On the Fuzzy Cognitive Map Attractor Distance

Alex Chong, *Member, IEEE*, and Kok Wai Wong, *Senior Member, IEEE*

***Abstract*—Fuzzy Cognitive Map (FCM) has commonly been used as a prediction tool. The FCM forward chains have been used to find answers to *what-if* questions. The process starts with the encoding of the what-if question into a stimulus vector. The vector goes through a series of vector-matrix multiplication until the FCM converges to one of the FCM attractors. The attractor is the answer to the initial question. There are several types of FCM attractors. The usefulness of the different types of attractors relies heavily on the user's objectives and interpretations. This paper presents the theoretical discussion on distance measurement among the various FCM attractor distances. Subsequently the FCM Attractor Distance (FCMAD) based on genetic algorithm is proposed. The use of this distance in FCM goal oriented analysis and FCM learning is discussed. Experiment results have confirmed the effectiveness of the proposed technique.**

I. Introduction

Fuzzy Cognitive Map (FCM) is proposed by Kosko [1] as a digraph-based tool for representing causal reasoning. FCMs are signed digraphs with feedback. Nodes represent concepts and the edges represent causal relationships among the concepts in the problem domain. Numerically, a FCM with *n* nodes can be represented as an *n* x *n* matrix (see Figure 2).

An FCM state vector gives a snapshot of events in time. For example, the state s_t = [0.9 0 0 0 0] at time step *t* represents the event where "the weather is bad". The state at the next time step s_{t+1} can be inferred by multiplying state s_t with the FCM matrix.

FCM uses forward chains to find answers to the *what-if* questions. Forward chaining involves a series of repeating vector-matrix multiplication and transformation. At each time step *t* the state vector is calculated as:

$$C_t = S(C_t\text{-}1 \text{ x } E) \quad \textbf{(1)}$$

where C_t is the state vector at time step *t*, *E* is the FCM matrix, and $S(x) = (s_1(x_1) \ldots s_n(x_n))$ is a function on the resulting vector of the multiplication. Each s_i is a function that transforms the i^{th} value of the resulting vector into a value within the valid range. The sigmoid function (2) is commonly used for this purpose.

$$S_i(x_i) = \frac{1}{1+e^{-cx_i}} \quad \textbf{(2)}$$

The vector-matrix multiplication is repeated until FCM converges or "settles down" to one of the following:

1. *A fixed point attractor*. A repeating state occurs. For example: [1100] [0100] [0100] [0100] … where [0100] is known as the fixed point attractor.
2. *A limit cycle*. A few states keep repeating, forming a cycle. For example: [0100] [0110] [0100] [0110] … where [0100] [0110] is a limit cycle.
3. *A chaotic attractor*. The states have different values with each step [1]. Repeating states will never be found.

Each of the attractors discussed above is an answer to a causal *what-if* question. In general, a stimulus vector *x* and a fixed-point attractor *y* indicate that the event(s) represented by *x* eventually leads to occurrence of event(s) represented by *y*.

The usefulness of the attractors (answers) relies heavily on the user's objectives. A fixed-point attractor provides straightforward answers to causal what-if questions. For example, a fixed-point attractor at [0 0 0 1 0] represent "Road Accident". A Limit Cycle provides the user with a deterministic behavior of the environment. It allows user to predict future events. A chaotic attractor can assist in simulation by feeding the environment with endless events (states) so that a realistic effect can be obtained.

Figure 1 shows the two-dimensional representation of the various attractors in the FCM state space. It can be observed that distance measurements between attractors are not straightforward. Without proper distance measurements, it is difficult to tell how the two FCM answers are different.

This paper presents a theoretical discussion on this subject. Subsequently, the FCM Attractor Distance (FCMAD) is proposed. The use of FCMAD to solve two important problems, namely the FCM goal oriented analysis and FCM learning. Experiments are carried out to validate the effectiveness of FCMAD and the results are reported in the Experiment Section of this paper.

Manuscript received March, 30 2007

Alex Chong is with ChongyLab, Western Australia (e-mail: cchong@ieee.org)

Kok Wai Wong is with Murdoch University, School of Information Technology, South St, Murdoch, Western Australia 6150 (corresponding author phone: +618 9360 6100; fax: +618 9360 2941; e-mail: k.wong@murdoch.edu.au).

.

