WCCI 2010 IEEE World Congress on Computational Intelligence
July, 18-23, 2010 - CCIB, Barcelona, Spain FUZZ-IEEE

On the Semantics and the Use of Fuzzy Cognitive Maps in Social Sciences

Joao Paulo Carvalho

***Abstract*—Fuzzy Cognitive Maps (FCM) have been around for more than twenty years, but the way how they have been used and the interpretation of their results are nowadays far from their original intended goal. This paper focus on discussing the structure, the semantics and the possible use of FCM as tools to model and simulate complex social, economic and political systems, while clarifying some issues that have been recurrent in published FCM papers.**

I. Introduction

DECISION makers, whether they are social scientists, politicians or economists, usually face serious difficulties when approaching significant, real-world dynamic social systems. Such systems are composed of a number of dynamic qualitative concepts interrelated in complex ways, usually including feedback links that propagate influences in complicated chains and make reaching conclusions by simple analysis an utterly impossible task. Axelrod pioneering work on Cognitive Maps (CMs) [1] in the 70's introduced a graphic way to express real-world qualitative dynamic social systems from the viewpoint of such decision makers. CMs consisted on ordered graphs representing concepts (the entities that are relevant for the system in question) and the relations between those concepts. For several years CM analysis was simply structural and consisted in methods to extract information based on the way the concepts were interconnected [14][15]. Even though CM were meant to represent dynamic systems, complete, efficient and practical mechanisms to analyze and predict the evolution of data in CMs were not available for years. Obvious early candidates were the use of mathematical differential equations and of System Dynamics tools like those developed by J.W.Forrester [5][27]. However, due to their qualitative nature, numerical data in social CMs is usually uncertain and hard to get, making formulation of a mathematical model difficult, costly or even impossible. Efforts to introduce knowledge on these systems should have relied on natural language arguments in the absence of formal models. This led to the introduction of Fuzzy Cognitive Maps (FCM) by Kosko [10] in the mid 80's.

FCM were developed as a qualitative alternative approach to system dynamics and presented as more adequate to model the social systems usually approached by CM [11].

Since then dozens of works have applied FCM or FCM variants in the most diverse fields, even if, as one will discuss below, FCM have not always been the most adequate tool to model and simulate the dynamics of CM, namely because, as it will be shown in this paper, FCM are quantitative (not qualitative) Causal Maps (a subset of Cognitive Maps) that only allow modeling of basic symmetric and monotonic causal relations.

Although FCM seem to have become quite successful, some issues and late trends should not be ignored:

- FCM scientific development is increasingly moving FCM away from social sciences;
- Some FCM works are starting to ignore some of the most basic FCM assumptions;
- Unfortunately a large amount of FCM works seems to be implemented without considering FCM inherent limitations and reach conclusions ignoring the semantics of FCM.

In this paper one will address the above mentioned issues and discuss the adequacy of FCM as tools to model and simulate social, economic and political systems as originally intended by Kosko.

II. Fuzzy Cognitive Maps

Fuzzy Cognitive Maps (FCM), as introduced by Kosko [10][11][12][13], are meant to be a combination of Neural Networks and Fuzzy Logic that allow us to predict the change of the concepts represented in Causal Maps. The graphical illustration of FCM is a signed directed graph with feedback, consisting of nodes and weighted interconnections. Nodes of the graph stand for the concepts that are used to describe the behavior of the system and they are connected by signed and weighted arcs representing the causal relationships that exist among concepts. Figure 1 represents a fuzzy causal map based on a work presented by Mohr [18].

Each concept represents the actors, entities and social, political, economic or abstract concepts that compose our system. Examples of concepts might be Inflation, the actions of an influent Politic, a Revolution, the Wealth of an individual or a nation, the Welfare of population, Road conditions, etc. Each concept is characterized by a value usually ranging from [0..1] or [–1..1] representing a transformation from its real world value.

Manuscript received February 7, 2010. This work was supported in part by the FCT - Portuguese Foundation for Science and Technology under project PTDC/EEA-ELC/66259/2006.

J. P. Carvalho is with the Technical University of Lisbon, Instituto Superior Técnico, and INESC-ID, Rua Alves Redol, 9, 1000-029 Lisboa, PORTUGAL (phone +351213100238; e-mail: joao.carvalho@ inesc-id.pt).

