Medical Progress through Technology **21:** 147–158, 1996.

Methodology for a medical expert system on fuzzy analog ganglionar lattices

Non-approximate reasoning with multiple antecendents of different relative importance and limited uncertainty

Carlos A. Holzmann, Alfonso Ehijo & Claudio A. Perez
Biomedical Engineering Group, Department of Electrical Engineering, University of Chile, Av. Tupper 2007, Casilla 412-3, Santiago, Chile

Accepted 16 August 1995

Key words: expert reasoning, explicative capacity, fuzzy analog ganglionar lattices, medical expert system, prospective capacity

Abstract

This work presents an Expert System based on fuzzy analog ganglionar lattices. Its reasoning scheme is designed analogously to the expert's mental organization and it is realized on an (analog) operator called the ganglionar lattice. It is a connectionist system that uses the medical knowledge to define its architecture. The operator evokes some similarities to higher order neural networks and performs as the knowledge base and inference engine of the expert system, in a unified manner. A main feature of this operator is that it exhibits the variables corresponding to all intermediate concepts identified by the expert; this characteristic is shown to be most valuable for assessing, explicating and prospecting in medical applications. Further, it is capable of (i) evaluating a consequent for a variety of non-approximate reasonings with multiple antecendents of different relative importance under limited uncertainty; (ii) explicating the conclusions at different levels of abstraction to suit the user; and (iii) prospecting for the best 'a priori' sequence of unevaluated antecedents, from which to choose following tests. These procedures are based on the objective criterion of the consequent's uncertainty decrease (entropy). All results are produced in numerical form and may be translated into restricted natural language. A simple example of this technology is fully developed. Finally the method's potentials are discussed for future applications.

Abbreviations: AAVG – arithmetic average, BR – basic reasoning, ci – cardiac insufficiency, cs – lung congestive signs, dy – dyspnea, ES – expert system, EXP – explicative capacity index, GL – ganglionar lattice, he – hemoptysis, hs – heart signs, le – lung edema, mGL – ganglionar lattice operator, nBR – basic reasoning operator, PP – probabilistic product, PRO – prospective capacity index, PS – probabilistic sum, S1 – only number one antecedent matters operator, S-T – triangular norm and co-norm, ta – tachycardia, UNC – uncertainty index, WAVG – weighted average, 4b – heart 4^{th} bruit

I. Introduction

This work presents an Expert System, ES, for medical applications based on fuzzy analog ganglionar lattices, GL. The mathematical operators defined and the overall structure of the model have some similarities to the activation function and architecture encountered in higher order neural networks [1]. Nevertheless the model shows important differences because units are in 1-1 correspondence with useful intermediate concepts identified by the expert, and its structure is defined on an order relation among the units. This structure allows the use of well established medical knowledge relating clinical manifestations, syndromes, pathological states, etc., and incorporating them to the architecture. Whenever this medical knowledge is not explicitly available, the model behaves similarly to a neural network relating a pattern of inputs to the output.

Thus the system could be defined as a connectionist model [2–4] using knowledge to define its architecture and data to identify its coefficients. Nevertheless this technology presented here is not knowledge-based in the classical sense. For comprehensive reviews of knowledge-based systems see [5–8].

As it is well known, experts do not agree in building an explicit model for their reasoning [9]. The assumption here is that the architecture of the ES will reflect the expertise of a single specialist, or that of a group of them agreeing on the reasoning strategy.

This ES for medical applications based on fuzzy analog GL has been under development for several years [10–13]. The main assumptions behind this ES are: (i) the reasoning can be simulated by an analogy to the expert's mental organization of his formal and heuristic knowledge, and (ii) a complex medical reasoning can be decomposed into Basic Reasonings, BR, each corresponding to a concept with useful significance to the expert. For example, the diagnosis of a disease can be decomposed into syndromes, pathological states, clinical findings, etc. A similar decomposition has been proposed earlier [14].

The inputs to this model are all the primary (directly observable) antecedents that are relevant to a decision. These antecedents are not precisely measurable, thus they are fuzzily specified as the patient's degree of membership of each primary antecedent's intensity. In medical diagnosis, for example, the primary antecedents are the patient's subjective sensations (symptoms), appreciation or measurement of signs and results of instrumental procedures, etc. The output of the model is the consequent's intensity; e.g., intensity of a disease in the patient.

The model relates the most abstract consequent (top), to the primary antecedents through a hierarchical structure made of *Basic Reasoning units, BRs*. This structure is defined by the order relation of 'inclusion'; thus it is the lattice [15] of the consequent. The consequent of any intermediate BR is an antecedent to superior BRs in the structure; and the antecedents of any BR are consequents of lower BRs and/or primary antecedents.

Up to now, each BR has been realized by a sublattice of units capable of operating on two antecedents only [10, 11]. Each unit was an operator specified by the kind of relationship among its inputs: associative type ('and'), non-associative type ('or'), and excluding operator, among others. This realization proved deficient because it forces the model to an artificial decomposition of each BR in units of the various types, which are difficult to distinguish, and produce a non-trivial growth of the lattice. This enlargement blurs the association to useful concepts which is needed to explicate and justify the conclusions. Moreover, this growth prevents a more precise identification of the coefficients.

An iterative computational procedure to identify the coefficients of the model, based on the gradient method, was first used in [11]. These parameters were proportionally adjusted, considering their sensitivity and the error between the model and the specialist, for a sequence of training cases. This iterative procedure is similar to the back-propagation method used in neural networks to adjust weights [16].

In the earliest work [10], the lattices to evaluate six cardiopathies were developed according to the specialist's reasoning. They were successfully tested with data from clinical records. Later this method was applied to determine the degree of normality of the ECG [11], and to diagnose, characterize essential hypertensive patients [17].

The main purpose of this work is to formalize a methodology and to develop the mathematical tools for an ES for medical applications based on ganglionar lattices, surmounting the above discussed deficiencies. This paper includes substantial improvements which will eventually permit treating some of the expert's abilities and limitations, such as the reasoning with imprecise antecedents, and his approximate reasoning, respectively [18].