1-4244-1340-0/07/$25.00 ©2007 IEEE

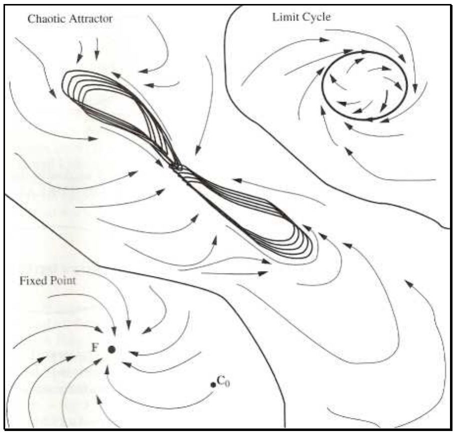
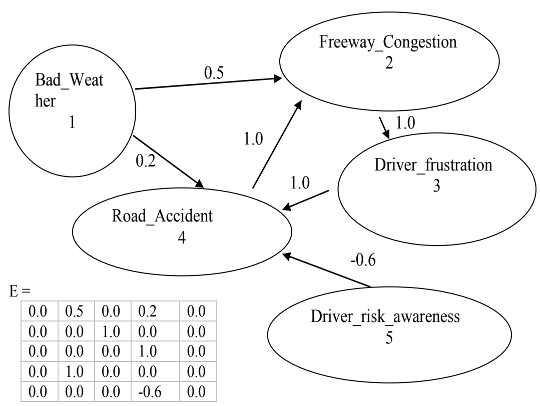


Figure 1 Two-dimensional representation of the fixed point, limit cycle and chaotic attractors extracted from [1].



0.0	0.5	0.0	0.2	0.0
0.0	0.0	1.0	0.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	1.0	0.0	0.0	0.0
0.0	0.0	0.0	-0.6	0.0

Figure 2 Simple Fuzzy Cognitive Map concerning the freeway condition at rush hour. Matrix E is the numerical representation of the FCM. The state vector s_t = [0.9 0 0 0 0] gives the snapshot of event at time step *t*. In this case, the event can be interpreted as "Bad weather" [5].

II. FUZZY COGNITIVE MAP ATTRACTOR DISTANCE

Fundamental to the problem of FCM training and goal oriented analysis is the distance measure of the FCM attractors. Let *A* and *B* be attractors in a FCM, a meaningful distance measure *d* should at least satisfy the following basic axioms:

Axiom 1: $d(A, B) = d(B, A)$
Axiom 2: $d(0, 1)$ is maximum
Axiom 3: $d(A, B) = 0$ when $A = B$

Here, we propose distance measures that satisfy the above axioms. First consider the simple scenario where $|A| = |B| = 1$, i.e. attractors with one state. In this case, the problem is reduced to the finding of distance between two vectors. The distance measure could be defined as some form of weighted squared distances between the vectors.

$$d_s(A, B) = TRACE((A - B)\, W\, (A - B)^{T}) \quad \textbf{(3)}$$

Alternatively, a maximum-based distance measure can be defined as:

$$d_s(A, B) = f((A - B)^{T} (A - B)) \quad \textbf{(4)}$$
$$\text{where } f(X) = \max_i (X_i)$$

In (4), X_i is the i^{th} diagonal element of *X*. What we have been discussing above is a special case where the attractors have only one state. In the general case, we are dealing with the comparison of arbitrary series of vectors. Fortunately, the series of vectors follows a pattern. Theorem 1 describes this.

Theorem 1: Let *A* and *B* be two *FCM* attractors. If there exist a pair of states $a \in A$, $b \in B$, such that $a = b$, then $A = B$, that is *A* and *B* are one and the same attractor.

Proof: Let a_i, be the i^{th} state in *A* and b_j be the j^{th} state in *B*. We have

$$a_{i+1} = f(a_i \times E) = f(b_i \times E) = b_{i+1} \quad \textbf{(5)}$$

and similarly,

$$a_{i-1} = f(a_{i-2} \times E) = f(b_{i-2} \times E) = b_{i-1} \quad \textbf{(6)}$$

where *E* is the *FCM* containing both the attractors *A* and *B*

Based on the theorem, the following distance measurement for two attractors *A* and *B* ($|A| \neq |B|$) is proposed:

$$d_A(A, B) = \min_{d \in D}(d) \quad \textbf{(7)}$$
$$\textit{Where } D = \{d_s(a, b) \mid \forall a \in A, \forall b \in B\}$$

In (7), *D* is the set of distances between the different combination of states in *A* and *B*. In the following, the behaviours of the distances proposed will be examined.

