ISSN 0980-1472

Revue Internationale de systémique

SYSTÉMIQUE MATHÉMATIQUE

Vol. 3, N° 4, 1989

afcet Dunod

AFSCET

Revue Internationale de systémique

Revue Internationale de Sytémique

volume 03, numéro 4, pages 421 - 436, 1989

Evidential reasoning on artificial neuron network

Su-Shing Chen

[23] J. D. Lowrance, T. D. Garvey and T. M. Strat, A framework for evidental-reasoning systems, in *Proc. national conference on artificial intelligence*, pp. 896-903, A.A.A.I., Menlo Park, California, 1986.

[24] R. Moore, *Reasoning about knowledge and action*, Technical Note 408, S.R.I., International, Menlo Park, California, 1980.

[25] J. Pearl, Fusion, propagation, and structuring in belief networks, *Artif. Intell.*, 29, 1986, pp. 241-288.

[26] J. Pearl, On logic and probability, *Computational intelligence*, 4, 1988, pp. 99-103.

[27] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann, San Mateo, California, 1988.

[28] R. Reiter, A logic for default reasoning, *Artif. Intell.*, 13, 1980, pp. 81-132.

[29] N. Rescher, *Many-Valued Logic*, McGraw-Hill, New York, 1969.

[30] E. H. Ruspini, Epistemic logic, probability, and the calculus of evidence, in *Proc. Tenth Intern. Joint Conf. on Artif. Intell.*, Milan, Italy, 1987.

[31] E. H. Ruspini, *The logical foundations of evidential reasoning*, Technical Note 408, Artificial Intelligence Center, S.R.I. International, Menlo Park, California, 1987.

[32] E. H. Ruspini, *On the semantics of fuzzy logic*, Technical Note No. 475, S.R.I. International, Artificial Intelligence Center, Menlo Park, California, December 1989.

[33] L. J. Savage, *The Foundations of Statistics*, Dover, New York, second revised edition, 1972.

[34] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey, 1976.

[35] P. Smets, Belief functions, in P. Smets, A. Mamdani, D. Dubois and H. Prade Eds., *Non-standard Logics for Automated Reasoning*, Academic Press, New York, 1988.

[36] C. A. B. Smith, Consistency in statistical inference and decision, *J. Roy. Stat. Soc. Ser. B*, 23, 1961, pp. 1-37.

[37] P. Suppes, The measurement of belief, *J. Roy. Stat. Soc. Ser. B.*, 36, 1974, pp. 160-175.

[38] E. Trillas and L. Valverde, On mode and implication in approximate reasoning, in M. M. Gupta, A. Kandel, W. Bandler and, J. B. Kiszka Eds., *Approximate Reasoning and Expert Systems*, Amsterdam, North Holland, 1985, pp. 157-166.

[39] L. A. Zadeh, A theory of approximate reasoning, in D. Michie and L. I. Mikulich Eds., *Machine Intelligence*, 9, New York: Halstead Press, 1979, pp. 149-194.

[40] L. A. Zadeh, A computational approach to fuzzy quantifiers in natural language, *Comput. Math.*, 9, 1983, pp. 149-184.

[41] L. A. Zadeh, Fuzzy sets, *Inform. Control.* 8, 1965, pp. 338-353.

[42] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *I.E.E.E. Trans. Systems, Man and Cybernetics*, SMC-3, 1973, pp. 28-44.

[43] L. A. Zadeh, Syllogistic reasoning in fuzzy logic and its application to usuality and reasoning with dispositions, *I.E.E.E. Trans. Systems, Man and Cybernetics*, SMC-15, 1985, pp. 754-765.

REVUE INTERNATIONALE DE SYSTEMIQUE
Vol. 3, N° 4, 1989, pp. 421 à 436

EVIDENTIAL REASONING ON ARTIFICIAL NEURAL NETWORKS

Su-Shing CHEN

University of North Carolina [1]

Abstract

In this information age, computers play an essential role in almost every field of our daily life. There is a great deal of uncertainty and incompleteness in the information and knowledge of real-world problem domains. Evidential reasoning is an area of machine intelligence which addresses this issue. Since human beings have the cognitive capability of dealing with uncertainty and incompleteness, one may ask how to model human intelligence in this aspect. A relaxation process is suggested for evidential reasoning on artificial neural networks. It is hoped that further investigation may show this system to be a viable model for biological neural networks as well.

Résumé

Dans cette ère de l'information, les ordinateurs jouent un rôle essentiel dans presque tous les domaines de la vie quotidienne. Il existe une importante part d'incertitude et de nombreuses informations incomplètes dans les connaissances concernant les problèmes du monde réel. Le raisonnement basé sur la théorie de l'évidence est un aspect de l'«intelligence» de la machine qui répond à ces préoccupations. Puisque les êtres humains ont la capacité cognitive de traiter incertitudes et imperfections, on peut se demander comment modéliser l'intelligence humaine à cet égard. Un processus de relaxation est suggéré pour le raisonnement basé sur la théorie de l'évidence dans les réseaux de neurones artificiels. On espère que des investigations futures permettront de montrer que ce système est également un modèle viable pour les réseaux de neurones biologiques.