In section II, the nBR operator is developed and realized, detailing some of its properties. Then the mGL operator is defined, supplying methods for parameter identification and emphasizing the role of its structure. In section III the advantages of this type of ES are demonstrated by developing methods for explicating the system's conclusions and for prospecting the unevaluated antecedents, both at different levels of abstraction to suit the user's needs. In section IV a detailed application example is shown. Finally, a discussion of results and conclusions for future improvements are provided.

II. Methodology

Preliminary definitions

Often a specific complex medical reasoning leading to evaluate a consequent, can be decomposed into a lattice of BRs [10, 11, 14]. These units are hierarchically organized by the order relation of 'inclusion', thus their structure is a lattice in the sense of Birkhoff [15]. The model proposed here, applicable when the reasoning problem yields to this decomposition, shall be called the *Ganglionar Lattice, GL*, of the specific consequent. For example, the diagnosis of some diseases can be decomposed into an organization of syndromes, clinical findings, pathological states, etc., which finally would depend on signs, symptoms, etc. [10, 11, 13, 14]. When the above decomposition is not known, the model reduces to a single BR. This last basic idea was earlier suggested by Zadeh [19] in connection with fuzzy theory applied to medical diagnosis. In what follows the methodology for defining and using the ES based on GLs is presented, illustrated, and discussed. It must be emphasized that the procedure to define the GL is most suited for incorporating a single expert's formal and heuristic knowledge to the model.

Consider the set of *relevant primary* antecedents to a top consequent. An antecedent is *relevant*, when the expert's formal or heuristic knowledge asserts this; and is *primary* when it is directly observable. Of this set, consider only those subsets bearing a useful significance to the expert in relation to the consequent, i.e., useful significance for reasoning with the antecedents, explaining the consequent, or communicating partial aspects of the problem. Following the definition of a lattice [15], this family usually contains a few of the 2^n possible subsets of the power set of primary antecedents, including always the full set at the cusp (the consequent), and the singletons next to the base (the primary antecedents). When the relation of 'inclusion' is used to order this family of subsets, the resulting structure is the ganglionar lattice of the top consequent, based on the relevant primary antecedents. This construction is exemplified in section IV, part 1.

Associate a *Basic Reasoning unit, nBR*, to each node of the above defined lattice, except to the nodes next to the base (primary antecedents). The outgoing arcs at any node represent a single consequent, which is antecedent of nBRs at higher levels in the lattice. The incoming arcs at any node represent the distinct antecedents specified by the lattice; each of these antecedents is identified with a consequent of an nBR at inferior level. For the top nBR the output variable is the top consequent, and for others, some input variables are primary antecedents. *Thus, an nBR is an operator realizing the specific relationship among the 'n' antecedents by which its consequent can be evaluated at each node of the lattice.* This operator is thought out to simulate different types of human reasoning, i.e., the types of relationships employed to assign values to the consequent. The nBRs are 'basic' in the sense that no further decomposition is meaningful to the expert. Each nBR, together with its consequent, bears the name of the concept associated with the subset of primary antecedents covered by the concept; this name and subset is previously identified by the expert.

Realization of a basic reasoning unit

For the realization of an nBR, a multivariable polynomial operator is proposed. This operator puts forth the favorable characteristics of a weighted average operator [20–23] to treat antecedents with different importance, and also exhibits the advantages of the triangular norm and co-norm operators [23, 24], using the probabilistic sum and product, to simulate different forms of reasoning. This family of operators allows the propagation of the intensity, and eventually of the uncertainty, from the antecedents to the consequent.

Definition

Mathematically, the *Basic Reasoning unit* is the operator nBR: $\underline{s} \in I^n \rightarrow nBR(\underline{s}) \in I$, where $I = [0,1]$, and $\underline{s} = [s_1,..,s_n]$ is a vector of input variables representing the 'n' antecedents. The explicit expression for the nBR is

$$nBR(\underline{s}) = \sum_{i=0}^{2^n-1} a_i(\prod_{j=1}^{n}(s_j)^{b_{ij}})$$

where the {0,1}-structural matrix $[b_{ij}]$, $i = 0,...,2^n-1$, and $j = 1,...,n$, represents the composition of the s-products among the 'n' antecedents. If the i^{th} term contains the factor s_j, then $b_{ij} = 1$ and otherwise $b_{ij} = 0$. For example,

$$2BR(s_1, s_2) = a_0 + a_1 s_1 + a_2 s_2 + a_3 s_1 s_2$$

$$3BR(s_1, s_2, s_3) = a_0 + a_1 s_1 + a_2 s_2 + a_3 s_3 + a_4 s_1 s_2 + a_5 s_1 s_3 + a_6 s_2 s_3 + a_7 s_1 s_2 s_3$$

The number of operations to compute the consequent of an nBR are $(2^n-1) + \Sigma^n{}_{j=2}(j-1)\{n!/(j!(n-j)!)\}$ products and (2^n-1) sums.

It is clear that the nBR operator includes not only the proportional contribution of each antecedent, but also that of all possible products among them. These products may be interpreted as dependencies or couplings among antecedents.

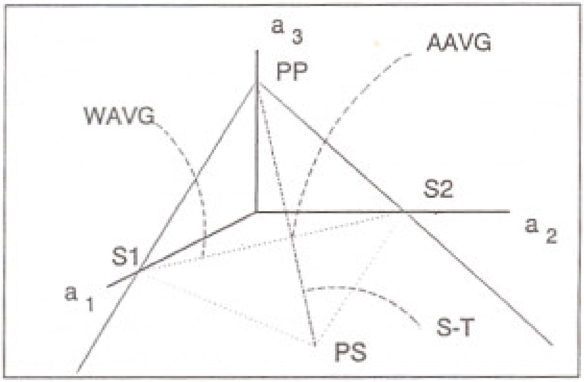

Fig. 1. Hyperplane determined by a1 + a2 + a3 = 1 on the coefficients space (a1, a2, a3) showing the locus for the allowed coefficient values of a 2BR. The coefficients permit identifying the type of reasoning obtained: locus for the Weighted AVeraGe operators, {WAVG: $\alpha(1,0,0) + (1-\alpha)(0,1,0)$, $\alpha\epsilon$I}; the locus of the triangular norm and co-norm operators, {S-T:$\beta(0,0,1)$+ $(1-\beta)$ (1,1,-1), $\beta\epsilon$I}; and the singular cases of: (1) the Probabilistic Product, PP: (0,0,1); (2) the Probabilistic Sum {PS: (1,1,-1)}; (3) the 'only s_1 matters' operator, {S1: (1,0,0)}; and (4) the Arithmetic AVeraGe operator {AAVG: (0.5,0.5,0)}.

