CAUSAL INDUCTIVE REASONING: A NEW PARADIGM FOR DATA–DRIVEN QUALITATIVE SIMULATION OF CONTINUOUS–TIME DYNAMICAL SYSTEMS

FRANÇOIS E. CELLIER

Department of Electrical and Computer Engineering, The University of Arizona, Tucson, Arizona 85721, U.S.A.
e_mail: Cellier@ECE.Arizona.Edu

JOSEFINA LÓPEZ

Institut de Cibernètica, Universitat Politècnica de Catalunya, Diagonal 647, 2na. planta, Barcelona 08028, Spain
e_mail: Lopez@IC.UPC.ES

Abstract Inductive reasoning attempts to induce future behavioral patterns of a system or time series from observation of the behavioral patterns of their past. No assumption is made about the underlying model structure. Inductive reasoning represents a paradigm of pure behavioral modeling and simulation of systems. Like all other qualitative approaches to reasoning, inductive reasoning fights a constant battle of generality versus specificity, a battle against ambiguity stemming from uncertainty. Causal inductive reasoning tries to win this battle by introducing redundancy into the reasoning process. Causal inductive reasoning can be compared to the president of a company who works with multiple independently operating marketing analysts. All of his or her analysts have to make predictions under uncertainty. If more than one of them comes to the same conclusion, the president will be more likely to heed their advice.

Keywords: Inductive Reasoning; Qualitative Modeling; Fuzzy Systems; Prediction; Redundant Reasoning.

INTRODUCTION

Research in qualitative reasoning has attracted much interest in the artificial intelligence community. This interest was spurred by the observation that humans are much more capable of operating appropriately in an environment of incomplete knowledge than any of the currently available automatic devices. It was perceived that this capability is somehow related to our ability to dealing with knowledge in qualitative terms.

Various qualitative reasoning paradigms have been developed over the past two decades, including expert systems, inductive reasoning, neural networks, qualitative physics, discrete–event systems, and fuzzy systems. Some are purely knowledge–based, others are strictly pattern–based. Yet others represent a merger between these two philosophies.

What they all have in common is that each one of them tries to deal with uncertainty in one way or another, that they all try to reach decisions without being provided with complete information. Qualitative reasoning is thus closely related to the problem of decision making under uncertainty. If everything is known, then qualitative reasoning does not offer any advantage at all over quantitative reasoning. In fact, if everything is known, then a good qualitative reasoning scheme should become quantitative. In other words, a good qualitative reasoning scheme should adapt itself to the amount of knowledge available to it.

The less is known about a situation, the more it is open to interpretation. If everything is known, then there is no margin for interpretation. With perfect knowledge, two experts must reach the same conclusion. However, this is not so in an environment with little knowledge available. Different experts may express different opinions, and there is a wide margin for interpretation of the few facts available, for guesswork, and for making mistakes. The signs are ambiguous, and so will be the conclusions.

If no errors are possible, if the one and only right answer is distinct, it may suffice to provide that answer to the one who asks. Yet, if there is margin for errors, then a decent reasoner should not content itself with providing an answer to a question. It should attempt at providing a measure of uncertainty of that answer as well. This is again a modeling task. This time, the output of the model is not the answer to a question asked, but the margin of error associated with that answer.

It turns out that this modeling task is the more difficult among the two. The reason is that, if the original question was plagued by a certain amount of uncertainty, then the uncertainty associated with the new problem is even larger. Since we are interested not in the answer, but in the error of that answer, the new output is of second order small. Yet, the uncertainty associated with this new output is the same as that associated with the original one. Thus, the relative uncertainty of the new output has grown by an order of magnitude. If a signal is subject to noise, then its derivative signal is subject to even more noise. If the original signal was barely discernible in the noise, than the derivative signal may be lost without a trace.

This is thence our curse: the more ignorant we are, the less likely it is that we shall be able to assess the magnitude of our own ignorance.

QUALITATIVE REASONING

Any form of reasoning requires the application of knowledge. Without any knowledge, we cannot reason. If the knowledge is complex, it needs to be organized in order to be usable. Organization of knowledge is just another name for *modeling.*[1] The use of such organized knowledge is another name for *simulation.*

Qualitative reasoning can thus be viewed as almost synonymous with a step of *qualitative modeling*, followed by one of *qualitative simulation.* In our everyday reasoning processes, this is how we humans reach decisions. We organize the facts available about the problem in question, i.e., we make a mental model of the situation at hand. Then, we design several scenarios that we play through in our minds, i.e., we perform several mental experiments on the available facts, or, in other words, we perform a set of mental simulation runs. We finally choose and implement in real life the one mental experiment that produced the most desirable results in our mental simulations.

