Dose.

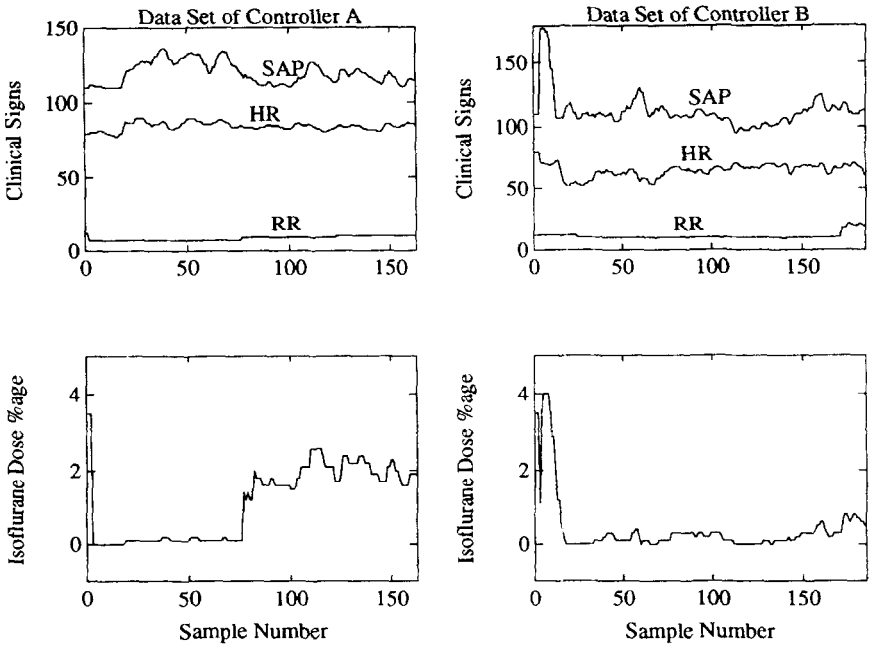


Fig. 4. Controller model measurement data.

For the controller model, we were able also to obtain two data sets from two different patients. These data are plotted in Fig. 4.

Looking at the plots, we can see that the output variable, Dose, varies here considerably more than in the data sets for the patient model. The patient model data are purely clinical data, i.e., data measured in the operating theater. The human anaesthetist didn't find the variations in the biological variables (HR, RR and SAP) alarming, and therefore, reacted very little during the entire operation. In contrast, the controller model data were obtained from RESAC, i.e., the true biological variables that had been observed during surgery were fed into RESAC, which, in turn, proposed a value for Dose. RESAC had been validated by showing the proposed Dose to several anaesthetists, who concluded that RESAC's recommendations were clinically meaningful.

Fortunately for us, RESAC was more 'industrious' than a human anaesthetist would ever be, and reacted to small variations in the biological variables by recommending a slightly modified Dose of the anaesthetic drug. Therefore, there now exists a direct causal relationship between the observed biological data and the recommended Dose, and it should, therefore, be possible to correlate the administered Dose with the biological variables, and come up with a *causal inductive model* that can be used to replace the anaesthetist (or RESAC) in his or her (its) decision making process.

The two data sets contain 163 and 185 records, respectively. They were sampled once per minute. According to information obtained from two anaesthetists whom we

consulted, the slowest time constant of interest in our system is in the order of 10 minutes, and the fastest time constant of importance is in the order of 1 minute.

In accordance with Shannon's sampling theorem, the data should thus have been sampled at least twice a minute. However, by the time, the data were received by us, it was too late to do anything about the problem. Before starting to identify an optimal model, we decided to use a Spline interpolation to find one new data record per interval, located exactly in the middle between the two neighbouring measurement data records. Thereby, the length of the data records was enhanced to 325 and 369 records, respectively. This helped satisfy the data hungriness of the method, but of course, could not recover the information that had already been lost through sampling with too slow a sampling rate.

It was then decided to recode (fuzzify) the variables SAP, HR, and Dose into three qualitative levels (classes), whereas RR was recoded into two qualitative levels only. Here, the landmarks (borders between neighbouring classes) were determined such that each class should contain the same number of recorded data points, rather than relying on expert knowledge.