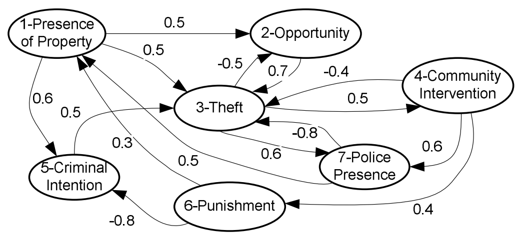

Fig. 1. Crime and Punishment causal map.

The concepts are connected by directed arcs that represent the causal relation between those concepts and are associated to a weight with range [-1..1]. A positive weight represents a causal increase and a negative weight represents a causal decrease (opposite effect). The way how one interpret the meaning (the semantics) of "causal increase" and "causal decrease" is one of the issues in this paper and will be discussed in section IV. For example, in Figure 1, a weight value of –0.8 (causal decrease) could be interpreted as the effect of a large decrease in Theft due to a Police_Presence increase (or the Theft increase due to the Police_Presence decrease). A smaller weight would represent a smaller effect.

To obtain the value of a concept, the value of each of its inputs (concepts) [0..1] or [-1..1] is multiplied by the respective weight [-1..1]; then the results are added and passed by a non-linear function used, among others, to limit the range of possible output values (Fig. 2).

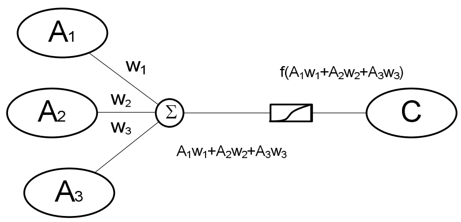

Fig. 2. Concept inference scheme on a FCM.

The dynamics of a FCM, which basically consist in its evolution in time, are modeled on an iterative way: time is discrete and the current value of each concept is computed based on the values of its inputs in the previous iteration. The update of the values of each concept for the current iteration must occur only after all concepts have been calculated. As the FCM evolves through time, the map might reach equilibrium, converging to a single state or a finite cycle of states. It is important to notice that "time" should be considered essential when modeling a FCM, since the rate of change on a social system (or in fact in most real world systems) cannot be infinite; i.e., when simulating a FCM, one cannot assume that the value of a concept can change from its minimum to its maximum value on a single iteration unless this iteration represents a large enough amount of time (see section IV.). However it is rare to find a FCM model where this concern is expressed, especially when considering social system models.

When seen as a whole, FCM allows the answer to what-if questions in causal maps: what happens to a system if some of its concepts change, or if new concepts are introduced or removed. Since the effect of each concept antecedent is <u>independent</u> from the effect of other concept antecedents, it is possible to remove or add concepts and links without having to change the rest of the system. This versatility in what concerns testing new scenarios represents one of the best qualities of FCM.

The graphical representation of the behavioral model of the system, as seen in Figure 1, allows a clear understanding of which concept of the system influences other concepts and to which degree, even if the overall system behavior is very hard to comprehend due to the feedback links.

It is also possible to use alternative representations to model FCMs, each with its own particular advantages. A square matrix representation, as exemplified in Figure 3., is a very compact representation containing all the relevant info to represent the FCM. Each position in the matrix contains the weight representing the causal relation between the concept indicated in the respective row and the concept indicated in the respective column. For example, w_{23}=0.7 represents the weight between concept 2 and concept 3. A '0' represents the absence of causal connection between the corresponding concepts. Although more compact, one must note that in this representation it is more difficult to understand the system and it is very difficult to verify the existence of feedback links.

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
C_1	0	0.5	0.5	0	0.6	0	0
C_2	0	0	0.7	0	0	0	0
C_3	0	-0.5	0	0.5	0	0	0.6
C_4	0	0	-0.4	0	0	0.4	0.6
C_5	0	0	0.3	0	0	0	0
C_6	0.3	0	0	0	-0.8	0	0
C_7	0.5	0	-0.8	0	0	0	0

Fig. 3. Crime and Punishment causal map represented as a weight matrix.

Usually it is accepted that causality is not self reflexive, i.e., a concept cannot cause itself [9][22], which means that the weight matrix always has '0' in its diagonal. However, as one will see below, sometimes authors claim that causality is not self-reflexive but inadvertently end up using self-reflex causality in their FCM.