Uniqueness of the nBR

It is important to notice that if a family of 2^n linearly independent inputs produces 2^n outputs through an nBR specified by a set of coefficients a_k then, there is no other set b_k that, for the same inputs will produce the same outputs. This can be proven by the linearity of the 2^n coefficients a_k. Thus an nBR operator is uniquely specified by its 2^n coefficients.

Restrictions to the nBR operator

In order to represent the absolute normality (abnormality) of a variable, the number '0' ('1') is used. Thus, if the absolute normality (abnormality) of all antecendents is to imply the absolute normality (abnormality) of the consequent, it is sufficient that $a_0 = 0$ ($\Sigma_i a_i$ = 1). In other words, $a_0 = 0$ ($\Sigma_i a_i = 1$) implies nBR($\underline{0}$) = 0 (nBR($\underline{1}$) = 1). The restriction $a_0 = 0$ indicates that the coefficient space is of dimension 2^n-1, and the restriction $a_1 + ... + a_q = 1$, for $q = 2^n$-1, describes an hyperplane in the coefficient space, intersecting every axis at the value '1'. Notice that, if $s_i = 1$ while all other inputs are '0', then nBR = a_i; thus $a_i \epsilon$I, for i = 1,...,n.

The type of reasoning performed by an nBR is specified by its coefficients. Figure 1 shows the allowed region in the 3-D coefficient space of a 2BR. Several types of reasoning can be identified: (i) decoupled equal importance antecedents; (ii) decoupled unequal importance antecedents; (iii) the consequent is never weaker than the strongest antecedent (pessimistic reasoning); and (iv) the consequent is never stronger than the weakest antecedent (optimistic reasoning), among others.

The ganglionar lattice operator

A *ganglionar lattice operator, mGL,* is a functional composition of the nBR operators composed as the nodes of a previously defined lattice for a particular consequent. Its output variables represent the consequent, the input variables are the 'm' primary antecedents, and the intermediate variables represent intermediate concepts identified by the expert. For instance, the ganglionar lattice for the diagnosis of 'cardiac insufficiency' is specified in section IV, part 2.

The structure of a ganglionar lattice operator, mGL, can be described by a {0,1} *Topology matrix, T[t_{ij}];* each row represents a subset of primary antecedents associated with a meaningful concept, and each column a primary antecedent, j = 1,,m. The entry t_{ij} = 1 if the i^{th} concept includes the j^{th} primary antecedent, and '0' otherwise. A sample of this matrix is shown in section IV, part 3.

Definition

Mathematically, the *Ganglionar Lattice* is the operator mGL($\underline{s}$):$\underline{s}$ ϵ I^n $\rightarrow$ s_0 ϵ I. The explicit expression of the mGL is specified by [nBR_k, T], where nBR_k is associated to the k^{th} node of the lattice of a particular consequent, and T is the lattice topology matrix.

Properties of an mGL

Some properties of nBRs are retained by any mGL. It is easy to show (i) if $nBR_k(\underline{0}) = 0$ and $nBR_k(\underline{1})$ = 1, for all operators of the lattice, then mGL($\underline{0}$) = 0 and mGL($\underline{1}$) = 1; (ii) if each nBR_k in the lattice is monotonously increasing, then the mGL operator is monotonously increasing; (iii) since all nBR operators in the lattice are differentiable then the operator mGL is differentiable.

An mGL is a composition of polynomial expressions thus resulting, possibly, in a more complex polynomial expression than that of an mBR. In fact, the degree of an mBR is of (s_1*...*s_m), while the degree of an mGL may be up to $(s_1*...*s_m)^{m-1}$; thus the mGL is far more complex than an mBR. Nevertheless, the mGL is specified by considerably less than the 2^m coefficients of an mBR. These facts suggest that *the mGL is most appropriate to model well established complex reasonings, with casted intermediate concepts and involving a large number of antecedents, which is often the case in medicine* [10, 11, 14].

It must be emphasized that the main advantage of an mGL is that it exposes all the intermediate variables corresponding to useful concepts, which at least in medicine, are of paramount importance to observe the behavior and eventually to explicate the outcome. These procedures will be most significant if achieved on solid, widely understood and accepted intermediate concepts, such as syndromes, clinical findings, pathological states, etc., since the primary antecedents alone are not so indicative.

The lattice is named *Ganglionar* in the same sense of a cluster of neurons processing multiple synaptic inputs; as performed by the lattice of nBRs. The GL is an *analogy* since its structure resembles the expert's mental scheme for reasoning. Moreover, it uses *analog* variables that can instantly and simultaneously (and continuously) produce the value for every consequent (if the inputs were continuous variables). The realization of this type of ES could be achieved on analog devices instead of a computer program.

The rooted tree structure

An extremely simple mGL is obtained when the structure is in the form of a rooted tree [25]. These structures can be defined on partitions of the primary set of antecendents, and successive refinements of the former partition, down to the singleton partition. An example is provided in section IV, part 2.

This special mGL, turns out to be an mBR composition of nBRs on disjoint variables, thus its degree is the same of an mBR, and behaves like one but with restricted coefficients. Furthermore, it can be proven that there always is an mBR that behaves exactly alike, and from which the mGL can be recuperated, provided T is given. In other words, *every mGL with rooted tree structure can be represented by an mBR.* Nevertheless the converse is not true. This mBR is a more powerful operator than the mGL, but on far more coefficients and not exposing the intermediate variables.