Due to the difference between the slowest and the largest time constants of importance, we decided to use a mask candidate matrix of depth 21 with nine zero rows in between rows that contain potential inputs.

$$
\begin{array}{l|cccc}
{}_t\backslash^x & SAP & HR & RR & DOSE \\
t-20\delta t & -1 & -1 & -1 & -1 \\
t-19\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-11\delta t & 0 & 0 & 0 & 0 \\
t-10\delta t & -1 & -1 & -1 & -1 \\
t-9\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-\delta t & 0 & 0 & 0 & 0 \\
t & -1 & -1 & -1 & +1
\end{array}
\qquad (4)
$$

In this way, one new forecast is produced every 0.5 min to satisfy Shannon's sampling theorem, and yet, the inductive reasoner looks at input values 5 min and 10 min back to capture the slowest time constant. This technique has proven successful in the past [4].

The first 270 (320) rows of the data matrix were used as past history data to compute the optimal mask. Fuzzy forecasting is used to predict new qualitative class and fuzzy membership values for Dose for the last 55 (49) rows of the raw data matrix, respectively.

For the first data set, the optimal mask obtained was the following:

$$
\begin{array}{l\backslash x} & SAP & HR & RR & DOSE \\
t-20\delta t & 0 & 0 & 0 & 0 \\
t-19\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-11\delta t & 0 & 0 & 0 & 0 \\
t-10\delta t & 0 & 0 & 0 & -1 \\
t-9\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-\delta t & 0 & 0 & 0 & 0 \\
t & -2 & 0 & 0 & +1
\end{array}
\tag{5}
$$

This mask denotes the relationship:

$$\text{Dose}(t) = \tilde{f}((t-10\delta t), \text{SAP}(t)) \tag{6}$$

For the second data set, the optimal mask obtained was:

$$
\begin{array}{l\backslash x} & SAP & HR & RR & DOSE \\
t-20\delta t & 0 & 0 & 0 & 0 \\
t-19\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-11\delta t & 0 & 0 & 0 & 0 \\
t-10\delta t & -1 & 0 & 0 & -2 \\
t-9\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-\delta t & 0 & 0 & 0 & 0 \\
t & 0 & -3 & 0 & +1
\end{array}
\tag{7}
$$

This masks denotes the relationship:

$$\text{Dose}(t) = \tilde{f}(\text{SAP}(t-10\delta t), \text{Dose}(t-10\delta t), \text{HR}(t)) \tag{8}$$

It turns out that the two masks obtained are different. Although RESAC used the same causal reasoning, SAPS decided that, by proposing a different causal relationship in the two cases, the quality of the forecast can be enhanced. The proposed controller is thus different for each of the two patients.

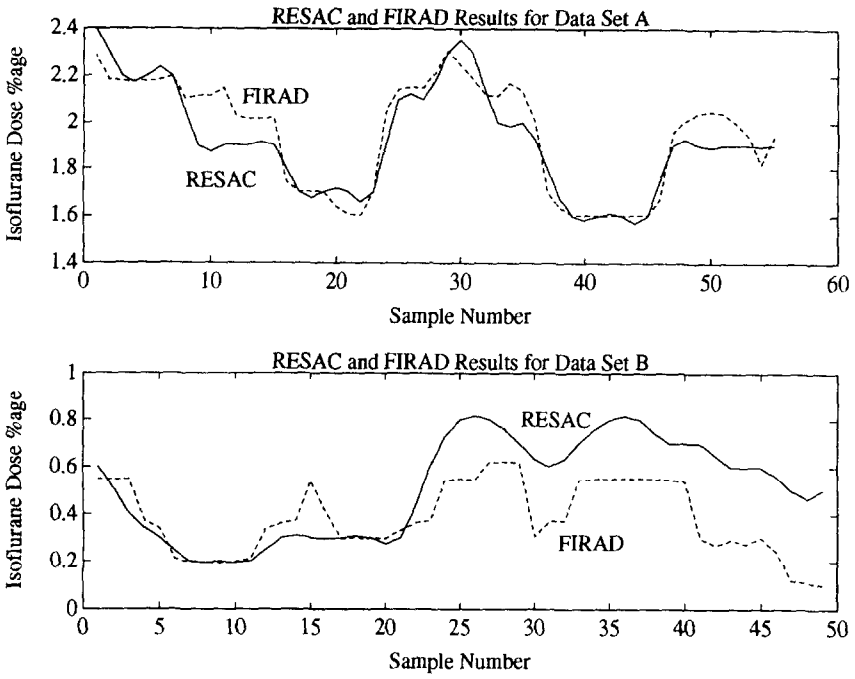


Fig. 5. Comparison RESAC/FIRAD.