The final representation of a FCM is its mathematical formulation. The most frequent is the one originally proposed by Kosko:

$$\forall \mathrm{j} \in \{1,..,\mathrm{n}\}, \mathrm{C_j}(t+1) = f\left(\sum_{\substack{i=1 \\ i\neq j}}^{n} \mathrm{w_{ij}} \times \mathrm{C_i}(t)\right), \quad (1)$$

where n is the number of concepts, $C_j(t)$ is the value of concept j in instant t, w_{ij} is the weight between concepts i and j, and finally f is a nonlinear function that limits the range of possible values for a concept. The most common nonlinear functions are step functions for bivalent or trivalent concepts, and the sigmoid (or logistic function) in the more interesting case of continuous concepts:

$$f(x) = \frac{1}{1+e^{-kx}}\ , \qquad (2)$$

where k controls how fast is the saturation of a concept.

Note that the restriction $i \neq j$ is only used when self causation is assumed as impossible.

An alternative approach has been introduced in [25], claiming to "introduce some kind of memory effect" to concepts in FCM and has become increasingly popular, although, as one will discuss in section IV., it completely alters the semantics of the original FCM:

$$\forall \mathrm{j} \in \{1, .., \mathrm{n}\}, \mathrm{C_j}(t+1) = f\left(\sum_{\substack{i=1 \\ i\neq j}}^{n} \mathrm{w_{ij}} \times \mathrm{C_i}(t) + \mathrm{C_j}(t)\right) (3)$$

If one considers a matrix of concepts $\mathrm{C} \in \mathbb{R}^{1\times n}$, and a matrix of weights $\mathrm{W} \in \mathbb{R}^{n\times n}$, then FCM inference can be expressed as the following product of the two matrixes:

$$\mathrm{C^{t+1}} = f(\mathrm{C^t} \times \mathrm{W}) \quad (4)$$

$$[C_1 \quad \dots \quad C_n]^{t+1} = f\left([C_1 \quad \dots \quad C_n]^t \times \begin{bmatrix} 0 & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & 0 \end{bmatrix}\right)$$

As one can see, FCM are not true "Fuzzy Systems", since they can be defined by a couple of matrixes and can be inferred using iterative standard algebraic operations. The simple fact that a system consists of variables defined with a continuous value ranging from 0 to 1 instead of Boolean values should not be enough to call it "Fuzzy". In fact they cannot even be truly called qualitative systems, since the only qualitative reasoning involved is (eventually) a direct mapping from a qualitative term to a number, e.g., "Large = 0.75" and uncertainty is not addressed in any way.

III. Fuzzy Cognitive Maps and Neural Networks

As it was shown in the previous section, FCM are not true fuzzy systems.

FCM are indeed a class of artificial Neural Networks (NN), where a concept is an artificial neuron. An artificial neuron is a mathematical function conceived as a crude model, or abstraction of biological neurons. Artificial neurons are the constitutive units in an artificial neural network. The artificial neuron receives one or more inputs (representing the one or more dendrites) and sums them to produce an output (synapse). Usually the sums of each node are weighted, and the sum is passed through a non-linear function known as an activation function or transfer function. The transfer functions usually have a sigmoid shape, but they may also take the form of other non-linear functions, piecewise linear functions, or step functions. When the threshold function is linear one has the original Rosenblat perceptron [21]. Usually a constant named bias is also considered as an extra term in the sum.

It is obvious that the description of an artificial neuron is exactly the same of what is presented in (1) and Fig.2, when considering bias = 0. Therefore, a concept on a FCM is an artificial neuron.

What makes a FCM distinct from a feedforward neural network is the existence of feedback cycles that imply a temporal dependence in the FCM. Therefore FCM are specifically a class of discrete time recurrent irregular neural networks. Kosko himself clearly states the feedback issue in [13], page 152, where he defines FCM as "fuzzy signed direct graphs with feedback". FCM without feedback are simply discrete feedforward NN with bias = 0.

Therefore, anyone using acyclic FCM, is using in fact NN and often is attempting something that has been attempted before using NN; by simply using the term "FCM" instead of "NN" they can claim to obtain original results where there are none. It is very important to state this fact once for all since recently there have been attempts of proposing new extensions for FCM that only apply when cycles are not present, and new applications of FCM that have been previously attempted with NN have even been accepted and published [6].