Hence, if we wish to gain a deeper understanding of the human reasoning processes, we must learn how mental models are made and how these models are being used in mental simulations.

INDUCTIVE REASONING

The inductive reasoning methodology had originally been developed by Klir[2] as a tool for general system analysis, to study the conceptual modes of behavior of systems. One implementation of this methodology is SAPS–II.[3]

Inductive reasoning models the behavior of time–dependent phenomena by a pure pattern–matching approach. In order to limit the complexity of available patterns, inductive reasoning simplifies the observed patterns in a process of discretization, called *recoding* in the inductive reasoning methodology. Real–valued signals are recoded into a –usually rather small– set of discrete classes. For example, instead of measuring temperature in degrees centigrade, we classify temperature as being either 'cold,' 'fresh,' 'moderate,' 'warm,' or 'hot.'

Evidently, some knowledge is lost in the process of recoding. The sentence: "Today, it was fairly hot." contains less information than "Today, the temperature peaked out at 27.9 degrees centigrade, the average afternoon temperature was 25.5 degrees, and now, at 8:15 p.m., the temperature is still at 21.3 degrees." The recoded model only contains the information that the afternoon had been 'hot,' and that, somewhere around 6:30 p.m., the temperature switched from 'hot' to 'warm.'

Evidently, any qualitative predictions made on the basis of such a qualitative model will be limited to the same five classes, i.e., we cannot hope to make predictions that are more precise than the qualitative model we are working with.

Selecting the number of discrete classes for representing each of the variables in the system relates to the struggle between generality and specificity. The more levels are chosen, the larger will be the expressiveness (specificity) of the qualitative model. However, this goes hand in hand with an increased difficulty of making predictions, with the need for more and more data. The smaller the number of levels chosen, the better will be the predictiveness (generality) of the model, but the less useful will these predictions be. If every variable is recoded into exactly one level, then the model will be infinitely predictive, yet infinitely useless.[4]

Inductive reasoning consists of a step of inductive modeling followed by a step of deductive simulation.[5] In the inductive modeling step, a qualitative model[6] is induced in the form of a finite state machine relating qualitative inputs to qualitative outputs. An abstraction mechanism[5] is employed that determines which input variables to look at when we wish to conclude something about a particular output variable.

A possible relation among the qualitative variables of a five–variable system could be of the form:

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

where $\tilde{\mathrm{f}}$ denotes a qualitative relationship. Notice that $\tilde{\mathrm{f}}$ 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 inductive reasoning 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 methodology), Eq.(1) is represented by the following matrix:

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

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 inductive reasoning, 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 (i.e., recoded) behavior to which it should be applied, and it has a certain number of rows, the *depth* of the mask.

The optimal mask is the one abstraction that leads to the most deterministic input/output behavior. The problem of finding the optimal mask relates also to the struggle between generality and specificity. If more m–inputs are added to the mask, the observed patterns become more and more specific. Yet, chances are that a newly observed input pattern has never been seen before, making a prediction impossible. Removing m–inputs from the mask leads to bolder, less specific, patterns that are likely to be ambiguous. The so obtained model no longer represents the true dynamics of the system, leading to non–deterministic input/output behavior, i.e., to ambiguities in the predictions made.

Once the optimal mask has been found and the corresponding Finite State Machine (FSM) generated, forecasting future system behavior is almost trivial. All that needs to be done is compare newly observed input patterns with those stored in the FSM, and read out the corresponding output patterns. If the FSM is not totally deterministic, i.e., if for the same input pattern several different output patterns have been observed in the past, then we have several choices: (i) we can report all possible outcomes together with their previous relative observation frequencies, (ii) we can limit the reporting to the most likely outcome, or (iii) we can draw a random number and predict any one of the previous observations with the correct probability of occurrence. The term "probability" is to be understood in this context in a statistical rather than theoretical probabilistic sense.

We can use the relative frequency of occurrence of an output pattern for any given input pattern as a measure of correctness of the prediction made. Of course, it is not only the prediction that contains an error, but the quantity that measures the correctness of that prediction contains itself an error, and as we meanwhile know, the trustworthiness of this measure is even lower than that of the original prediction. Yet, it is the best we can do under the given circumstances, and it is certainly better to provide this measure of correctness than none at all.