One fact that is common to both optimal masks is that the output of the controller model depends on the amount of previously administered anaesthetic agent. This is clinically plausible since the chemical substance accumulates in the patient for some time. The anaesthetic agent used in all these operations was isoflurane. The forecast results for the two data sets are shown in Fig. 5.

The results are quite acceptable. The optimal masks contain sufficient information about the behaviour of the anaesthetist (or RESAC) to be used as a valid controller of the dosage of isoflurane given to the patient. In contrast, the neural network gave good responses for the controller model for only one of the data files.

From these results, we can conclude that the SAPS methodology is fairly robust, i.e., it consistently generates a satisfactory inductive model whenever the data allows it to, and it categorically will not generate a model if the available data do not permit to validate an inductive model.

The neural network approach is different in this respect, since it uses a gradient technique (backpropagation) for optimisation in the original (i.e., continuous) search space, whereas SAPS uses an exhaustive search in a reduced (discrete) search space. Therefore, it is perfectly feasible that the neural network does not converge (as it happened with one of the data records), whereas SAPS will come up with the 'best possible' model (within the framework of the discrete search space) whenever the data justifies a model.

Would it have been possible then to simply invert the data, i.e., use the same data records to generate a patient model as well as a controller model? The answer to this

question is no. Causal modelling is an extension to the concept of univalued functions. Given the function $y = \sin(x)$, it is always possible to find a unique value of y for any given value of x, because $\sin(x)$ is a *univalued function*. On the other hand, $x = \sin^{-1}(y)$ is multivalued, and therefore, it is not possible to conclude a unique value of x given a value of y. It therefore makes sense to call a univalued function a *causal function*, whereas a multivalued function is a *non-causal function*.

Causal modelling is an extension of this concept. Causal models are univalued functions that accept, as inputs, not only current values of its input variables, but also past values of all its inputs as well as its outputs. The controller model is obviously a causal model, since the anaesthetist (or RESAC) bases his or her (its) decision making in a semi-deterministic (fully-deterministic) fashion on the available inputs. The reverse, however, is not true. It is not evident that it is possible to conclude the current (and future) value(s) of the few recorded physiological signals in a unique fashion from measurements of their own past, and from current (future) as well as past Dose values, and SAPS indeed concludes that this is not a meaningful proposition. Considerably more recordings of various physiological signals would be needed to base a valid patient model upon.

4. Comparison of results from the two modelling methodologies

Before comparing the results obtained from ANNAD, and FIRAD, we wish to make a comment about the SAPS methodology. The original idea was that FIRAD should forecast the Dose during 63 min in order to obtain the same plot length as was shown in the previously published ANNAD report [16]. This was not possible because SAPS seemingly needed more data points (training data) than the neural network methodology to identify a model.

Previous investigations involving SAPS have led to a recommendation with respect to the minimum number of data records to be used in the identification of an inductive model. This rule is based on statistical considerations, and states that, in any class analysis, each (discrete) state should be recorded at least five times. Thus:

$$n_{\text{rec}} \geq 5 \cdot \prod_{\forall i} k_i \tag{9}$$

where n_{rec} denotes the total number of recordings, i.e., the total number of observed states, i is an index that loops over all variables and k_i denotes the number of levels (i.e., discrete class values) of the variable i.

In the given application and using the first data stream, the number of suggested records is:

$$n_{\text{rec}} \geq 5 \cdot (3 \cdot 3 \cdot 3 \cdot 2) = 270 \tag{10}$$

Consequently, the first 270 data records should be used for model identification, which leaves us with only 55 records, or 27 min worth of measurement data for forecasting.

To improve the situation, tests were made to find the minimum number of records needed to identify the same controller model that was found using the set of 270 records.