The coefficient identifiction of an mGL with rooted tree structure, is unique; in fact it can be proven that two mGLs with the same rooted tree structure and with the same input-output behavior must have the same coefficients. This fact is crucial for explicating the consequent value. In general an mGL without a rooted tree structure, cannot be expressed as an mBR, and the conditions under which uniqueness extends, are not known.

Identification of the ganglionar lattice

To successfully use an mGL operator it must be specified to fit a particular reasoning, i.e., the structure and the coefficients must be identified. The structure is completely identified by the expert when specifying the subsets associated with concepts of useful significance. The coefficient identification may be approached in practice by successive approximations, as is later described. Nevertheless, in the special case of a rooted tree structure, the coefficients can be exactly computed, given an appropriate set of at least 2^m-1 linearly independent 'training' cases.

Parameter identification by successive approximations

Consider an mGL of a specific consequent with its structure of intermediate concepts already identified by the expert. Take arbitrary initial values for the coefficients of all nBRs included in the mGL. Consider also an appropriate variety of training cases supplied by the expert, including values for all the primary antecedents and corresponding consequents. The algorithm employed to identify the coefficients corresponds to that used in System Theory to minimize the square error by an iterative procedure on the difference equation of the mGL, using the gradient method. Furthermore this method is similar to the back-propagation that is widely used in neural networks [16]. Its limitations in the non-linear case of rooted tree structure are well known [16, 26].

This procedure has been applied successfully to identify the mGL corresponding to 'cardiac insufficiency' using real outcomes certified by the expert, as shown in section IV, part 4. Also the procedure has performed well identifying simple monotonous mGLs without rooted tree structure although the solution may not be unique. Most important is that the method has performed with satisfaction identifying the ESs to evaluate the normality of ECG [11] and for the diagnosis, characterization and treatment of hypertensive patients [17].

III. Explication and prospection

Basic ideas

One of the main criticisms of ES focuses on their limited explicative and prospective capacity [27, 28]. Here, a new technology based on analog simulation of the expert's ability for reasoning is proposed, which turns out to be suitable to answer the following questions at different levels of abstraction: (i) Which retrospective reasons the ES can furnish to explicate the consequent's value?, and (ii) Which prospective reasons the ES can foresee to select the most appropriate successive evaluation to optimally improve the consequent quality?

Actually, the expert builds his answer to the above questions only on a few antecedents of different conceptual levels. He will choose his arguments regarding whom he is explicating to, on the following basis: (i) the arguments must be adapted in level to the user (specialist, student, etc.); (ii) the arguments must reduce the user's uncertainty up to satisfaction, and (iii) the arguments must be as simple as possible. These criteria essentially identify the explicative and prospective method developed here, for this ES based on fuzzy analog GLs.

The problem at hand reduces to classify the subsets of antecedents according to their explicative or prospective capacity. The basic idea to measure the explicative capacity (prospective capacity) of subsets of primary antecedents is by the uncertainty change that would have occurred (that would occur) in the previous (next) state of these antecedents.

Suppose that some primary antecedents of a monotonous increasing mGL are *evaluated*, i.e., each one has been measured or estimated and consequently assigned a punctual value. The rest of the antecedents remain *unevaluated*, i.e., for them a measured or estimated punctual value, is not known. A way to handle this unevaluated condition is by assigning to each unevaluated antecedent the range I, of equally possible values and then, the possible values of the consequent range in an *interval* of I. The upper limit of this interval is the *pessimistic* value, and the lower limit, the *optimistic* value. Since the mGL is assumed monotonous, these extreme values are obtained by assigning the value '1', and assigning the value '0', to all unevaluated antecedents at once, respectively, keeping all other antecedents on their respective punctual values.

Define the *UNCertainty of the consequent* of $mGL(s^*)$ at the operating state $\underline{s}^*$, denoted by $UNC(s_0 = mGL(\underline{s}^*))$, as the interval length of s_0. Thus, $UNC(s_0) = |\max mGL(\underline{s}^*) - \min mGL(\underline{s}^*)|$, where $\underline{s}^*$ represents all possible values of each component. For a monotonous mGL, $UNC(s_0) = |$ (pessimistic value)-(optimistic value) $|$.

It should be noticed that if the consequent is a fuzzy variable, the computation of the area under the possibility distribution is a measure of its uncertainty. This concept reduces the evaluation of UNC when real variables are used. If all the antecedents are unevaluated (evaluated pointwise) the consequent is totally uncertain (certain), and $UNC(s_0) = 1$ ($UNC(s_0) = 0$).

Explicative capacity

Consider only monotonous mGLs at an operating state specified by $\underline{s}^*$. *Define* the *EXPlicative capacity of a subset* S_i of primary antecedents of an $mGL(\underline{s}^*)$ at the operating state $\underline{s}^*$, denoted by $EXP(S_1)$, as the uncertainty decrease in the consequent, when a state $\underline{s}^-$ is forced retrospectively. This previous state is obtained by changing the values of all the antecedents in S_i to range I, while the rest remain as given. Thus $EXP(S_i) = UNC(mGL(\underline{s}^-)) - UNC(mGL(\underline{s}^*)$.

The modification on $\underline{s}^*$ to obtain $\underline{s}^-$ consists of the assumption that all the antecedents in S_i become retrospectively unevaluated, i.e. if $s_j \epsilon S_i$ then $s^-_j =$ I.

Prospective capacity

Consider only monotonous mGLs at an operating state specified by $\underline{s}^*$. *Define* the *PROspective capacity of a subset* S_i of primary antecedents of an $mGL(\underline{s}^*)$, at the operating state $\underline{s}^*$, denoted by $PRO(S_i)$, as the average uncertainty decrease in the consequent when all the unevaluated antecedents in S_i take all possible values, while the rest remain as given. Thus $PRO(S_i) = < UNC(mGL(\underline{s}^*)) - UNC(mGL(\underline{s}^+)) >$, where $<\cdot>$ is the average integral operator computed for all unevaluated antecedents in S_i, and where $\underline{s}^*$ are the given antecedent values. The values $\underline{s}^+$ contain, as integration variables, all the unevaluated antecedents in S_i. The unevaluated antecedents not in S_i, simultaneously take the value '1', to compute the pessimistic case, and the value '0', to compute the optimistic case.