In truth, the inductive reasoning methodology works amazingly well. Two accounts of successful applications of this methodology to a linear system and a Boeing 747 airplane in high altitude horizontal cruise flight were reported.[7–8]

FUZZY INDUCTIVE REASONING

As was mentioned earlier, the problem of the ambiguity of predictions is intimately linked to the problem of incomplete knowledge. Since we are living in an imperfect world, we may not be able to avoid the problem of incomplete knowledge, and hence we must live with the realities of ambiguous signs and predictions. However, we can minimize the problem by at least exploiting all the information given to us.

It was mentioned before that, in the process of recoding, a lot of precious information about the system under study is thrown away. This is a pity, and means should be sought that would avoid throwing away this information. One such means is the introduction of fuzzy measures into the methodology.[9–12]

In our own dialect of Fuzzy Inductive Reasoning (FIR), real–valued variables are recoded into qualitative triples, consisting of the previously introduced class value augmented by a fuzzy membership function value and a side value. As before, the class value represents a coarse discretization 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. 1 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 neighboring 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.

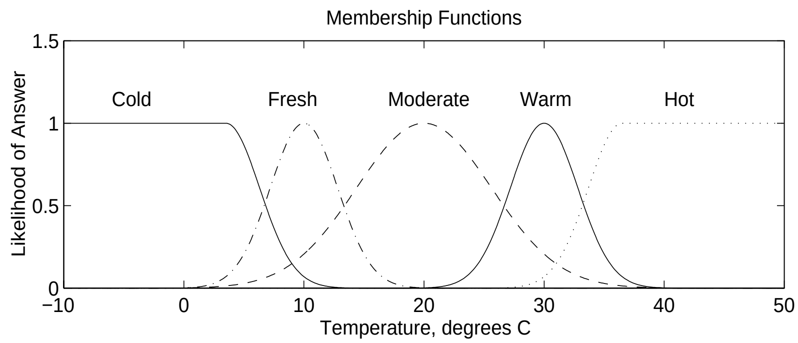

Figure 1: Fuzzy Recoding of Outside Temperature

In the fuzzy systems literature, the process of converting a quantitative (real–valued) variable into a qualitative (fuzzy) variable is called *fuzzification.* Clearly, fuzzy recoding qualifies as a fuzzification method.

The determination of the optimal mask and the process of forecasting future behavior remain basically the same as in the original inductive reasoning methodology. However, instead of measuring the likelihood of a certain outcome through the frequency of its occurrence, the fuzzy membership function values are now used to compute a measure of confidence in that outcome. Rather than first throwing valuable information away in the process of recoding, information that we shall have difficulties later on estimating again, the confidence information is kept throughout the qualitative analysis. This sharpens the discrimination power of the methodology, and thereby significantly reduces the ambiguity of predictions made by it.[13]

Although it is correct that fuzzy inductive reasoning sometimes leads to the selection of a different optimal mask than the original inductive reasoning technique, this fact is not important. What *is* however important is that the additional information contained in the fuzzy membership function values of the past history recordings that are stored in the FSM leads to a much improved selection of relevant past observations in determining a new prediction. The *five–nearest–neighbors (5NN)* method[14] is used to select the most relevant similar patterns as they had been observed in the past.

Equally if not more important is the fact that the enhanced methodology enables us to make predictions not only of the class values of an output variable, but also of its associated membership function and side values. This in turn makes it possible to make *quantitative predictions* of output variables by means of inverting the process of fuzzy recoding. In the fuzzy systems literature, this process is called *defuzzification.* It has been demonstrated that the predictions obtained in this fashion are good enough for mixed quantitative and qualitative simulations,[15–17] as well as for the systematic design of fuzzy controllers.[18–19]

It should also be mentioned that the additional information available to the inductive reasoner reduces the uncertainty of the estimation of the prediction error. If it is decided to pursue only the most likely estimation, this will increase the confidence in the correctness of the prediction made. If it is decided to pursue the envelope of all feasible predictions, it will reduce the angle of the prediction cone, thereby allowing us to predict over a longer time span into the future before the prediction becomes pure speculation.