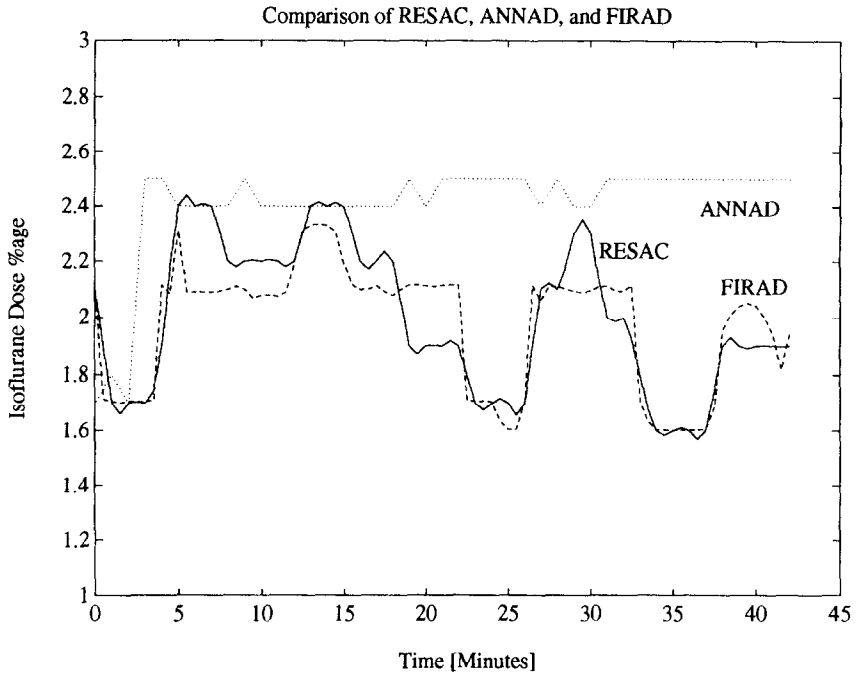


Fig. 6. Comparison RESAC/ANNAD/FIRAD.

It was determined that, if at least 240 records were used for identification, the same controller model could still be found. This then allows us to forecast the system over the last 43 min of the recorded data. This forecast can be compared with the forecast obtained from ANNAD and with the original Dose recommendations made by RESAC. The comparative results are given in Fig. 6.

As can be seen from this plot, FIRAD forecasts the Dose quite well, in fact considerably better than ANNAD. Thus, the FIR methodology has been shown to be able to synthesise inductive biomedical models at least as well as a neural network in this application.

5. Elimination of patient-specific behaviour

In the previous section, it was demonstrated that the FIR methodology is indeed able to obtain a qualitative model for controlling the anaesthetic agent to be administered to a particular patient. However, is this of practical use for the physician?

It does not make much sense, from a medical point of view, to first have to identify a model for a given patient during surgery to be able to predict his or her behaviour at some later time. A reliable model must be ready for use before surgery begins. It is therefore important to be able to synthesise a generic model that is valid for a specific type of patient undergoing a given kind of surgery.

This section presents a knowledge combination technique that allows to merge the knowledge stemming from different patients in order to obtain a general knowledge

base. This knowledge base can then be used for the prediction of future states of a new patient with characteristics similar to those of the patients used for obtaining the knowledge base.

5.1. Combination technique

In order to combine information stemming from different patients, their data sets are concatenated one behind the other. Since the research focuses on models of dynamic behaviour, the advocated methodology searches for causal relationships between variables measured at different points in time. Therefore, if the data set stemming from one patient is placed immediately adjacent to the data set stemming from another patient, fake causal relationships will be created at the seam of the two data streams. These fake relationships can cause a severe degradation of the forecasting power of the derived qualitative model. The solution is to add gaps of 'missing data' [15] between neighbouring data streams stemming from different patients, thereby preventing the methodology from extracting from the combined data set contaminated data records containing mixed information from different data sources.

An important factor to take into account when data sets stemming from different patients are combined is the *normalisation* of the data. Usually, different patients will have different mean values for each variable. Therefore, if data stemming from such patients are to be combined, it is necessary to normalise the data. This process is called *prefiltering* of the data. Both linear and nonlinear prefiltering procedures are known, but only linear prefiltering should be applied in order to prevent a degradation of the relevant correlation functions. In this paper, the mean value of each variable is subtracted from all elements of the corresponding trajectory. This simple normalisation procedure is applied to the data of each patient separately, prior to concatenating their data records.

In order to improve the quality of the prediction and reduce the risk of coming up with entirely incorrect forecasting values, a *voting procedure* is adopted [14]. Instead of working with a single optimal mask, as was done in the earlier parts of this paper, the three best masks are evaluated, and three different state transition matrices are obtained. In the forecasting process, three separate forecasts are obtained using the three state transition matrices.