It can be shown that for monotonous mGL with rooted tree structure, the computation of the average uncertainty decrease is equivalent to assigning the values 0.5 to all unevaluated antecedents in S_i. Thus, the formidable computation to obtain $PRO(S_i)$ can be avoided.

If S_i contains evaluated antecedents only, then $PRO(S_i) = 0$; clearly no uncertainty decrease can be achieved by S_i. The opposite occurs if S_j contains all antecedents and everyone is unevaluated, then $PRO(S_j) = 1$.

Meaningful partitions
To advantageously use $EXP(S_i)$ and $PRO(S_i)$ it is required an organization of the set of primary antecedents, $S = \{s_1, s_2,, s_m\}$, according to the mGL structure. Consider any partition of S and call it *meaningful* if all its members correspond to meaningful concepts in the mGL, including the primary antecedents as concepts; and denote it by P. Thus P can be viewed as a set of concepts of different levels of abstraction, and each concept is a subset of S.

Explication and prospection

The explication of the consequent in precise terms other than the consequent value itself, is fundamental for the acceptance of an ES. Any procedure to explicate assumes that there is a threshold, U, that must be surpassed to achieve explicative satisfaction. On the other hand, the cyclic procedure to evaluate a consequent needs an objective strategy to propose the subsequent evaluation or estimation to be performed, and with assurance of an optimal consequent's quality improvement. This prospective strategy assumes that there is a threshold, V, that must fail to end the cyclic approximation to the decision. Both procedures use the consequent uncertainty decrease as the basic criterion.

Basic explication and prospection
Consider the singleton partition of the set of primary antecedents and construct the power set of this partition. It must be regarded as a set of meaningful concepts: the primary antecedents.

Explication
Rank each member of the power set by EXP and select those subsets containing a minimal number of members (primary antecedents) and with EXP > U. The concepts included in any solution constitute an explication of no less than U-completeness of the consequent value. Note that this selection is not unique, in general, since more than one member of the power set may satisfy the above conditions.

Optionally, the chosen explication may be transformed into a fuzzy explicative rule in restricted natural language, if an intensity-to-words rule, such as the one given in Table 4, is provided. This procedure is exemplified for the diagnosis of cardiac insufficiency in section IV, part 6.

Prospection
Rank each member of the power set by PRO, and select the highest in the ranking of the group of 'p' primary antecedents. This last condition suggests to perform 'p' tests at once. Note that PRO of any set of evaluated antecedent is null.

This criterion by no means assures the test(s) outcome, it only constitutes the best 'a priori' choice, from the point of view of informatics, which objectively would optimally reduce the consequent's uncertainty. This strategy can be successively used until the threshold V is failed, presumably using the minimum number of tests to achieve satisfaction. An example of this procedure is shown in section IV part 5.

Higher level explication and prospection
These procedures are identical to the previous ones except that the meaningful partition, P, includes concepts of higher level of abstraction. There are several meaningful partitions on which these procedures can be applied. The lattice itself suggests how to identify them to suit the user. Each partition is of a different level of abstraction and simplicity, in relation to the concepts involved and with respect to the number of arguments or tests used.

The procedure for explicating is most valuable to fit the different user's needs: from a specialist to a student. An example is provided in section IV part 7.

The strategy of prospecting in terms of abstract concepts renders the system friendly to the user, because it substitutes several observations by one estimation; that of the chosen concept. The procedure leads to fail the threshold V easier and faster. The main drawback of this strategy is that higher order concepts cannot be properly measured, thus only estimated by expertise.

IV. Application to cardiac insufficiency

This example is taken from an Expert System to aid in the diagnosis and characterization and treatment of hypertensive patients [17]. The Cardiac Insufficiency is one of the parenchymatous damage caused by this disease [29].

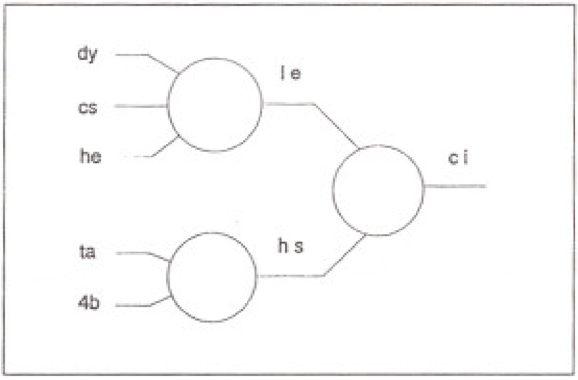

Fig. 2. The ganglionar lattice of cardiac insufficiency (ci), exposing the intermediate concepts of 'lung edema' (le) = {dyspnea (dy), lung congestive signs (cs), hemoptysis (he)}; and 'heart signs' (hs) = {tachycardia (ta), heart 4^{th} bruit (4b)}.

Part 1: the structure
Consider the medical diagnosis of Cardiac Insufficiency, 'ci', the top consequent for this diagnostic decision. According to the expert, the primary antecedents required to evaluate the 'ci' are [29]: 'dyspnea' = dy, 'lung congestive signs' = cs, 'hemoptysis' = he, 'tachycardia' = ta, and 'heart 4^{th} bruit' = 4b. Thus the set {dy, cs, he, ta, 4b} = ci. Among all possible subsets of 'ci', the cardiologist identifies the subsets dy, cs, he = 'lung edema' = le, and ta, 4b = 'heart signs' = hs. Thus the family of subsets is: < {dy};{cs};{he};{ta};{4b};{ta, 4b};{dy, cs, he};{dy, cs, he, ta, 4b} >. By the order relation of 'inclusion' the lattice of Fig. 2 is defined. The top of the lattice is the 'ci' and the atoms in the base are the primary antecedents.