IV. Semantics of Fuzzy Cognitive Maps

In the previous sections FCM were thoroughly described and some issues were raised to be discussed.

A. *FCM Semantics and Causality*

FCM are about representing causality in dynamic systems. However they represent a very different approach from what is called the "Logic approach to causality" which basically consists in defining "Necessary Causes" and "Sufficient Causes"[22], and leads often to what is known as the "Causation Vs. Correlation" problem: the belief that correlation proves causation, is a logical fallacy by which two events that occur together are claimed to have a cause-and-effect relationship. This fallacy is also known as *cum hoc ergo propter hoc* ("with this, therefore because of this"). The FCM approach relies on the universally agreed facts that to test for causality it is necessary to guarantee that a cause must precede an effect and that an alteration on a cause alters the effect: there is a causal relation between two given concepts whenever a relative variation in one of those concepts cause a relative variation on the other one. For example, there is a negative causal relation between Police Presence and Theft: a major increase in Police Presence will probably cause a large decrease in Theft. Causal relations in causal maps should always involve change: the result of a causal effect is always a variation in one or more concepts.

Therefore relations and concepts on a FCM must semantically represent variations and when one models or interprets a FCM, it is necessary to have in mind the meaning of a concept and of the relations:

- Relations should represent causal increases or causal decreases;
- Concepts should somehow integrate the effects of the causal changes modeled by the relations.

However, is that what one thinks when looking at a FCM as the one depicted in Figure 1? Is that what one models when using the FCM mechanism described in (1), (2), (4)? It possible to answer these questions by analyzing the FCM inference process:

Let us consider only the concepts 7 - Police Presence and 3 - Theft. If Police Presence is High, i.e, equal to 1, then by applying (1), since w_{73} = -0.8, the result of Theft, depending on the sigmoid function used, should be around 0 if concepts range from [0..1] or -1 if concepts range from [-1..1]. How is this to be interpreted?

a) A high presence leads to very low theft on the next iteration - “If Police Presence is High then Theft will be Very Low”;
b) A high increase in police presence leads to a very high decrease of theft on the next iteration - “If Police Presence Increase Much then Theft will Decrease Very Much”;

Choice a) represents the usual approach when modeling and interpreting FCMs. In each instant each concept is defined by its absolute value (e.g., “Theft is presently Low”), and concepts often range from [0..1]. Although this interpretation is valid, it goes against the semantics of the above presented causality definition, where one is supposed to represent that a change on a concept causes a change in the consequent. One must notice that the meaning of *w* is supposed to be “how much a change in A causes an increase or decrease in C”, but in this case the meaning is obviously different; it should state “how much the value of A affects the value of C”. In this case, when someone refers to the weight in the relationship it is false to say “positive *w* represent causal increase and negative *w* represent causal decrease” (as Kosko states in its FCM definition [13]). The only possible true statement is “positive w cause higher values, and negative w cause lower values”, which is different from increase or decrease. One can argue that in the end there is a change in both antecedent and consequent (by stating Police presence decreased and is now 0, one obtains Theft is now 1, which means it increased), but semantically we have different meanings, since what one is really modeling is that an absolute value on the antecedent causes (imposes) a given absolute value on the consequent.

This approach leads to an implementation of causality more similar to the “Suficient cause” logical approach (x implies y), although there is an accumulation of effects in the consequent when there is more than one antecedent.

Another disadvantage of this approach regards the drastic dynamic changes it implies: it is perfectly possible that in consecutive instants a concept changes from its maximum possible value to its minimum or vice-versa. In real world social systems this can only be accepted if the time span between iterations is large enough to allow such change. However this time span between iterations is hardly ever mentioned in FCM approaches to social systems.

On the other hand, choice b) is the obvious semantic representation of the above definition of causality and the obvious representation for the original meaning of w. However it means that the state of the FCM in any system never indicates the absolute values of a concept. Instead what one has in any instant is the rate of change of a concept, i.e., how much it is currently changing. One will never know if theft is currently Low or High, one simply knows that it is currently decreasing or increasing.

Obviously both approaches have shortcomings. The ideal solution would be to be able to combine both solutions and have both the absolute value and variation of a concept on a given iteration (which is the common sense when representing the dynamics of a given instance).