Observation 1

Distance (3) and (4) satisfies Axiom 1. The squared term ensured the commutativity:

$$\sum_i w_i (A_i - B_i)^2 = \sum_i w_i (B_i - A_i)^2$$

and

$$\max_i w_i (A_i - B_i)^2 = \max_i w_i (B_i - A_i)^2$$

Observation 2

Distance (3) and (4) satisfies Axiom 2. Since 0 and 1 are minimum and maximum state values in FCM respectively, the squared term $(A_i - B_i)^2$ reaches its maximum, and, A and B are zeros and ones matrices/vectors.

Observation 3

Distance (3) and (4) satisfies Axiom 3. When A = B, the term (A-B) will be zero and consequently $d_s(A, B) = 0$.

Observation 4

Distance (7) satisfies Axiom 1. Since $d_s(a, b)$ in (7) satisfies the axiom, the set of distances D will be the same for both $d_A(A, B)$ and $d_A(B, A)$. It follows that $d_A(A, B)$ must produce the same value as $d_A(B, A)$ through the *min* operator in (7).

Observation 5

Distance (7) satisfies Axiom 2. This can be observed in a way similar to Observation 4. The distance will achieve its maximum value as $d_s(a, b)$ produces the maximum value (as shown in Observation 2).

Observation 6

Distance (7) satisfies Axiom 3. In (7), when $d_A(A, B) = \min_{d \in D}(d) = 0$, there exist a pair of state vectors $a = b$ such that $a \in A$, $b \in B$. It follows from Theorem 1 that:

$$d_A(A, B) = 0 \rightarrow A = B.$$

III. FCM Attractor Distance For Goal Oriented Analysis

FCMs are predominantly used as predictive tools. FCM prediction is made possible by the encoding of snapshot of events into state vectors. Given the state vector s_t at time step t, FCM forward chains to predict future events via a series of vector matrix multiplication:

$$s_{t+1} = f(s_t \text{ x } E) \quad \textbf{(8)}$$

where f is a transformation function. FCM backward chaining to perform goal oriented analysis was first outlined in [2,3]. Keeping a goal state in mind, we can back step to find the state that leads to the desired state. The process involves the finding of the state s_{t-1} by multiplying s_t with the inverse matrix E^{-1}. Unfortunately, it is well known that not all FCM matrixes have an inverse. Specifically, an inverse exists only if E is non-singular, a condition that cannot be guarantee in all situations. FCM backward chaining remains an unsolved problem.

A more general backward chaining approach is possible with the use of the proposed FCM Attractor Distance (FCMAD). Instead of finding s_{t-1}, it is often sufficient, for practical applications, to find $\sim s_{t-1}$, an approximation state that leads to the state $\sim s_t$ that is sufficiently close to s_t. Fundamental to this problem is the distance measurement between $\sim s_t$ that is sufficiently close to s_t. This is the distance between FCM hidden patterns. In the following experiment section, we show that (7) can be used for this purpose. From here, the backward chaining is then reduced to the distance optimisation problem.

In general, FCM forward chaining answers *what-if* questions whereas backward chaining answers *why* and *how-to* questions. Clearly both are equally important operations for most practical applications. [4] gives a good overview of FCM structure and operation.

IV. FCM Attractor Distance for FCM Learning

FCM learning refers to automatic FCM performance tuning based on a set of training data. The goal is to adapt the FCM edges so that the resulting model produces the desired state vector sequences for a given input vector.

The Differential Hebbian Learning (DHL) in [2] is proposed for this purpose. Training is done by going through each state vector and modifying the FCM matrix based on the DHL law. The DHL law correlates the changes of two concepts. If concept A and concept B move in the same direction (e.g. B increases when A increases), the edge strength between the two concepts is increased; otherwise the edge strength is decreased.

DHL is an unsupervised FCM learning scheme. It gives FCMs the ability to adapt causal link strengths based on sequences of state vectors (training data) but does not mathematically guarantee the encoding the training sequences into the FCM [1]. For more effective learning, a supervised scheme is needed. In the experiment section of this paper, it is shown that the proposed FCM Attractor Distance (FCMAD) can be used in conjunction with an optimization algorithm (e.g. Genetic Algorithm) to perform supervised learning for FCMs.

Given a set of training data, the learning scheme will iteratively modify the FCM edge values until it produces result sufficiently close to the desired output based on (7).

V. Experiments

In this section, the practical use of the proposed FCM Attractor Distance in goal oriented analysis and automated training are demonstrated. Figure 2 shows the classic FCM and its corresponding edge matrix drawn from the literature [5]. This FCM will be used throughout the experiments. The experiments are described in the following two subsections.

A. FCM Attractor Distance in Goal Oriented Analysis.

In this experiment, the FCM Attractor Distance (FCMAD) is used to perform the goal oriented analysis. The following steps are used to perform the goal oriented analysis.