1. Department of computer Science, Charlotte, 28223 North Carolina, U.S.A.

1. Introduction

The reasoning process in human intelligence is adaptive, incremental and evolutive. On the contrary, in the information processing literature, the reasoning process usually employs logical inference and symbolic processing. This approach is not readily amenable to the characteristics of human intelligence.

In this paper, the scope of reasoning is broadened. In addition to the usual logical and symbolic methodologies, we use cognitive maps (weighted graphs of relations) and iterative relaxation techniques on artificial neural networks. Artificial neural networks are computational models that are similar to how human brain works. This idea frees us from the limitations of conventional AI techniques and enables us to extend the usual symbolic reasoning results to more adaptive and evolutive systems. While symbolic processing of our system still relies on the usual AI techniques, computationally intensive tasks (e. g. constraint satisfaction and evidential reasoning) are carried out by artificial neural networks. Thus, we favor a hybrid system of symbolic and neural processing modules which are complementary to each other.

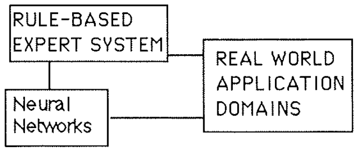

Figure 1. *A hybrid AI and neural network system.*

Human intelligence has the capability of deriving definitive and conclusive knowledge from uncertain and incomplete information. From the information processing point of view, a quite successful theory of evidential reasoning has emerged from several different approaches – fuzzy reasoning [19], probabilistic inference [15], and frames of discernment [18]. In this paper, evidential reasoning is performed as an iterative process [16], rather than a single collection of arithmetic and logical operations for computing certain probabilistic or fuzzy measures ([15], [18], [19]). Moreover, the correct prescription of initial probabilistic and fuzzy measures are not necessary, because the system is self-correcting.

A knowledge base is dynamically encoded in cognitive maps represented by artificial neural networks. The link weights represent the relationships among concept nodes or their attribute nodes. The node values of the neural networks are updated from some initial input node values representing certainty factors of evidences and observations. The iterative process reaches some limiting output node values representing certainty factors of hypotheses and diagnoses. This approach is adaptive and flexible, because not only the node values (certainty factors) are updated but also the link weights are modified (learned from external world) during the computation. It seems to be a rational way of managing uncertain knowledge. In principle, we should make the least commitment to the information and knowledge structure in evidential reasoning before we can draw any conclsuion. As we know more and more about the situation, the structure can gradually be hardened. Neural computing provides this kind of "plasticity". In other approaches to evidential reasoning, a lot of structures are assumed in order to get to the uncertain information.

2. Artificial neural networks

There are many artificial neural network models ([9], [10], [17]). The essential characteristics are analog (real-valued) and parallel computation unlike the digital, sequential nature of ordinary computers. Basically, an artificial neural network model should contain the following components:

(1) A set of processing units.

(2) A state of activation.

(3) An output function for each unit.

(4) A pattern of connectivity among units.

(5) A propagation rule for propagating patterns of activities throughout the network.

(6) An activation rule for combining inputs to a unit with the current state of that unit to produce a new level of activation for that unit.

(7) A learning rule to modify patterns of connectivity by experience.

(8) An environment within which the system must operate.

The processing units have only simple computational functions. There should be a significant number of these units to carry out computations in a collective way. Artificial neural networks can be considered as very fine grained parallel processors. These units may be used to represent conceptual objects such as features, letters, words and concepts. Since an artificial neural network consists of a large number of processing units, a group of units may be used to represent conceptual objects distributively. In fact, neurophysiologists have hypothesized that information in our brains are distributedly represented [17]. In this case, units may represent feature entities which form feature patterns or vectors. There are input units, output units and hidden

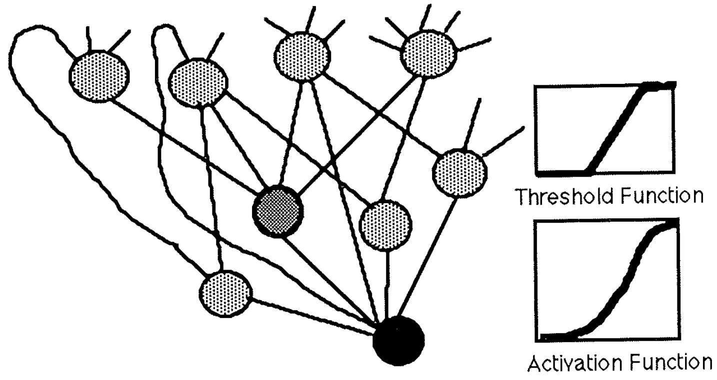

Figure 2. An artificial neural network and its activation and threshold functions.

units in an artificial neural network. The hidden units are the system units that encode a wide body of knowledge.

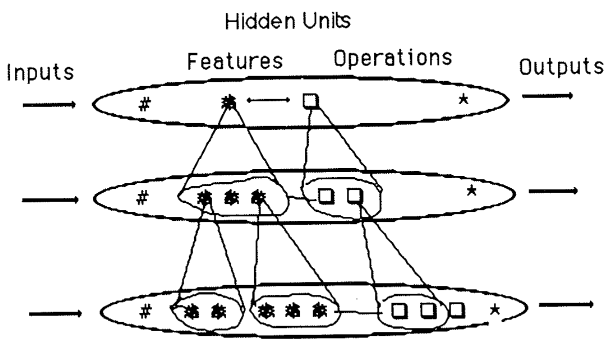

Figure 3. Input, hidden and output units.