CAUSAL INDUCTIVE REASONING

Although *Fuzzy Inductive Reasoning (FIR)* makes use of all the information available, this does not mean that there is no longer any margin of uncertainty or risk of making mistakes. There are two causes for this:

1. Although all the information originally available has been preserved for use by the inductive reasoner, this does not mean that complete information had originally been available. One would need infinitely long data files to capture every possible comportment of a system through a collection of behavioral input/output data alone.

2. Although the inductive reasoner has all originally available information at its disposal, this does not necessarily mean that it will make use of the available information in an optimal fashion, i.e., information can get lost in the process of reasoning. For example, why do we propose to use exactly five neighbors in the process of prediction? Why not four or six? Should we not use all training data points for the prediction? Don't we again throw

potentially valuable information away by restricting ourselves to precisely five neighbors?

There is little than can be done about the former of the two problems. We must strive to compute an informed estimate of the percentage of information missing, since this knowledge will certainly influence the quality of our estimation of the prediction error. This can be done in various ways, but we don't have the space here to go into any details. We shall only mention that the argument can also be turned around. Our estimation of the prediction error tells us something about the degree of completeness of the available information. If the estimate of the prediction error is very large, the software may alert the user to this fact and suggest to collect more data before proceeding with the qualitative analysis.

The latter problem is the focus of our concern in this paper. There is some good reason for not including all the available data points in the formula for predicting the membership function value of an output variable. The larger the distance of a previously observed (training) input pattern is from the current (testing) input pattern for which we wish to predict an output value, the less relevant will that information be for the case at hand. We can also say, that the correlation between past and current patterns decreases with distance (not in time, but in similarity of input patterns).

The term "correlation" carries the association of *statistics*. Isn't what we are doing in inductive reasoning just a very strange way of performing nonlinear statistics on the available data? The answer to this question must be an enthusiastic *yes*. "Performing statistics on measurement data" is just another expression for "inducing knowledge from data." The classical (linear) statistical methods are simply the oldest and best established routes to doing so; thus, if we propose alternate routes to reaching the same goal, then we should not ignore previous knowledge, but study it and exploit what is salvageable from it. Indeed, the inductive reasoning methodology carries many common threads with nonlinear statistics.

In principle, everything that can be accomplished by a neural network or an inductive reasoner, can also be accomplished by nonlinear statistics, and vice versa. The question is not, which of the different approaches can accomplish something, but which of them can accomplish it most easily. In this respect, inductive reasoning fares quite well in comparison with the other two approaches in the context of predicting future behavior of dynamic systems, as our previous publications have shown.

Yet, the problem of making uncertain predictions remains, and we must ask ourselves, what can be done to reduce it. Since we already have preserved all the information originally available, evidently, there is nothing we can do to gain more information except ask the user to collect more data. Thus, we must make use of the available knowledge in optimal ways. The new idea here is to introduce *redundancy* into the reasoning process. This is exactly what a human decision maker would do under the same circumstances: employ the services of several independently operating advisors and hear them all before reaching a final decision.

We can do precisely the same in SAPS. In a recent paper,[20] it was proposed to make use of three suboptimal masks, and use a *voting scheme* in reaching a decision about the final prediction. A distance measure between the three predictions is introduced, the "advisor" with the largest sum of distances to its two competitors is eliminated, and a mean value between the survivors is used as the final prediction.

It would be easy to improve on this scheme. The "boss" can increase the "salary" of an advisor who performed well in the past. In terms of SAPS–II, we could take new incoming data into account to check how well the three "advisors" did in predicting the correct outcome, and increase the influence of the most successful "advisor" by a certain percentage, whereas that of the least successful one is reduced by the same amount.

However, we can also introduce redundancy at other places. In *Causal Inductive Reasoning (CIR)*, it is proposed to recode real–valued variables into qualitative quadruples. Each quadruple contains the same three pieces of information that were used in FIR, plus a qualitative derivative value that indicates whether the recoded variable is currently increasing, decreasing, or staying at about the same level. We can then select the five nearest neighbors on the basis of a modified formula for the distance function that includes the additional information retrieved from similar past derivative patterns as an additional piece of information.

In FIR, the distance function was computed in the following manner. Given the class value, $class_{ij}$, the membership value, $Memb_{ij}$, and the side value, $side_{ij}$, of the i^{th} m–input of the j^{th} training data record, we can compute a position function as follows:

$$pos_{ij} = class_{ij} + side_{ij} \cdot (1.0 - Memb_{ij}) \tag{3}$$

This is possible because, in SAPS–II, variables of enumerated types, such as the class and side variables, are represented through integers.