This leads us to (3), the solution proposed by Stylios [25]. Notice that the main difference between this implementation of FCM, and the original from Kosko (1), is that the previous value of a concept is always added to itself in the present iteration. This has a few interesting implications, since the proposed solution is <u>equivalent to a Kosko FCM where each concept has a self-feedback link with *w*=1</u>:

$$\sum_{\substack{i=1 \\ i\neq j}}^{n} \mathrm{w_{ij}} \times \mathrm{C_i}(t) + \mathrm{C_j}(t) \equiv \sum_{\substack{i=1 \\ i\neq j}}^{n} \mathrm{w_{ij}} \times \mathrm{C_i}(t) + 1 \times \mathrm{C_j}(t)$$

$$\equiv \sum_{i=1}^{n} \mathrm{w_{ij}} \times \mathrm{C_i}(t), \mathrm{w_{jj}} = 1$$

When considering the matrix representation, the diagonal of the weight matrix consists of ‘1’ instead of ‘0’ in this approach. Our example FCM would become:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
C_1	1	0.5	0.5	0	0.6	0	0
C_2	0	1	0.7	0	0	0	0
C_3	0	-0.5	1	0.5	0	0	0.6
C_4	0	0	-0.4	1	0	0.4	0.6
C_5	0	0	0.3	0	1	0	0
C_6	0.3	0	0	0	-0.8	1	0
C_7	0.5	0	-0.8	0	0	0	1

Fig. 4. Crime and Punishment causal map represented as a weight matrix, when the FCM uses Stylios inference (3).

An implication of this approach is that one must take care when defining self-causality, since here all concepts cause themselves.

Another implication lies in the fact that the semantics of this system are completely different from the semantics of the original approach. In this case one ends up with an approach that lies roughly between those described in a) and b): the concepts are represented by their absolute values, but the relations implement semantics where the consequent is represented as a variation. Taking the previous example, one will have:

"If Police Presence is High then Theft will Decrease Very Much".

This is how the effect of accumulating the effects of the antecedents with a concept previous value must be interpreted. Although this presents an interesting approach towards implementing causality, extra care must be taken when interpreting and modeling such systems, since one is using an absolute value as an antecedent to model a variation in the consequent. It is not possible in this case to think in terms of "an increase in A will cause a large decrease in C" when defining the relationships.

Even if there are three rather different ways of interpreting FCM, it is very rare for an author to indicate how he/she modeled the system and what semantics he used. Note that not even Stylios considered the semantic implications of its novel proposed approach, simply referring to his proposal as a meaning to give the concepts some memory and accelerate saturation. Obviously one cannot expect to obtain valid results and conclusions if these semantic issues are not addressed.

B. Neurons as Concepts

As we have seen, the main goal of a FCM should be the dynamic simulation of qualitative social systems, and the role of the concepts would be to model the main entities and actors in such systems. However, one must not forget that on a FCM a concept is an artificial neuron (section III.). Thus, when using FCM to model social systems, one is reducing what might be complex entities (like a politician, social unrest, etc.), to a single neuron, and modeling the relationships among the several entities using a perceptron based mechanism [21]. In 1962 Novikoff [19] proved that the perceptron algorithm converges after a finite number of iterations if and only if the data set is linearly separable. A consequence of this is that if the data set is not separable, then the perceptron cannot model it. Therefore it is not possible to model systems where the several antecedents of a given concept are not linearly separable. In 1969 Marvin Minsky and Seymour Papert in their book Perceptrons [17], showed that it was impossible for a single layer perceptron to learn an XOR function. For example given A, B and C, if C is supposed to be an exclusive or (XOR) of A and B, then a FCM isn't capable of modeling that relationship. Obviously when considering individual concepts and their antecedents, FCM can't be used as universal aproximators. Therefore, when using FCM to model relationships between concepts (especially when trying to learn FCM automatically), one should be aware of these limitations, and this certainly show that it should not be conceivable to use a FCM to model social systems. Note that the problem here is that on FCM each node is a concept that represents a relevant and complex actor or entity.

The concepts and the relationships between concepts on social FCM are expected to be complex. But as we have seen, when using perceptrons complexity can only be achieved when using multiple perceptron layers, which is never the case on a FCM. To model such complexity it would be needed to separate the notion of concept and node on a FCM, and model each concept as a NN (or a FCM) of its own. Note that this is basically what Kosko and Dickerson did after proposing the FCM to model a herd of dolphins [11].