The state of a system of n units at time t is given by the vector

$$A(t) = (a_1(t), a_2(t), \ldots, a_n(t)),$$

where each $a_i(t)$ is the activation value of unit u_i at t. The processing of the system is determined by the dynamics of $A(t)$:

$$dA(t)/dt = F(A(t), \Phi(W(t), O(t))),$$

where $\Phi(W(t), O(t))$ is the network input, depending on the connectivity matrix $W(t)$ and the output vector $O(t)$ at t. In ordinary models, the network input to unit u_i is $\Sigma_j w_{ij}(t) o_j(t)$. The function F is a nonlinear function, such as a sigmoid function.

The output function $f_i(a_i(t)) = o_i(t)$ may be a threshold function or a stochastic function which produces the output. The output vector $O(t)$ of the system is denoted by $(o_1(t), o_2(t), \ldots, o_n(t))$.

The pattern of connectivity describes the topology of an artificial neural network by a sparse matrix $W(t) = (w_{ij}(t))$. It is called the matrix of link weights which represent the strengths of connections among units. A positive (negative or zero) weight represents excitatory (inhibitory or no) connection respectively. The propagation in the network is assumed to be asynchronous, that is, local computations are performed simultaneously. For simplicity, computations are assumed to be instantaneous.

Modification (learning) of patterns of connectivity consists of the following three kinds: (1) development of new connections, (2) loss of existing connections, and (3) modification of strengths of existing connections. The simplest learning rule was observed by Hebb in 1949: "If a unit u_i receives an input from another unit u_j, and if both are highly active, the weight w_{ij} from u_j to u_i should be strengthened". The change Δw_{ij} is given by $\xi a_i o_j$ (ξ is the learning rate) in the delta rule. In general, Δw_{ij} is a nonlinear function $g(a_i(t), t_i(t), o_j(t), w_{ij}(t))$, where $t_i(t)$ is the teaching input from the external world. In the simple delta rule, there is no teaching input given.

The representation of an external world and input/output pairs depends on the specific problem domain. For instance, a Markov random field model may be suitable to visual perception. In knowledge acquisition and evidential reasoning, we represent knowledge in a collection of cognitive maps.

3. Cognitive maps

Cognitive maps are (uni- or bi-) directed graphs representing relations (by arcs) among concepts/attributes (by vertices) [1]. Cognitive maps include several knowledge representation schemes. Semantic networks or frames form a special class of cognitive maps. Inference networks and causal networks form other classes of cognitive maps. Other cognitive maps include spatiotemporal, terminological, and evidential cognitive maps. Spatiotemporal cognitive maps represent knowledge at different locations and times. Terminological cognitive maps represent different terminologies and their equivalences. Evidential cognitive maps represent conflicting evidences and their combination, different viewpoints and their relative weights, uncertain and partial information and knowledge.

In cognitive maps, link weights may be assigned to relations representing their degrees, and node values may be assigned to concepts and attributes representing certainty factors or belief measures. They have the structure of a fuzzy graph. The positive relational weights represent excitatory associations, and the negative relational weights represent inhibitory associations. There is a natural mapping from cognitive maps to artificial neural networks.

Moreover, cognitive maps as graphs can be extended to allow both inputs and outputs from the external world. First, input and output nodes may be added to a cognitive map as follows:

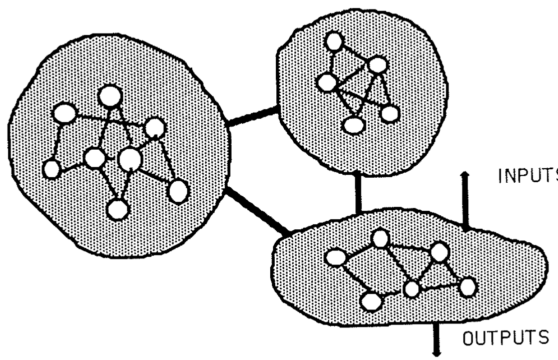

Figure 4. Cognitive maps with inputs and outputs.

This is obvious when we know exactly which are the input and output nodes. Alternatively, a cognitive map receives inputs by clamping (or instantiating) a certain number of nodes, and it generates outputs by sending values of certain nodes to the external world. The only difference between the two is that neural network updating dynamics is applied to all nodes in the second case, but not to the input and output nodes in the first case.

A cognitive map has either a single representation or a distributed representation. A production may be represented distributedly by its feature entities and a relationship between two propositions by multiple links among feature entities.

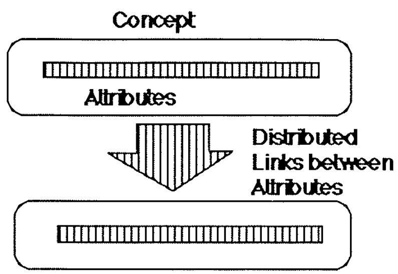

Figure 5. Distributed representation.