The pos_{ij} values are quantitative (real–valued) variables that can be used to represent the relative magnitude of a particular qualitative triple. However, they are *not* regenerations (defuzzifications) of the original quantitative signals. They are normalized variables. Irrespective of whether an original signal was very small, ranging from -10^{-15} to $+10^{-14}$, or very large, ranging from 10^6 to 10^{12}, the corresponding pos_{ij} signal ranges *exactly* from 0.5 to 1.5 for values in class '1,' from 1.5 to 2.5 for values in class '2,' etc. Consequently, different pos_{ij} signals can be compared to each other or can be summed up, without weighing them relative to each other, something that would not be meaningful using the original or regenerated quantitative signals. The normalization function *is* a transformation from a qualitative triple to a quantitative variable, but this variable lives in a different space from the original quantitative variable.

The pos_{ij} values corresponding to the different variables of an input state are then concatenated to form the vector:

$$\mathbf{pos_j} = [pos_{1j}, pos_{2j}, \ldots, pos_{nj}] \tag{4}$$

assuming, the training data record contains n m–inputs. We call the vector $\mathbf{pos_j}$ the *norm image* of the j^{th} training data record.

Finally, the distance function dis_j is computed as the $\mathcal{L}_2$ norms of the differences between the $\mathbf{pos}$ vector representing the norm image of the testing data record and the $\mathbf{pos_j}$ vectors representing the training data records:

$$dis_j = \|\mathbf{pos} - \mathbf{pos_j}\| \tag{5}$$

The five training data records with the smallest distance functions are then chosen as the five nearest neighbors of the new testing data record.

In CIR, we take also the qualitative derivative information into consideration in the process of selecting the five nearest neighbors. To this end, we define a qualitative derivative vector:

$$\mathbf{der_j} = [der_{1j}, der_{2j}, \ldots, der_{nj}] \tag{6}$$

where der_{ij} assumes a value of '-1,' if the i^{th} m–input of the j^{th} training data record has a decreasing tendency, '$+1$,' is the tendency is increasing, and '0,' if the quantitative value of the variable changes relatively little.

We then modify Eq.(5) as follows:

$$dis_j = \|\mathbf{pos} - \mathbf{pos_j}\| + k \cdot \|\mathbf{der} - \mathbf{der_j}\| \tag{7}$$

The constant k determines how much weight is to be assigned to the derivative information in the overall computation of the distance function.

As the qualitative derivative information contains much less correlation (relevance) than the qualitative state information, we must be cautious when taking this information into consideration. Somehow, this is similar to allowing more and farther away neighbors to contribute to the equation, or allowing more and more "advisors" (masks with lower and lower quality) to contribute to the final decision. Evidently, this is a sword with two edges. k is a tuning parameter of the algorithm. At the time of writing this paper, we are still experimenting with different examples to determine the optimal value of k to be used. Although we could leave k as a user–tunable parameter in the software, we don't want to do so, unless we can provide the user with a clear recipe of how to determine, which k value is likely to perform well in his or her application.

Up to this point, we have made use only of the qualitative derivatives of the m–inputs. What about the derivative information for the m–outputs? Evidently, we can compute the qualitative derivative of m–outputs for the training data records. Since this is qualitative information, we can predict the qualitative derivative of the m–output of any testing data record from the information contained in its five nearest neighbors. Since we also predict the state values, from which we can again derive values for the qualitative derivatives, we now have *redundant information.* We have estimated the qualitative derivative of the output of the testing data record in two different ways.

How are we going to use this information? We could reject predictions that lead to inconsistent derivative information. However, this is dangerous. The derivative information contains implicitly a lot of assumptions about the continuity of the output signal. Discrepancies in the predictions will occur precisely at times when the signal changes fast, whereas the predictions will be consistent when not much happens. This is like passing the output through a low–pass filter. Evidently, this reduces the sensitivity to noise, but it simultaneously reduces the alertness to fast changes in the signal. We can't have it both ways. Either we are alert, with the risk of receiving many false alarms, or we dull ourselves, with the greater risk of missing the true action when it takes place.