C. The Semantics of Monotonic and Symmetric Causality

Using the above described neural structure to model concepts and relations restricts FCM to the representation of systems that use only simple monotonic and symmetric causal relations between concepts. The symmetric restriction is due to the symmetric characteristics of the non-linear function. This limitation seriously hinders FCM capabilities to model social systems because many real world causal relations exhibit non symmetric or non monotonic semantics. Positive and negative causality often produce different non-symmetric results as we can see in [2]. For example, the amount of increase on a concept "Anti Pollution Measures" due to an increase in concept "Industrial Pollution" could be larger than the decrease in "Anti Pollution Measures" caused by a similar decrease in Industrial-Pollution (the measures could even not diminish at all). One could argue that the use of a non-symmetrical non-linear function would allow FCM to model these relations, however, this only holds if all the antecedents of a given concept behave similarly in what concerns the non-linearity, which basically makes this solution useless.

D. The Semantics of Time in FCM

One very important issue that was ignored (or avoided) when approaching the original FCM is time. Considering that all relations involved in FCM are causal relations, and that the effect of a causal relation always involves a change (represented by an amount of variation in a given period of time), one might make considerations about the validity and utility of the models when time is never mentioned.

Time is obviously essential in the study of system dynamics (SD). When we are dealing with quantitative SD like [5] or [27], timing issues are naturally solved, since time is explicitly expressed in the mathematical equations used to describe the relations between concepts. However, in FCM time is not an explicit entity and must be somehow included in the model.

The only way of including time in the original FCM is to establish a base time for the iterations (an iteration should correspond to a fixed amount of time – 1 day, 1 month, etc.), and consider the base time implicitly when modeling the system. This implicit time can be included in the semantics of the FCM when defining the weights and by considering the maximum possible change in the value of the concepts during that period of time. Without imposing this implicit timing knowledge it is impossible to guarantee a minimally acceptable simulation. A similar approach was proposed in [4], where details can be found.

It is also worth to mention that several FCM extensions that include time as an explicit variable have been more or less successfully proposed over the last years [7][16][20] [26][28].

V. CONCLUSION

FCM have evolved, and so have their meaning and intended use. There should be no doubt that FCM have interesting properties and can be used in multiple applications, especially when combined with the most recent learning mechanisms. As an example one can cite some excellent works [8][23][24]. However, as it was shown in this paper, their mechanisms and semantics are far from ideal to be used when modeling “socio-political-economical” systems, and the name FCM is definitely not the most appropriate since FCM are not Fuzzy and not adequate to model the dynamics of qualitative social Cognitive Maps as originally introduced by Axelrod. For that purpose, different and better mechanisms are needed, and some have been proposed before [2][3][4]. On the other hand one must note that the term Cognitive Map (not as coined by Axelrod) fits quite nicely the structure and aspect of a FCM: several artificial neurons irregularly connected as a directed graph containing cycles. Maybe a different term should start to be used instead of “FCM”, like for example “Dynamic Neural Maps” or “Dynamic Concept Maps”.

REFERENCES

[1] Axelrod, R., *The Structure of Decision: Cognitive Maps of Political Elites*, Princeton University Press, 1976

[2] Carvalho, J.P. and Tomé J.A.B., *Rule Based Fuzzy Cognitive Maps – Fuzzy Causal Relations*, Computational Intelligence for Modelling, Control and Automation: Evolutionary Computation & Fuzzy Logic for Intelligent Control, Knowledge Acquisition & Information Retrieval, edited by M. Mohammadian, IOS Press, 1999

[3] Carvalho, J.P., Tomé,J.A., *Fuzzy Mechanisms for Qualitative Causal Relations,* Views on Fuzzy Sets and Systems from Different Perspectives. Philosophy and Logic, Criticisms and Applications. (Studies in Fuzziness and Soft Computing), Chapter 19, Edited by. Rudolf Seising, Springer, 2009

[4] Carvalho, J.P., Tomé,J.A.,*Rule Based Fuzzy Cognitive Maps - Expressing Time in Qualitative System Dynamics*, Proceedings of the 2001 FUZZ-IEEE, Melbourne, Australia 2001

[5] Forrester, J.W., several papers available online at http://sysdyn.mit.edu/sd-intro/home.html

[6] Gonzalez, J., Castillo, O., Aguilar, L., *Performance Analysis of Cognitive Map-Fuzzy Logic Controller Model for Adaptive Control Application*, Proceedings of the 2008 IEEE World Congress on computational Intelligence, WCCI2008, Hong Kong, 2008, pp. 2375-2381

[7] Hagiwara, M., *Extended Fuzzy Cognitive Maps*, in proceedings of IEEE International Conference on Fuzzy Systems, 1992, pp. 795–801.