4. Evidential reasoning

Artificial neural networks are useful to evidential reasoning. We shall formulate some evidential reasoning problems in cognitive maps and artificial neural networks. There are several interesting advantages of this approach. We shall compare our evidential reasoning scheme with Bayesian networks [15] and probabilistic logic ([13], [14]).

What is the significant differences between our system and other evidential reasoning systems? In other systems, certainty factor and belief measure can only be updated by new evidences. In our system, certainty factor or belief measure may also be improved upon through a relaxation scheme of constraint satisfaction without any new evidence introduced. Thus, there are two components: (1) reasoning based on additional external evidences, and (2) reasoning with internal existing constraints among propositions in the knowledge base.

Bayesian networks are directed acyclic graphs with nodes

$$\{x_1, \ldots, x_n\}$$

representing propositions and links representing direct relationships of propositions whose weights are conditional probabilities of these propositions. The joint probability distribution $p(x_1, \ldots, x_n)$ can be computed by conditional probabilities, such as

$$p(x_n \mid x_{n-1} \ldots x_1)\, p(x_{n-1} \mid x_{n-2} \ldots x_1) \ldots p(x_3 \mid x_2 x_1)\, p(x_2 \mid x_1)\, p(x_1).$$

Conversely, conditional and marginal probabilities are computed from joint probabilities. Bayesian inference is performed by computing the marginal probabilities of hypotheses (propositions) $\{H_1, \ldots, H_s\}$, given the probabilities of evidence $\{e_1, \ldots, e_m\}$. Interpretation of input data is to instantiate a set of proposition variables and to compute the probabilities of hypotheses variables. That is, to find the most probable instantiation of all hypotheses variables, given the observed data.

Although probability axioms provide a rigorous calculus for evidential reasoning, their relevance to human intelligence as well as real-world applications is disputable. If human intelligence is modeled by neural networks, then conditional and marginal probabilities can not be computed in a natural way from joint probabilities. However, joint probabilities are computed by the conditional probabilities as link weights and the marginal probabilities as node values of a particularly structured neural network of joint probabilities.

Nevertheless, artificial neural networks have the capability to manage uncertainty in an intrinsic manner. Evidential reasoning is represented by evidential cognitive maps which are mapped onto artificial neural networks. The computational models of artificial neural networks are used for the incremental computation in evidential reasoning.

As input nodes, evidences are the instantiated (clamped) proposition nodes of an artificial neural network. As output nodes, each proposition in the collection of hypotheses is assigned a certainty factor. The probability axioms are not required. Evidences and their certainty factors are propagated through the networks so that, in the equilibrium state, hypotheses are assigned some limiting belief values. Similarly, an artificial neural network supplies a mechanism to revise belief values of a collection of hypotheses to reach the most satisfactory explanation of evidences.

The probabilistic logic [14] is a semantic generalization of the first order logic, in which the truth values of propositions are probability values between 0 and 1. It applies to a (finite) knowledge base of consistent propositions. This probabilistic inference reduces to the ordinary logical inference when the probabilities of all propositions in the collection are either 0 or 1. In [2], the probabilistic logic was extended to evidential logic in the framework of Dempster-Shafer theory [18]. In [3], we formulated the probabilistic logic as a probabilistic consistent labeling problem ([11], [12], [16]).

For a knowledge base of propositions $\{x_1, \ldots, x_n\}$, the space of all possible worlds is given by the collection of all consistent binary valuation vectors (the cardinality k is $\leqq 2^n$). The main idea of probabilistic logic is to assign a joint probability distribution $(p_1, \ldots, p_k)$ to this space, called a

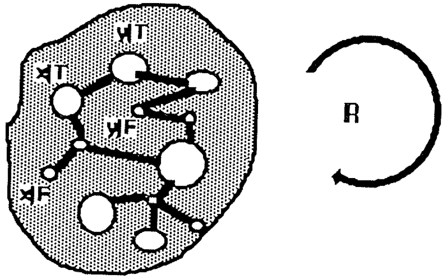

Figure 6. Relaxation of probabilistic labeling.

permissible probabilistic interpretation vector. The generalized truth value (or probability) of x_i is the sum of probabilities of all the possible worlds that satisfy x_i. Probabilistic inference is performed by a linear analysis of convex sets in vector spaces or by propagation on the binary semantic networks by conditional probabilities. The probabilistic logic (as a probabilistic labeling problem) is realized as evidential cognitive maps in the next section.

5. A probabilistic reasoning system

Let S be a collection of propositions $\{x_1, \ldots, x_n\}$, and let Λ be a collection of labels $\{\lambda_1, \ldots, \lambda_m\}$ with any mathematical structure. The labeling problem is to find a consistent labeling of propositions in S by Λ, given a set of relations among propositions and a set of constraints among propositions and their labels. In classical and probabilistic logics, the collection of labels is only $\{0, 1\}$.

Since our model allows multiple semantic values, it is useful in various kinds of evidential reasoning. For instance, conflicting points of view may be represented by different labels. The conflict is resolved by a relaxation scheme to reach a consistent labeling. The learning capability of artificial neural networks permits us to learn about the resolution of conflicts.