1. Define the *user target vector, G*. This is the goal or desired state vector. Feeding this vector into the back inference process is equivalent to asking the question "what leads to the occurrence of this state?"
2. Define the *user weight matrix, W* to specify the importance of each node to achieve its goal state. The relative importance of each node is encoded into a diagonal matrix. The i^{th} diagonal value of the matrix, $d_i \in [0,1]$, defines the importance of the i^{th} FCM node.
3. Define the *user policy vector, P*. This vector specifies, for each node, whether it can be a policy node. If the i^{th} node is allowed to be a policy node, the corresponding i^{th} value of the vector will be 1. Otherwise, the value is 0.
4. Optimize the objective function in equation (7). The weight matrix *W* in the equation is substitute as the user weight matrix defined in step 2.

Table 1 shows the parameters used in this experiment. The question being encoded by the *user target vector* is "What leads to road accident". Next the FCM nodes driver risk awareness (5) and bad weather (1) are set as suggestive policy nodes as they have no incoming links.

The optimization is performed using Golberg's Genetic Algorithm (GA) [6]. The algorithm starts with randomly generating the set of stimulus vectors and the corresponding policy vectors in the form of binary coded population (termed chromosomes). Using the stimulus vectors and policy vectors, forward chaining is performed to find the attractors. This is followed by the evaluation of the objective function in equation (7). Only individuals within the population that achieve good performance are kept. The processes of crossover, mutation and reinsertion are carried out to reproduce new individuals in the population. The process repeats until satisfactory performance is obtained or a preset number of iterations is reached. Upon completion, the population is decoded into the stimulus vectors and stimulus policy vectors.

TABLE 1
PARAMETERS USED IN EXPERIMENT ONE

Parameter	Value
User target vector *G*	[0.0 0.0 0.0 1.0 0.0]
User weight matrix *W*	$\begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$
User policy vector *P*	[1.0 0.0 0.0 0.0 1.0]

Figure 3 shows the GA objective function values (error) over the first 500 iterations. It can be seen that the error stop decreasing after 150 iterations. Table 2 shows some of the stimulus and policy vectors found by the optimization process. The vectors are not significantly different from one another. They resulted in the same output attractor. The attractor has a high value for the FCM node 'Road Accident'.

TABLE 2
STIMULUS & POLICY VECTORS FOUND IN EXPERIMENT ONE

Stimulus Vectors and Policy Vectors	Output Attractors
Stimulus: [1 1 0.9 0.43 0] Policy: [1 0 0 0 1]	[1 0.94 0.87 0.89 0]
Stimulus: [1 1 0.92 0.44 0] Policy: [1 0 0 0 1]	[1 0.94 0.87 0.89 0]
Stimulus: [1 1 0.6 0.43 0] Policy: [1 0 0 0 1]	[1 0.94 0.87 0.89 0]

The result stimulus vectors have high values for all FCM nodes except "Driver risk awareness". That is to say that "Bad Weather", "Freeway Congestion", and "Driver Frustration" have high chance of leading to "Road Accident". As for "Driver risk awareness", this should prevent the "Road Accident", which can also be shown in the FCM as it is the only node that has a negative correlation with the target node "Road accident". Intuitively, having high driver risk awareness will result in low road accident rate, which can be shown by the FCM. Through this experiment, it can be shown that the use of the FCM Attractor Distance in goal oriented analysis can produce intuitive and useful results.

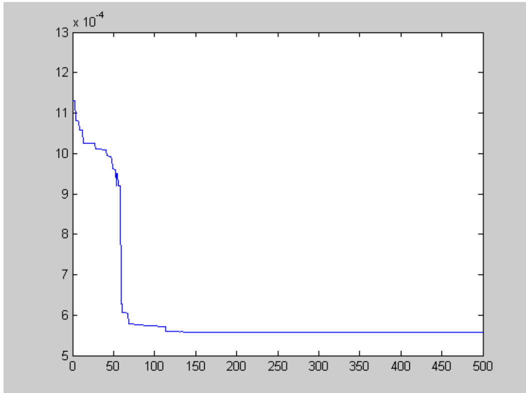


Figure 3 Genetic Algorithm Objective Function Value over five hundred iterations for FCM Goal Oriented Analysis.

B. FCM Attractor Distance in Automated Fuzzy Cognitive Training

In this experiment, the FCM Attractor Distance (FCMAD) is used to perform automatic FCM weight tuning. The following steps are used to perform the tuning.