[8] Hengjie, S., Chunyan, M., Zhiqi, S., *Implementation of Fuzzy Cognitive Maps Using Fuzzy Neural Network and Application in Numerical Prediction of Time Series*, Transactions on Fuzzy Systems, accepted for publication, 2010

[9] Huff, A., *Mapping Strategic Thought*, John Wiley and Sons, 1990

[10] Kosko, B., *Fuzzy Cognitive Maps*, International Journal of Man-Machine Studies, 1986

[11] Kosko, B., *Fuzzy Engineering*, Prentice-Hall International Editions, 1997

[12] Kosko, B., *Fuzzy Thinking*, Hyperion, 1993

[13] Kosko, B., Neural Networks and Fuzzy Systems*: A Dynamical Systems Approach to Machine Intelligence*, Prentice-Hall International Editions, 1992

[14] Laukkanen, M., *Comparative Cause Mapping of Management Cognition: A computer Database For Natural Data* , Helsinki School of Economics and Business Publications, 1992

[15] Laukkanen, M., *Conducting Causal Mapping Research: Opportunities and Challenges*, in Managerial and Organisational Cognition, edited by Eden, C. and Spender, Sage, 1998

[16] Miao, Y., Liu, Z., Siew, C., Miao, C., *Dynamical Cognitive Network - an Extension of Fuzzy Cognitive Map*, IEEE Transaction on Fuzzy Systems, vol. 9, no. 5, 2001, pp. 760–770.

[17] Minsky M L and Papert S A, *Perceptrons*, MIT Press, 1969

[18] Mohr, S. T., *The Use and Interpretation of Fuzzy Cognitive Maps*, Master Project, Rensselaer Polytechnic Institute, 1997 (http://www.voicenet.com/~smohr/fcm _white.html)

[19] Novikoff, A. B., *On convergence proofs on perceptrons*, Symposium on the Mathematical Theory of Automata, 12, Polytechnic Institute of Brooklyn , 1962, pp. 615-622, .

[20] Park, K., and Kim, S., *Fuzzy cognitive maps considering time relationships*, International Journal of Human-Computer Studies, vol. 42, no. 2, pp. 157–168, 1995.

[21] Rosenblatt, F., *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*, Spartan Books, 1962

[22] Spirtes, P., Glymor C., Scheines, R., *Causation, Prediction, and Search*, 2nd Ed., MIT Press, 2000

[23] Stach W., Kurgan L., Pedrycz W., and Reformat M., *Genetic learning of fuzzy cognitive maps*, Fuzzy Sets and Systems, vol. 153, no. 3, 2005, pp. 371–401

[24] Stach W., Kurgan L.A., and Pedrycz W., *Numerical and linguistic prediction of time series with the use of fuzzy cognitive maps*, IEEE Transactions on Fuzzy Systems, vol. 16, no. 1, 2008, pp. 61-72

[25] Stylios, C. D., Groumpos, P. P., *A Soft Computing Approach for Modelling the Supervisor of Manufacturing System*, Journal of Intelligent and Robotic Systems 26, 1999, pp. 389–403

[26] Tsadiras, A. and Margaritis, K., *Cognitive Mapping and Certainty Neuron Fuzzy Cognitive Maps*, Information Sciences: an International Journal, vol. 101, no. 1–2, pp. 109–130, 1997.

[27] Vensim., *The complete toolkit for building and using high quality dynamic feedback models*, Vensim

[28] Zhong, H., Miao, C., Shen, Z., Feng, Y., *Temporal Fuzzy Cognitive Maps*, Proceedings of the 2008 IEEE World Congress on computational Intelligence, WCCI2008, Hong Kong, 2008, pp. 1831-1840.