For each x_i, let Λ_i be the subset of Λ that is compatible with x_i. For any pair $\{x_i, x_j\}$ of propositions (i, j distinct), let Λ_{ij} be the subset of compatible pairs of labels in $\Lambda_i \times \Lambda_j$. A labeling $L = \{L_1, \ldots, L_n\}$ is an assignment of a set of labels Λ_j in Λ to each x_i. L is consistent if for each i, j and all λ in Λ_i, $(\{\lambda\} \times \Lambda_j)$ intersects with Λ_{ij}. L is unambiguous if it is consistent and assigns only a single label to each proposition.

The probabilistic labeling is to assign a probability distribution $p_i(\lambda)$ to the statement "λ is the correct label of x_z". In the case of probabilistic logic, it is the generalized truth value of x_i (there are only two labels-true and false). An arbitrary labeling of a knowledge base may not be consistent and unambiguous, because the constraint satisfaction is required among either propositions in the knowledge base or a combination of new input evidences with the knowledge base.

The interaction with external worlds (human-machine interface) is through a symbolic (or AI-based) reasoning subsystem. For instance, query is still better performed in a logic-based fashion. At the initial stage, the probability distributions $p_i(\lambda)$ is either estimated by the user or is provided by the symbolic reasoning scheme. Now the initial probability distributions (evidences) go through a constraint satisfaction checking by the artificial neural network subsystem. This is a relaxation scheme which iterates the process until the convergence to final probability distributions is reached. The final probability distributions are sent back to the symbolic reasoning subsystem for either interaction with external worlds or further symbolic reasoning. Thus, evidential reasoning is carried out by the artificial neural network subsystem.

The relaxation scheme is described as follows. An initial assignment of probability distributions $\{p_i^{(0)}(\lambda)\}$ to $\{x_i\}$ is given. A relaxation operator **R** is defined to transform one set $\{p_i^{(k)}(\lambda)\}$ of probability distributions to another set $\{p_i^{(k+1)}(\lambda)\}$. The limit $\{p_i^{(\infty)}(\lambda)\}$ of $\{p_i^{(k)}(\lambda)\}$ gives the unambiguous labeling under compatibility constraints, as k approaches to infinity. In reality, we expect the limit to be attained after a finite number of iterations.

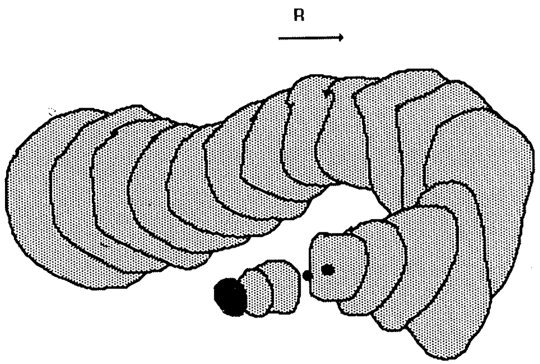

Figure 7. The relaxation operator R making certainty measures converging to A limit.

In practice, the limit $\{p_i^{(\infty)}(\lambda)\}$ may not be unique (we are not getting an unambiguous labeling). The multiple labelings are sent back to the symbolic reasoning subsystem so that its knowledge base can select an appropriate result for further reasoning.

There are several ways to define the relaxation operator **R**. A relaxation operator **R** should produce a $p_i^{(k+1)}(\lambda)$ from the combination of $p_i^{(k)}(\lambda)$ and $s_i^{(k)}(\lambda)$ by some update equations, where $s_i^{(k)}(\lambda) = \sum_j d_{ij} \sum_{\lambda'} r_{ij}(\lambda, \lambda')\, p_j^{(k)}(\lambda')$, $\sum_j d_{ij} = 1$, $r_{ij}(\lambda, \lambda')$ $(-1 \leqq r_{ij}(\lambda, \lambda') \leqq 1)$ is the compatibility function of "label λ is assigned to x_i and label λ' is assigned to x_j", and j-indices are indices of source nodes leading to the i-th node. In particular, $r_{ij}(\lambda, \lambda')$ can be the conditional probability $p_{ij}(\lambda | \lambda')$ of "x_i has label λ and x_j has label λ'". A requirement of the update equations is that $p_i^{(k+1)}(\lambda)$ should remain to be a probability distribution.

A relaxation operator **R** is defined by the following update equations ([3], [4]):

$$p_i^{(k+1)}(\lambda) = \min[1, \max(0, p_i^{(k)}(\lambda) + s_i^{(k)}(\lambda))],$$

$$s_i^{(k)}(\lambda) = \sum_{\lambda', j} (p_{ij}^{(k)}(\lambda, \lambda') + \Delta p_{ij}^{(k)}(\lambda, \lambda'))\, p_j^{(k)}(\lambda'),$$

$$\Delta p_{ij}^{(k+1)}(\lambda, \lambda') = a_{ij} \Delta p_{ij}^{(k)}(\lambda, \lambda') + b_{ij}\, p_i^{(k+1)}(\lambda)\, p_j^{(k)}(\lambda'),$$

where a_{ij} and b_{ij} are learning parameters. The first equation makes sure that $p_i^{(k+1)}(\lambda)$ stays between 0 and 1. The second equation provides the network input to the (i, λ)-th node. The third equation includes the Hebbian learning rule as a special case.