Part 2: the ganglionar lattice
On the lattice of ci the following nBRs are defined: le = 3BR(dy, cs, he); hs = 2BR(ta, 4b); and ci = 2BR(le, hs). The composition defined by the structure is the 5GL(dy, cs, he, ta, 4b) = 2BR(3BR(dy, cs, he), 2BR(ta, 4b)).

In this case the structure is of the form of a rooted tree, so the variables partition in the same form as the structure branches. This fact induces a different perspective to view the rooted tree structure by partitions of the set of primary antecedents, as: the one-member partition is < {dy, cs, he, ta, 4b} = ci >; a two-member refinement is < {dy, cs, he} = le, {ta, 4b} = hs >; and finally the atom partition < {dy}, {cs}, {he}, {ta}, {4b} > is reached.

Table 1. Data of thirty cases.

#	dy	cs	he	ta	4b	ci
1	0	0	0	0	1	0.10
2	0	0	0	1	0	0.08
3	0	0	0	1	1	0.20
4	0	0	1	0	0	0.15
5	0	0	1	0	1	0.25
6	0	0	1	1	0	0.23
7	0	0	1	1	1	0.35
8	0	1	0	0	0	0.45
9	0	1	0	0	1	0.50
10	0	1	0	1	0	0.60
11	0	1	0	1	1	0.70
12	0	1	1	0	0	0.60
13	0	1	1	0	1	0.70
14	0	1	1	1	0	0.75
15	0	1	1	1	1	0.85
16	1	0	0	0	0	0.08
17	1	0	0	0	1	0.15
18	1	0	0	1	0	0.12
19	1	0	0	1	1	0.25
20	1	0	1	0	0	0.22
21	1	0	1	0	1	0.30
22	1	0	1	1	0	0.30
23	1	0	1	1	1	0.40
24	1	1	0	0	0	0.55
25	1	1	0	0	1	0.70
26	1	1	0	1	0	0.68
27	1	1	0	1	1	0.75
28	1	1	1	0	0	0.85
29	1	1	1	0	1	0.95
30	1	1	1	1	0	0.93

Part 3: the topology matrix
The rows corresponding to single primary antecedents are superfluous, thus omitted. Then the topology matrix of the 5GL is a 3 concepts by 5 primary antecedents, as

$$[T] = \begin{array}{r|ccccc|l} & dy & cs & he & ta & 4b & \\ ci & 1 & 1 & 1 & 1 & 1 & \text{fifth level} \\ le & 1 & 1 & 1 & 0 & 0 & \text{third level} \\ hs & 0 & 0 & 0 & 1 & 1 & \text{second level} \end{array}$$

Part 4: identification of coefficients
The following parameter values are obtained adjusting the 5GL by the successive approximation method, using data from 30 patients with true diagnosis certi-

Table 2. Successive computations of PRO(ci/{antecedents}).

	PRO_0	PRO_1	PRO_2	PRO_3
PRO(ci/dy)	0.131	0.139	0.154	×
PRO(ci/cs)	0.517	×	×	×
PRO(ci/he)	0.176	0.197	×	×
PRO(ci/ta)	0.096	0.093	0.092	0.092
PRO(ci/4b)	0.092	0.089	0.088	0.088
UNC(ci)	1.000	0.518	0.335	0.181
Assign value	cs = 0.9	he = 0.7	dy = 0.5	etc.

Table 3. Subsets with EXP > 0.8.

Case N°	Subsets S_i of primary antecedents					Cardinal	Explicative capacity
	dy	cs	he	ta	4b		
32	1	1	1	1	1	5	1.000
31	1	1	1	1	0	4	0.912
30	1	1	1	0	1	4	0.909
16	0	1	1	1	1	4	0.881
28	1	1	0	1	1	4	0.825
29	1	1	1	0	0	3	0.823
15	0	1	1	1	0	3	0.792

fied by the expert given in Table 1. The chosen data covers most of the possible cases.

Extending the notation, the coefficients of the respective nBRs are: {le_1 = 0.063; le_2 = 0.569; le_3 = 0.178; le_4 = 0.069; le_5 = 0.011; le_6 = 0.021; le_7 = 0.089}; {hs_1 = 0.504; hs_2 = 0.482; hs_3 = 0.014}, and {ci_1 = 0.826; ci_2 = 0.199; ci_3 = -0.025}. The adjustment quality, measured by the error between the model and the expert, is RMS = 2.5%.

To explore on the method's adjustment quality, a 5BR(dy, cs, he, ta, 4b) was identified with the same data. The resulting coefficients are: 0.000, 0.079, 0.449, 0.149, 0.079, 0.098, 0.022, - 0.008, - 0.038, - 0.028, 0.002, 0.072, - 0.048, 0.002, 0.002, 0.022, 0.154, 0.014, 0.125, 0.034, 0.005, 0.035, - 0.006, 0.045, 0.025, - 0.005, - 0.082, - 0.103, - 0.163, - 0.033, - 0.043, 0.144. On this 5BR realization, intermediate concepts are not be exhibited, thus reducing the model's explicative capacity. Nevertheless, the highly increased number of coefficients improves the adjustment quality fifty times to RMS = 0.05%.

Table 4. Restricted language-to-membership value translation.

Restricted language	Concept membership
Highest	1.0
Very high	[0.85–0.99)
High	[0.70–0.85)
Medium	(0.30–0.70)
Low	(0.15–0.30]
Very low	(0.01–0.15]
Lowest	0.0

Table 5. EXP for the set (le, hs).