However, we can still make use of this information. It was mentioned earlier how important it is to obtain, together with the output, an estimate of the correctness of that prediction. Evidently, if the redundant derivative computation is consistent, this will raise our confidence in the prediction made. If it is inconsistent, we should be suspicious that our prediction might be wrong. In the original inductive reasoning algorithm, we had to use the frequency of past occurrences of the same input/output pattern as an indicator of the correctness of a prediction made, which is a fairly poor indicator function. In FIR, we were able to improve the acuteness of estimating the prediction error by using the confidence information available through the fuzzy membership functions. In CIR, we can sharpen the resolution further by making use of the *redundancy* inherent in the computations.

CONCLUSIONS

In this paper, an enhancement to the fuzzy inductive reasoning methodology was presented. This enhancement introduces redundancy into the process of qualitative computations. This redundancy can then be exploited to sharpen the resolution of qualitative decision making based on quantitative observations of dynamic phenomena. This is important since all reasoning based on incomplete information is plagued by uncertainty and margins of error. Hence every prediction made should be accompanied by an estimate of its accuracy. Since this estimate itself is imprecise, and in fact, even more so than the original prediction made, one must strive to exploit the limited available knowledge in optimal ways to minimize the margin of error. Introducing redundancy into the qualitative computations accomplishes the desired goal.

ACKNOWLEDGMENTS

We wish to express our thankfulness to Prof. Dr. Gabriela Cembrano and Prof. Dr. Rafael Huber of the Universitat Politècnica de Catalunya for their continued support of this research effort. Their readiness to discuss these and other research questions with us at all times as well as their invaluable help with organizational matters are highly appreciated.

REFERENCES

1. B. P. Zeigler, Multifaceted Modeling and Discrete Event Simulation (Academic Press, London, 1984).

2. G. J. Klir, Architecture of Systems Problem Solving (Plenum Press, New York, 1985).

3. F. E. Cellier and D. W. Yandell, International Journal of General Systems, 13, 4, 307-322 (1987).

4. F. E. Cellier, Continuous System Modeling (Springer–Verlag, New York, 1991).

5. H. S. Sarjoughian: Inductive Modeling of Discrete–Event Systems: A TMS–Based Non–Monotonic Reasoning Approach, Ph.D. dissertation, Dept. of Electr. & Comp. Engr., University of Arizona, Tucson, AZ (1995).

6. F. E. Cellier, in Proceedings Winter Simulation Conference (Phoenix, AZ, December 8-11, 1991) pp. 1086-1090.

7. F. E. Cellier, International Journal of General Systems, 13, 4, 333-344 (1987).

8. P. J. Vesanterä and F. E. Cellier, Simulation, 52, 3, 111-121 (1989).

9. G. J. Klir and T. A. Folger, Fuzzy Sets, Uncertainty, and Information (Prentice–Hall, Englewood Cliffs, N.J., 1988).

10. G. J. Klir, in Modelling and Simulation Methodology: Knowledge Systems' Paradigms, edited by M. S. Elzas, T. I. Ören and B. P. Zeigler (Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1989).

11. D. Li and F. E. Cellier, in Proceedings Winter Simulation Conference (New Orleans, LA, December 5-8, 1990) pp. 527-538.

12. Z. Wang and G. J. Klir, Fuzzy Measure Theory (Plenum Press, New York, 1992).

13. A. de Albornoz and F. E. Cellier, Simulation, 62, 6, 354-364 (1994).

14. F. Mugica and F. E. Cellier, in Proceedings International Symposium on Artificial Intelligence (Monterrey, Mexico, September 20-24, 1993) pp. 372-379.

15. F. E. Cellier, A. Nebot, F. Mugica and A. de Albornoz, in Proceedings IFAC Symposium on Intelligent Components and Instruments for Control Applications (Málaga, Spain, May 20-22, 1992) pp. 589-593.

16. F. E. Cellier, A. Nebot, F. Mugica and A. de Albornoz, International Journal of General Systems (1995) to appear.

17. A. Nebot, Qualitative Modeling and Simulation of Biomedical Systems Using Fuzzy Inductive Reasoning, Ph.D. dissertation, Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain (1994).

18. F. Mugica and F. E. Cellier, in Proceedings European Simulation MultiConference (Barcelona, Spain, June 1-3, 1994) pp. 523-528.

19. F. Mugica, Diseño Sistemático de Controladores Difusos Usando Razonamiento Inductivo, Ph.D. dissertation, Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain (1995).

20. A. Nebot and F. E. Cellier, in Proceedings First Joint Conference of International Simulation Societies (Zurich, Switzerland, August 22-25, 1994) pp. 584-588.