6. A fuzzy reasoning system

The probabilistic reasoning system may be extended to a fuzzy reasoning system which relies also on an artificial neural network subsystem. Let S be a knowledge base of propositions $\{x_1, \ldots, x_n\}$. Each x_i assumes one of labels $\Lambda_i = \{\lambda_{1(i)}, \ldots, \lambda_{m(i)}\}$ (the label set is no longer assumed to be a fixed set). For each i, we define a fuzzy subset A_i of Λ_i which is a mapping from Λ_i to [0, 1] representing the degree of compatibility of $\lambda(i)$ in Λ_i with x_i. For each pair $\{x_i, x_j\}$ (i, j distinct), we define a fuzzy subset A_{ij} of $\Lambda_i \times \Lambda_j$ which represents the degree of compatibility of $\lambda(i)$ to x_i with $\lambda'(j)$ to x_j. Similarly, we may define fuzzy subsets A_{ijk} and etc. for triples and other tuples to represent compatibilities of higher-order networks (sigma-pi networks [17]).

Similarly, the interaction with external worlds (human-machine interface) is through a symbolic (or AI-based) reasoning subsystem. At the initial stage, the fuzzy value $A_i(\lambda)$ is either estimated by the user or is provided by the symbolic reasoning scheme. Now the initial fuzzy values (evidences) go through a constraint satisfaction checking by the artificial neural network subsystem. This is a relaxation scheme which iterates the process until the convergence to final fuzzy values is reached. The final fuzzy values are sent back to the symbolic reasoning subsystem for either interaction with external worlds or further symbolic reasoning. Thus, evidential reasoning is carried out by the artificial neural network subsystem.

The relaxation scheme is described as follows. An initial assignment of fuzzy values $\{A_i^{(0)}(\lambda)\}$ to $\{x_i\}$ is given. A relaxation operator **R** is defined to transform one set $\{A_i^{(k)}(\lambda)\}$ of fuzzy values to another set $\{A_i^{(k+1)}(\lambda)\}$. The limit $\{A_i^{(\infty)}(\lambda)\}$ of $\{A_i^{(k)}(\lambda)\}$ gives the crisp (either 0 or 1) labeling under compatibility constraints, as k approaches to infinity. In reality, we expect the limit to be attained after a finite number of iterations. The update equations of the relaxation operator **R** is defined as follows.

$$A_i^{(k+1)}(\lambda) = \min[1, \max(0, A_i^{(k)}(\lambda) + s_i^{(k)}(\lambda))],$$

$$s_i^{(k)}(\lambda) = \sum_{\lambda', j} (A_{ij}(\lambda, \lambda') + \Delta A_{ij}^{(k)}(\lambda, \lambda'))\, A_j^{(k)}(\lambda'),$$

$$\Delta A_{ij}^{(k)}(\lambda, \lambda') = a_{ij}\, \Delta A_{ij}^{(k)}(\lambda, \lambda') + b_{ij}\, A_i^{(k+1)}(\lambda)\, A_j^{(k)}(\lambda'),$$

where s_i is the support function, A_{ij} and ΔA_{ij} are weight and its increment respectively, and j-indices are the indices of source nodes leading to the i-th node. a_{ij} and b_{ij} may be adjusted for convergence results.

In practice, the limit $\{A_i^{(\infty)}(\lambda)\}$ may not be unique (we are not getting a crisp labeling). The multiple labelings are sent back to the symbolic reasoning subsystem so that its knowledge base can select an appropriate result for further reasoning.

7. A dynamic knowledge acquisition system DYKAS

DYKAS, a knowledge acquisition system that receives inputs from and generates outputs to an external world and forms new pieces of knowledge from existing knowledge and input information, is being developed. The system is described as follows:

This system has deductive and inductive capabilities. From input data, new concepts and propositions are formed through inductive clustering algorithms. However, the consistency of newly formed propositions with the existing

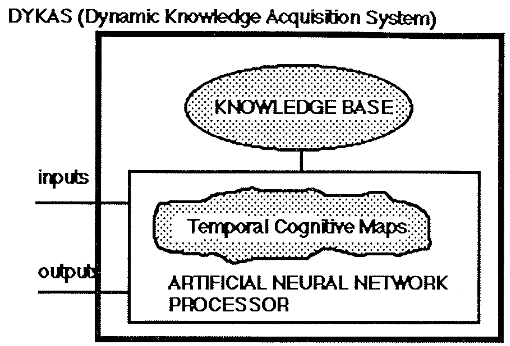

Figure 8. DYKAS: A knowledge acquisition system using artificial neural networks.

knowledge base is a question that can not be addressed by traditional AI learning algorithms. The artificial neural network subsystem supports the constraint satisfaction requirement.

The knowledge base of the system is encoded in a dynamic collection (over time t) of cognitive maps. These cognitive maps are temporal cognitive maps. Let $\{x_1(t), \ldots, x_n(t), \ldots\}$ be a collection of propositions over time t, each of which may assume a finite set $\Lambda_i(t)$ of labels (or semantic values). The relations among these propositions are represented by arcs in temporal cognitive maps. Under this dynamic assumption, label sets $\Lambda_i(t)$'s vary according to time t. This implies that semantic values of each proposition vary in time. For instance, a proposition, such as "it is raining", may change from true to false alternately for different hours in one day. Also, the intensity value of a pixel in the image frame of a video camera varies along the time axis.