Antecedent		Cardinal	EXPlicative capacity
le	hs		
1	1	2	1.000
1	0	1	0.823
0	1	1	0.181

Part 5: prospection for the next evaluation test

Suppose it is the first interview of a patient, i.e., all primary antecedents of ci are unevaluated. Thus dy = cs = he = ta = 4b = I, and the consequent ci = I. It is not known which would be the best next test to perform so that the uncertainty of ci is maximally reduced. A sequence of best tests cannot be absolutely specified since the second choice will depend on the outcome of the first. The best a priori choice at each j^{th} stage is evaluated in the column PRO_j, j = 0,1,2, of Table 2. Subsequent columns show the best next choice, once the present evaluation is known. The last row in Table 2 shows a fictitious outcome for the selected evaluation. The order of the following column depends on the operating point determined by the evaluation performed now. This choice reduces the consequent's uncertainty to the value indicated in the next to the last row (UNC(ci)). Notice that the evaluation of 'cs' followed by 'he' and 'dy', reduces the uncertainty of ci from total uncertainty, UNC(ci = [0.0; 1.0]) = 1.0, to UNC(ci = [0.614; 0.795]) = 0.181.

Part 6: explication on primary antecedents

Suppose that the patient is completely evaluated for ci in terms of the primary antecedents: dy = 0.5, cs = 0.9, he = 0.7, ta = 0.1, 4b = 0.2, thus computing the ci = 0.64. For example choose the threshold U = 0.8. An explication suited for a non specialist or student is in terms of primary antecedents only. This explication is detailed but long. The procedure to obtain it requires computing EXP of the 32 subsets of the primary set of antecedents. Table 3 shows the computed value of EXP for the 7 most significant groups. Only one of these is of smallest cardinality, 3, surpassing the threshold of 0.8; that is N° 29 of EXP({dy, cs, he}) = 0.82. The explicative rule, translated into the restricted natural language by Table 4, is the simplest satisfactory explication to 80% of the full explication, and reads: 'The Cardiac Insufficiency is of medium intensity (0.64) because the Dyspnea is of medium intensity (0.5), the Heart Congestive Signs is of very high intensity (0.9), and the Hemoptysis is of high intensity (0.7)'.

Part 7: explication on concepts of higher abstraction

The main advantage of an mGL is that it exposes the higher concepts, not only to understand the model, but to supply a more conceptual explication. For this purpose it is adequate to explicate ci = 0.64, in terms of the clinical findings of 'lung edema' = le, with resulting value of 0.74; and 'heart signs' = hs, with value of 0.15. The method proceeds evaluating EXP({le}) = 0.82, EXP({hs}) = 0.18, and EXP({le, hs}) = 1.00; as shown in Table 5. All values are computed at the operating point. Clearly, the simplest choice for 80% explication, translated to restricted language, reads: 'The Cardiac Insufficiency is of medium intensity (0.64) because the Lung Edema is of high intensity (0.74)'.

The possibility of producing these variety of explications is a paramount characteristic of this technology since it renders the ES friendly and acceptable to the user.

V. Discussion and conclusions

An Expert System based on fuzzy analog ganglionar lattices was presented. It is capable of simulating complex reasonings modeled by a functional composition of Basic Reasoning operators. This nBR operator allows different type of association among its antecedents, generalizing a combination of weighted average, triangular norm and co-norm operators (probabilistic sum and product, respectively). This lattice is capable of exhibiting the variables corresponding to intermediate reasonings, which is valuable in medical applications for explicating at different levels of abstraction, and for prospecting the best sequence of unevaluated antecedents. The lattice serves as knowledge base and inference engine in a unified manner. The procedures to explicate and prospect are founded on the uncertainty decrease in the consequent respect to antecedents of different levels of abstraction, thus constituting a friendly method to explicate and an objective criterion to approach a conclusion.

Each node of the lattice corresponds to a BR unit capable of processing 'n' antecedents of different relative importance. The association of nBRs to each node of the lattice permits to model the intermediate concepts with useful meaning for the expert. The application example shows that the structure of the ganglionar lattice (mGL) can represent the mental model of the expert when reasoning, and exhibit more or less detail according to the user's preference.

The number of parameters of the nBR operator, grows exponentially with the number of antecedents. Even though a large number of parameters results in good adjustment quality, the computational complexity for the consequent also increases. For example, a 7BR (n = 7) needs 448 products and 126 sums. Real time applications may be limited by the computational complexity as antecedents grow in number. Fortunately, it is unusual to find basic reasonings with more than seven active inputs [30]. Furthermore, in relation to the preceding comments, the analog model presented here uses analog variables, so that it could be realized on analog nBRs, thus taking no time for processing.

Only in the case of lattices in the form of a rooted tree the mGL is recuperated by an mBR; in this sense the mBR operator is more complete. Also in this case, the realization of the mGL is unique, so that the exhibited intermediate variables have a unique value, for a given input. The possibility of having multiple output connections in the lattice structure of an mGL, makes the complexity grow beyond that of its component nBRs, thus obtaining operators that are exponentially more powerful.

In the case of the mGL with rooted tree structure, the parameters can be identified by calculation, given a complete set of input-output data covering the space of parameters. The iterative procedure to identify the mGL works adequately even for cases without rooted tree structure, but the solutions may not be unique. For antecedents with moderate inconsistencies the parameters are identified within tolerable accuracy.

The ES on fuzzy analog ganglionar lattice, compared to technologies based on production rules, does not require to verify the consistency of the basic reasoning functions. This is assured by the procedure to define the lattice architecture and by the coefficients adjustment, based on data supplied by the expert.

The 'a priori' average decrease in the consequent's uncertainty, used as the criterion to measure the prospective capability of an unevaluated group of antecedents, allows to choose the subset of antecedents that protentially produces the maximum decrease in the consequent's uncertainty. The method to select the next test offers an objective criterion to rank the prospective capability of different members of a partition of primary antecedents in order to get as much information as possible with fewer tests. Analogously, the 'a posteriori' average decrease in the consequent's uncertainty, used as the criterion to measure the explicative capability of a group of antecedents, introduces a very human aspect in the generation of explications. Results of the identification process of the subset that produced the maximum decrease in the consequent's uncertainty, have been presented in the application example.

The explicative capacity of the model is useful for the medical user (physician, student) who wants to know why the consequent has reached a certain state. It also allows the user to communicate to the patient an explication about his condition. In both situations, it is possible to choose different levels of explication abstraction according to the user knowledge level.

In the generation of explications using intermediate concepts, it is required that the mGL is unique, otherwise the patient's condition could accept multiple explications in terms of intermediate concepts.