Let M_a, M_b and M_c be the three best masks. Each of these masks leads to a different forecast. Let them be called F_a, F_b and F_c. Three distance measures are computed in the following way:

$$D_a = \mathrm{abs}(F_a - F_b) + \mathrm{abs}(F_a - F_c) \tag{11}$$

$$D_b = \mathrm{abs}(F_b - F_a) + \mathrm{abs}(F_b - F_c) \tag{12}$$

$$D_c = \mathrm{abs}(F_c - F_a) + \mathrm{abs}(F_c - F_b) \tag{13}$$

Once the distance measures have been computed, the predicted value with the largest distance measure is refused. The new forecast value will be the mean value of the two predicted points obtained with the two remaining masks. For instance, if $D_b > D_a$ and $D_b > D_c$, then forecast F_b is rejected, and the new forecast is computed as:

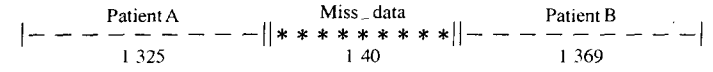
$$F = \frac{F_a + F_c}{2} \tag{14}$$

This technique offers a systematic way to compute predictions for all patients in the patient/operation class.

5.2. Combination technique applied to FIRAD

As was shown in the first part of this paper, the two qualitative models that were obtained for the two patients were distinct. It was not possible to apply either of the two qualitative models to the other patient and obtain meaningful predictions of that patient's future behaviour.

Therefore, it was decided to combine the data from the two patients in order to extract a set of models that are able to offer acceptable predictions for both patients. The two individual data sets contain 325 and 369 records, respectively. A gap of 40 'missing values' were inserted connecting the two data sets. Consequently, a single data set of 734 values resulted to be used in identifying a set of three suboptimal masks.

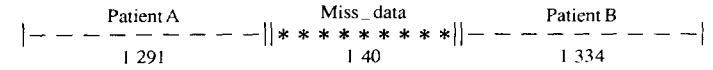


As was to be expected, the mean values of the data sets from the two patients were different, thus, it was necessary to normalise the data in the manner previously described.

At this point, the data were ready to start the model identification process. As before, the variables SAP, HR and Dose were recoded into three qualitative classes each, whereas RR was recoded into two qualitative classes only.

The first 291 rows of patient A combined with the gap of 'missing values' together with the first 334 rows of patient B were used as past history data to compute the optimal mask.

Past history data



The optimal mask obtained for the combined data set was the following:

$$
\begin{array}{c|cccc}
{}_{t}\backslash^{x} & SAP & HR & RR & Dose \\
t-20\delta t & 0 & -1 & 0 & 0 \\
t-19\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-11\delta t & 0 & 0 & 0 & 0 \\
t-10\delta t & 0 & 0 & 0 & -2 \\
t-9\delta t & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
t-\delta t & 0 & 0 & 0 & 0 \\
t & -3 & 0 & 0 & +1
\end{array}
\qquad (15)
$$

This mask denotes the relationship:

$$\text{Dose}(t) = \tilde{f}(\text{HR}(t - 20\delta t), \text{Dose}(t - 10\delta t), \text{SAP}(t)) \tag{16}$$

Notice that the optimal mask is different from either of the two optimal masks found for the two patients separately.

Fuzzy forecasting was then used to predict new qualitative class and fuzzy membership values for variable Dose for the last 34 rows of each patient. It turned out that the prediction obtained using this optimal mask alone was not good enough, and therefore, the previously described voting method had to be applied.

To this end, a set of three suboptimal masks had to be defined. One of the three masks is the optimal mask obtained for the combined data set. The other two masks could be chosen as suboptimal masks from the mask history. However, a different approach was taken. The second mask was obtained using 'common sense'. It has been shown that the two optimal masks obtained for the two patient data sets separately were different. Whereas one reaches the best forecast for patient A, the other does the same for patient B. However, neither of them gave acceptable results for the other patient. It makes sense to think that if the input patterns of the two masks are combined, a good forecast for both patients could be encountered. Following this reasoning, the second mask was constructed as the superposition of the individual masks found earlier in this paper:

$$\begin{array}{c|cccc} {}_t\backslash^x & SAP & HR & RR & Dose \\ t-20\delta t & 0 & 0 & 0 & 0 \\ t-19\delta t & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t-11\delta t & 0 & 0 & 0 & 0 \\ t-10\delta t & -1 & 0 & 0 & -2 \\ t-9\delta t & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t-\delta t & 0 & 0 & 0 & 0 \\ t & -3 & -4 & 0 & +1 \end{array} \tag{17}$$