For each t, we have a mapping (or a fuzzy set) $A_i(t)$ from $\Lambda_i(t)$ to the real numbers $\mathbb{R}$. It gives the certainty factors of labels in $\Lambda_i(t)$ for the proposition $x_i(t)$. A state of the knowledge base of temporal cognitive maps is

$$A(t) = (A_1(t). \ldots, A_n(t), \ldots),$$

which evolves with time t. Temporal relations in cognitive maps are given by directed arcs $\{x_i(t), x_j(t)\}$ which are assigned link weights. Over time t, existing relations may disappear, and new relations may emerge. Similarly

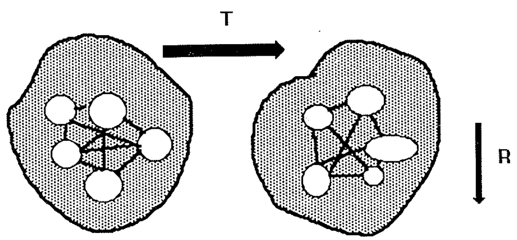

Figure 9. Relaxation operator R and temporal operator T form two degrees of freedom in DYKAS.

nodes x's may disappear or be idle, and some idle nodes may become active again.

The relaxation scheme is described as follows. For each fixed time step t, the relaxation operator **R** refines the knowledge structure as in the last section. An initial assignment of fuzzy values $\{A_i^{(0)}(t)(\lambda)\}$ to $\{x_i(t)\}$ is given. The relaxation operator **R** transforms one set $\{A_i^{(k)}(t)(\lambda)\}$ of fuzzy values to another set $\{A_i^{(k+1)}(t)(\lambda)\}$. The limit $\{A_i^{(\infty)}(t)(\lambda)\}$ of $\{A_i^{(k)}(t)(\lambda)\}$ gives a crisp (either 0 or 1) labeling under compatibility constraints, as k approaches to infinity. A set of compatibility coefficients $\{A_{ij}^{(k)}(t)\}$ which are mappings from $\Lambda_i(t)\times\Lambda_j(t)$ to R gives the degrees $\{A_{ij}^{(k)}(t)(\lambda,\lambda')\}$ of compatibility of label $\lambda'(t)$ to $x_j(t)$ with label $\lambda(t)$ to $x_i(t)$. The support $s_i^{(k)}(t)(\lambda)$ is a mapping from $\Lambda_i(t)$ to $\mathbb{R}$ which gives the support of new external evidence or internal state update for label $\lambda(t)$ at proposition $x_i(t)$.

The support $s_i^{(k)}(t)(\lambda)$ at the i-th node is given by

$$s_i^{(k)}(t)(\lambda)=\sum_j\sum_{\lambda'}A_{ij}^{(k)}(t)(\lambda,\lambda')\,A_j^{(k)}(t)(\lambda'),$$

which is a linear sum of compatibility coefficients of interconnecting proposition-labels weighted by the present certainty values for those labels.

The update equation of the states is defined by

$$A_i^{(k+1)}(t)(\lambda)=\min[1,\max(0,A_i^{(k)}(t)(\lambda)+s_i^{(k)}(t)(\lambda))].$$

The learning mechanism is provided by an update equation (as fuzzy sets):

$$A_{ij}^{(k+1)}(t)=f(A_{ij}^{(k)}(t),A_i^{(k+1)}(t),A_j^{(k)}(t)),$$

where f is a nonlinear functional or a polynomial mapping. The learning rule is more general than those in the probabilistic and fuzzy reasoning models so that more general knowledge acquisition tasks are supported.

The knowledge acquisition system has two operators-the relaxation operator **R** and the temporal operator **T**. In this paper, we shall not discuss **T** in detail. After each time step under **T**, the system generates to the artificial neural network subsystem an initial state $A_i^{(0)}(t)$. Iteratively, new evidence is obtained externally from input data or internally using update equations on the present state with compatibility constraints yielding new states. The limiting results of the relaxation scheme are passed on to the knowledge acquisition system as existing knowledge items at time t. Then, the system moves on to the next time step of reasoning. Overall, the certainty or belief measure at each node is highly dynamic. Moreover, nodes and links are dynamically created and detected by the two operators **R** and **T**. In a subsequent paper, we shall present a formal theory of temporal cognitive maps. Here, we can only describe briefly the issues involved.

8. Conclusion

In our evidential reasoning system, not only new evidences will update the certainty or belief measures of a collection of propositions, but also the constraint satisfaction among those propositions in cognitive maps will revise the measures. This mechanism is somewhat similar to human reasoning which is an evolutive process converging to the most satisfactory result. Under this model, inference and causal networks are represented by cognitive maps. Thus, we have generalized Bayesian networks and probabilistic logic to a more flexible framework which is realized on artificial neural network models. Different evidential reasoning algorithms are replaced by a single neural computing scheme. Finally, we discussed a general knowledge acquisition system-DYKAS, which uses the relaxation scheme for refining the internal knowledge structure of temporal cognitive maps.

References

[1] R. Axelrod, *Structure of Decision*, Princeton University Press, 1976.