In relation to the iterative coefficient identification process, based on successive presentation of input patterns, information about the expert's reasoning is gradually gained. This procedure is an 'external instructor learning process' and the learning is reflected on the coefficients. If this procedure is used to identify the coefficients of an mGL made of a single mBR, as shown in section IV part 4, not only will the coefficients be identified but several alternative structures could be inferred by clustering antecedents through the relative coefficient magnitudes. Then the GL could be synthesized on several approximate ganglionar lattices, including each of the several nBRs. This procedure can serve to identify possible concepts (nBRs) in fields where these concepts are not yet defined. Besides these alternative realizations may provide an explication for the known fact that experts do not necessarily agree on a concept structure since, in general, there is no unique solution.

The model's mathematical formalization has been thought out as a conceptual basis for future extensions that will incorporate more realistic aspects of the human reasoning to the ES, such as the uncertainty treatment of both the expert's approximate reasoning and the antecedents' imprecision or ignorance.

Acknowledgments

Recognition is due to Professor Eduardo Rosselot, MD, Dean of the School of Medicine, University of Chile, for his enthusiastic contribution as the expert cardiologist. Also appreciation is expressed to Pablo Estevez, PhD, for his enlightening discussions, and to all members of the research group for their work and encouragement. This work has been partially funded by the Department of Electrical Engineering of the University of Chile and by FONDECYT project N° 1930-889 without which it would not have been possible.

References

1. Wyard PJ, Nightingale C. A single layer higher order neural net and its applications to context free grammar recognition. In: Linggard R, Meyers DJ, Nightingale C, eds. Neural Network for Vision, Speech and Natural Language. Chapman & Hall, 1992; 203–34.
2. Gallant SI. Neural network expert systems. In: Neural Network Learning and Expert Systems. Cambridge, Massachusetts: MIT Press, 1993; 253–94.
3. Sima J. Neural expert systems. Neural Networks 1995; 2 (8): 261–71.
4. Jones D. Neural networks for medical diagnosis. In: Maren A, Harston C, Pap R, eds. Handbook of Neural Computing Applications. Academic Press, 1990; 309–17.
5. Feigenbaum EA. Survey of expert systems. In: Firebaugh MW, ed. Artificial Intelligence: A Knowledge-based Approach. PWS-Kent, 1989; 333–68.
6. Hayes-Roth F, Waterman DA, Lenat DB. An overview of expert systems. In: Hayes-Roth F, Waterman DA, eds. Building Expert Systems. Addison-Wesley, 1983; 3–30.
7. Jackson P. Expert systems and artificial intelligence. In: Introduction to Expert Systems. Addison-Wesley, 1986; 1–18.
8. Frank Puppe. Characterization and history of expert systems. In: Systematic Introduction to Expert Systems: Knowledge Representations and Problem-solving Methods. Springer-Verlag, 1993; 3–8.
9. Pauker SG, Gorry GA, Kassires JP, Schwartz WB. Towards the simulation of clinical cognition: taking a present illness by computer. Am J Med 1976; 60: 981–96.
10. Holzmann CA, Perez CA, Rosselot E. A fuzzy model for medical diagnosis. Med Prog Tech 1988; 13: 171–8.

11. Holzmann CA, Hasseldieck U, Rosselot E, Estevez P, Andrade A, Acuña G. Interpretation module for normal ECG screening. Med Prog Tech 1990; 16: 163–71.
12. Holzmann CA, Avaria M. Entropy aided diagnosis in genetics. IEEE/Eng Med Biol Mag 1992; 11 (3): 35–40.
13. Holzmann CA, Ehijo A. A model for expert reasoning: fuzzy ganglionar lattices. IJCNN-93, Nagoya, Japan, 3, Oct 1993; 2991–2994.
14. Charniac E, McDermott. Bayesian inference networks. In: Introduction to Artificial Intelligence. Massachusetts: Addison-Wesley, 1987; 477–82.
15. Birkhoff G. Lattice theory. Amer Math Soc Colloq Publ; Providence 1967 (25), 3rd ed.
16. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. In: Rumelhart DE, McClelland, eds. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations. MIT Press, 1986; 318–62.
17. Holzmann CA, Estevez P, Mella P. Organic characterization of the hypertense: a connectionist approach. In: Cornelis, Peeters, eds. Proc North Sea Conf Biomed Eng IFMBE. Antwerp, 1990, Topic-8, N° 6.
18. Yager R. Approximate reasoning as a basis for rule-based expert systems. IEEE Trans Sys Man Cyber 1984; Vol SMC-14, 4: 636–43.
19. Zadeh LA. Biological applications of the theory of fuzzy sets and systems. In: Proctor LD, ed. Biocybernetics of the Central Nervous System. Boston, Massachusetts: Little Brown, 1969; 199–212.
20. Yager R. Fuzzy decision making including unequal objectives. Fuzzy Sets Syst 1978; 1: 87–95.
21. Yager R. On a general class of fuzzy connectives. Fuzzy Sets Syst 1980; 4: 235–42.
22. Dyckhoff W. Generalized means as model of compensative connectives. Fuzzy Sets Syst 1984; 14: 143–54.
23. Dubois D, Prade H. A review of fuzzy set aggregation connectives. Information Sciences 1985; 36: 85-121.
24. Weber S. A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. Fuzzy Sets Syst 1983; 11 (2): 115–34.
25. Harary F. Graph Theory. London: Addison-Wesley, 1969.
26. Kreytzig E. Advanced Engineering Mathematics. New York: Wiley, 1993.
27. Shortliffe E, Buchanan B, Feigenbaum E. Knowledge engineering for medical decision making: a review of computer-based clinical decision aids. Proc IEEE, Sept 1979; 67 (9): 1207–24.
28. Duda RO, Shortliffe EH. Expert systems research. Science 220, N° 4594, April 15, 1983: 261–268.
29. Hurst JW. The Heart, Arteries and Veins. MacGraw-Hill, Kogakusha Ltd, 1974: 416–65.
30. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956; 63: 81-97.