This mask denotes the relationship:

$$\text{Dose}(t) = \tilde{f}(\text{SAP}(t - 10\delta t), \text{Dose}(t - 10\delta t), \text{SAP}(t), \text{HR}(t)) \tag{18}$$

This mask, when used alone gives worse results than the previous one, but in concert with the other two voting masks, it turns out to be acceptable.

Finally, the optimal mask obtained for the second patient when the two data sets were treated separately was chosen as the third mask in the voting set.

The forecasting results for the two data sets using the voting scheme are shown in Fig. 7.

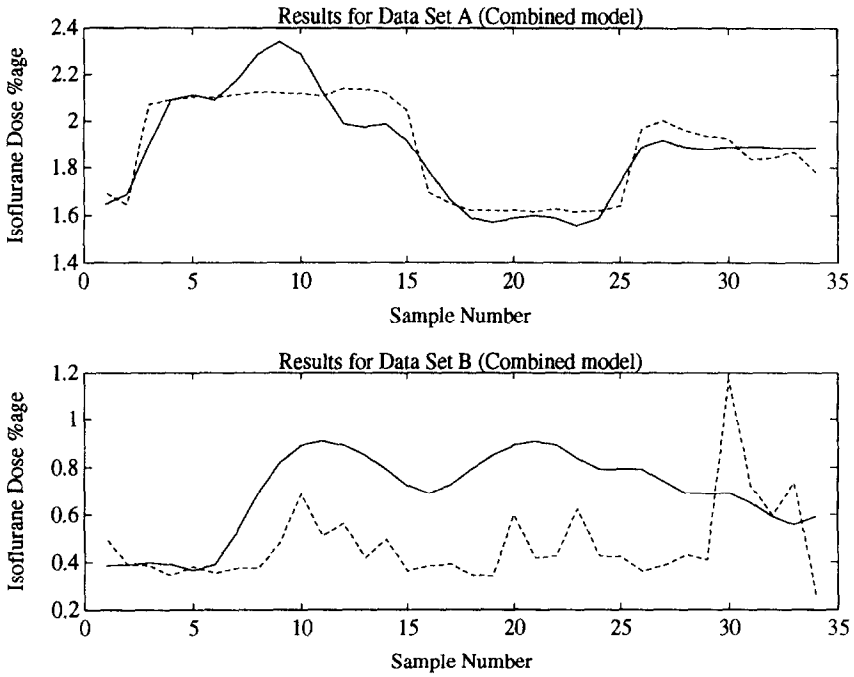


Fig. 7. Prediction results using combined models.

As can be seen, the prediction curve follows the real curve in an acceptable way. The predictions are not as good as those obtained from the individual models shown in Fig. 5, but they are clinically meaningful.

Computing the least square error of the predictions for the two patients, the following results are obtained:

- The forecast *Dose* to be applied to patient A when the individual model is used has an error of 0.4886, whereas the error is 0.5449 when the combined model is used.
- The forecast *Dose* to be applied to patient B when the individual model is used has an error of 1.4156, whereas the error is of 1.8224 when the combined model is used.

Therefore, it is clear that the predictive power has decreased, but not to an unacceptably large extent.

6. Conclusions

The results shown in this paper confirm the ability of the FIR methodology to produce good control performance of the anaesthetic agent delivery system. The FIRAD system not only replicates the advice from RESAC, but it performs better, more reliably, and more consistently than the ANNAD system for the given application.

As demonstrated in this paper, one of the strengths of SAPS is that it contains an inherent model validation mechanism inside the modelling methodology, which the

neural network approach does not. This is why the SAPS methodology would not generate a patient model that was not justifiable from the given data. We consider this intrinsic model validation mechanism a distinctive advantage in comparison with the neural network methodology, especially in the context of soft sciences.

The FIR methodology needs large amounts of rich data to work with. This is necessary in order to be able to record, in the past experience data base, all physically feasible behavioural patterns so that they may later be recognised when they are met again. Evidently, it is important that the data stem from a causal system, i.e., that the outputs of the perceived 'system' are indeed causally related to its inputs.