[2] S. Chen, Some extensions of probabilistic logic, *Proc.* AAAI Workshop on Uncertainty in Artificial Intelligence, Philadelphia, Penn., August 8-10, 1986, pp. 43-48; An extended version appeared in *Uncertainty in Artificial Intelligence*, Vol. 2, L. N. Kanal and J. F. Lemmer Eds., North-Holland, pp. 221-227.

[3] S. CHEN, Automated reasoning on neural networks: A probabilistic approach, *I.E.E.E. First International Conference on Neural Networks*. San Diego CA, June 21-24, 1987, pp. II-313-318.

[4] S. CHEN, Knowledge acquisition on neural networks, *Unvertainty and Intelligent Systems*, B. BOUCHON, L. SAITTA and R. R. YAGER Eds, *Lecture Notes in Computer Science*, Springer-Verlag, Vol. 313, 1988, pp. 281-289.

[5] C. A. CRUZ and H. J. MYERS, Associative Networks, *I.B.M. Palo Alto Scientific Center Report ZZ*20-6459, October 1982.

[6] C. A. CRUZ and H. J. MYERS, Associative Networks II, *I.B.M. Palo Alto Scientific Center Report G*320-3446, 1983.

[7] A. P. DEMPSTER, A generalization of Bayesian inference, *J. R. Stat. Soc.*, Series B, 30, 1968, pp. 205-247.

[8] J. GORDON and E. H. SHORTLIFFE, A method for managing evidential reasoning in a hierarchical hypothesis space, *Artif. Intell.*, 26, 1985, pp. 323-357.

[9] S. GROSSBERG, *The Adaptive Brain I: Cognition, Learning, Reinforcement, and Rhythm, and The Adaptive Brain II: Vision, Speech, Language, and Motor Control*, Elsevier/North-Holland, 1986.

[10] J. J. HOPFIELD and D. W. TANK, *Computing with neural circuits: A model, Science*, Vol. 233, 1986, pp. 625-633.

[11] R. A. HUMMEL and S. W. ZUCKER, On the foundations of relaxation labeling processes, *I.E.E.E. Trans. PAMI*, 5, 1983, pp. 267-287.

[12] M. S. LANDY and R. A. HUMMEL, *A brief survey of knowledge aggregation methods*, New York University. Courant Institute Technical, Note 177, September 1985.

[13] J. LUKASIEWICZ, Logical foundations of probability theory. *Jan Lukasiewicz Selected Works*, L. BERKOWSKI Ed., North-Holland, 1970, pp. 16-43.

[14] N. J. NILSSON, Probabilistic logic, *Artif. Intell.*, 28, 1986, pp. 71-87.

[15] J. PEARL, *Probabilistic reasoning in intelligent systems: Networks of plausible inference*, Morgan Kaufmann, 1988.

[16] A ROSENFELD, R. A. HUMMEL and S. W. ZUCKER, Scene labeling by relaxation operations, *I.E.E.E. Trans. Systems. Man and Cybernetics*, 6, 1976, pp. 420-433.

[17] D. E. RUMELHART and J. L. MCCLELLAND, *Parallel distributed processing*, Vol. I and II, *M.I.T. Press*, 1986.

[18] G. A. SHAFER, *Mathematical Theory of Evidence*, Princeton University Press, 1979.

[19] L. A. ZADEH, Fuzzy logic and approximating reasoning, *Synthese*, 30, 1975, pp. 402-428.

REVUE INTERNATIONALE DE SYSTEMIQUE
Vol. 3, N° 4, 1989, pp. 437 à 458

LA NOTION DE SINGULARITÉ ET SES APPLICATIONS

C. P. BRUTER
Université Paris-XII [1]

Résumé
Singularité et bifurcation sont des notions essentielles non seulement en mathématiques, mais également dans la vie quotidienne, dans l'organisation des mondes physique et biologique. Les définitions de ces notions, leurs principales propriétés sont illustrées dans ce texte par des exemples très variés. Si la singularité est le lieu géométrique ou physique associé à la genèse ou à la disparition, la bifurcation décrit les modalités des transformations qui adviennent en cette singularité.

Abstract
Singularity and bifurcation are essential notions, not only in mathematics, but also in the daily life, and in the organization of the physical and biological universes. The definitions of these notions, their main properties are illustrated in this text through various examples. If the singularity is the geometrical or physical place associated with genesis or vanishing, the bifurcation describes the modalities of transformations which happen in the singularity.

A celui qui serait tenté de se pencher sur cet article, l'auteur ne saurait trop recommander la lecture préalable et complémentaire de la fresque philosophique constituant la première partie de *Topologie et Perception* [2]. On y traite d'objets, appelés systèmes dans cette revue. On en dégage des propriétés générales qui fournissent le soubassement à maintes explications, à une compréhension renouvelée de nombreux faits de sociétés. On fait appel dans cette étude à deux concepts, celui d'extrémalité d'une part, celui de centre

1. Mathématiques, U.F.R. Sciences, avenue du Général-de-Gaulle, 94010 Créteil Cedex.

Revue internationale de systémique. 0980-1472 Vol. 3/89/04/ 437 /22/$ 4.20/© Gauthier-Villars