For the controller model, two different masks have been found from two different data sets. This is not of much practical use from a medical point of view. In order to eliminate patient-specific behavioural patterns from the measurement data, and thereby be able to derive a single model characterizing a specific class of similar patients undergoing similar operations, a technique based on the combination of data stemming from different patients has been presented.

Using the FIR methodology, it has now become feasible to generate a single qualitative model that can be used to predict the future behaviour of patients within an entire class of similar patient/operation pairs. The predictions are not as good as those obtained from individual models, but they are still clinically meaningful.

Acknowledgements

The research presented in this paper has been financially supported, in parts, by CICYT under grant TIC 93-0447. We wish to thank Dr. Rafael Huber of the Institut de Cibernètica, Universitat Politècnica de Catalunya for his support of this project.

References

[1] F.E. Cellier and D.W. Yandell, SAPS-II: a new implementation of the systems approach problem solver, *Int. J. Gen. Syst.* 13 (1987) 307–322.

[2] F.E. Cellier, *Continuous System Modelling* (Springer, New York, NY, 1991).

[3] F.E. Cellier, Qualitative modelling and simulation – promise or illusion, *Proceedings of the Winter Simulation Conference*, Phoenix, AZ, 1991, pp. 1086–1090.

[4] F.E. Cellier, A. Nebot, F. Mugica and A. de Albornoz, Combined qualitative/quantitative simulation models of continuous-time processes using fuzzy inductive reasoning techniques. Int. J. Gen. Syst. 24 (1996) 95–116.

[5] S.G. Greenhow, *A knowledge base system for the control of depth of anaesthesia*, Ph.D. dissertation, University of Sheffield, Sheffield, 1990.

[6] G.J. Klir, *Architecture of Systems Problem Solving* (Plenum Press, New York, NY, 1985).

[7] D. Li and F.E. Cellier, Fuzzy measures in inductive reasoning, *Proceedings of the Winter Simulation Conference*, New Orleans, LA, 1990, pp. 527–538.

[8] D.A. Linkens, Adaptive and intelligent control in anesthesia, *IEEE Control Syst.* 12 (1992) 6–11.

[9] D.A. Linkens, S.G. Greenhow and A.J. Asbury, Clinical trails with the anaesthetic expert advisor RESAC, *Expert Systems in Medicine*, 6th Annual Meeting, London, 1990, pp. 11–18.

[10] D.A. Linkens and H.U. Rehman, Nonlinear control for anaesthetic depth using neural networks and regression, *Proc. IEEE International Symposium on Intelligent Control*, Glasgow, 1992, pp. 410–415.

[11] M.J. Navabi, R.C. Watt, S.R. Hameroff and K.C. Mylrea, Integrated monitoring can detect critical events and improve alarm accuracy, *J. Clin. Eng.* (1991) 295–306.

[12] A. Nebot, *Qualitative modelling and simulation of biomedical systems using fuzzy inductive reasoning*, Ph.D. dissertation, Universitat Politècnica de Catalunya, Barcelona, 1994.

[13] A. Nebot, F.E. Cellier and D.D. Linkens, Controlling an anaesthetic agent by means of fuzzy inductive reasoning, *Proceedings of QUARDET'93, IMACS International Workshop Qualitative Reasoning and Decision Technologies*, Barcelona, Spain, 1993, pp. 345–356.

[14] A. Nebot and F.E. Cellier, Preconditioning of measurement data for the elimination of patient-specific behaviour in qualitative modelling of medical systems, *Proc. CISS'94, First Joint Conference of the International Simulation Societies*, Zurich, Switzerland, 1994, pp. 584–588.

[15] A. Nebot and F.E. Cellier, Dealing with incomplete data records in qualitative modelling and simulation of biomedical systems, *Proc. CISS'94, First Joint Conference of the International Simulation Societies*, Zurich, Switzerland, 1994, pp. 605–610.

[16] H.U. Rehman, D.A. Linkens and A.J. Asbury, Neural networks and nonlinear regression modelling and control of depth of anaesthesia for spontaneously breathing and ventilated patients, *Comput. Methods Prog. Biomed.* 40 (1993) 227–247.

[17] H.J. Uyttenhove, *SAPS-System Approach Problem Solver*, Ph.D. dissertation, State University of New York, Binghamton, NY, 1979.