1. Define the list of *training input vectors I* and their corresponding *desired output vectors O*. Table 3 shows the training data used. The output vectors are generated by feeding the input vectors into the FCM in Figure 2.
2. Define the *user policy vector, P*. This vector specifies, for each input vector $i \in I$, the set of nodes that should be set as policy node to produce the corresponding result vector $o \in O$.
3. Optimize the objective function in equation (7). The weight matrix W is note used in this experiment (i.e. set to all ones).

The optimization process is done in a way analogous to the previous experiment except GA populations are used to encode the FCM edge matrix rather than the stimulus vectors. The process starts a set of random population, each representing an FCM matrix. To simplify the training, some optimization constraints have been added to the process. Positive-value edges in Figure 2 are constrained to have values 0 to 1. The edge connecting node 5 to node 4 is constrained to have values 0 to -1. The rest of the edges are constrained to the value 0.

The constraints can be considered as expert knowledge used to facilitate the training. They prevent counter intuitive edges to be formed and ensure edge values to be adjusted in the right directions.

TABLE 3
PARAMETERS USED IN EXPERIMENT ONE

Training Input	Training Output
[1 0 0 0 0]	[1 0.93 0.87 0.82 0.50]
[0 1 0 0 0]	[0.5 1 0.88 0.80 0.50]
[0 0 1 0 0]	[0.5 0.90 1 0.83 0.5]
[0 0 0 1 0]	[0.5 0.92 0.86 1 0.50]
[0 0 0 1 1]	[0.5 0.86 0.85 0.67 1]

Figure 4 shows the GA objective function values (error) over the first 2000 iterations. It can be seen that the error is close to zero as the number of iteration approaches 2000.

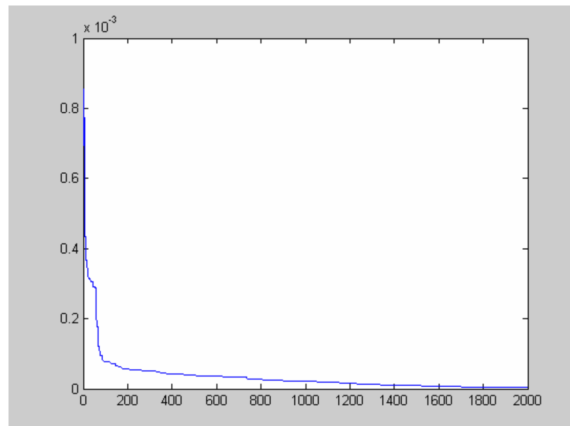


Figure 4 Genetic Algorithm Objective Function Value over five hundred iterations for FCM edge values training.

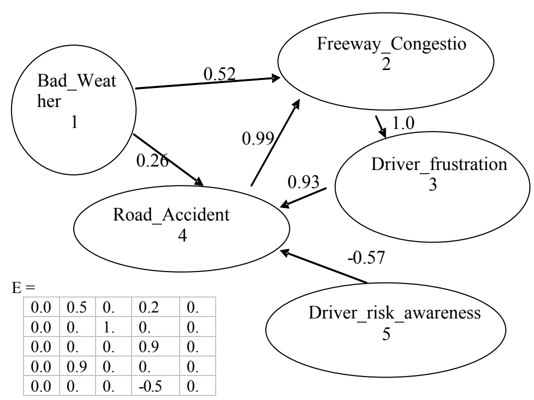


Figure 5 Genetic Algorithm Objective Function Value over five hundred iterations for FCM edge values training.

The result of the experiment is shown in Figure 5. It can be observed that the tuned FCM resembles closely the original (desired) FCM in Figure 2. This has also shown that the proposed learning algorithm can help to fine tune the performance of the FCM using a set of training data. As we are using genetic algorithm in the learning algorithm, it can perform the learning automatically with increasing improvement on the search over time.

VI. CONCLUSIONS

FCM has been very popular in many decision support systems. Most of the application of FCM is based on the forward chaining analysis, which is used to predict future behavior. Backward chaining analysis can help to solve the goal oriented problem. In this paper, we proposed the FCM Attractor Distance using genetic algorithm to facilitate the backward chaining analysis. From the experiment results, it has shown that the proposed technique can provide reasonable and intuitive results. At the same time, it has also shown that the proposed technique can be used as a good alternative as a learning algorithm to fine tune the FCM.

REFERENCES

[1] B. Kosko, *Fuzzy Engineering*, 1 ed. N.J.: Prentice Hall, 1997.

[2] B. Kosko, and Dickerson, J., "Hidden Patterns in Combined and Adaptive Knowledge Networks," *International Journal of Approximate Reasoning*, vol. 2, 1988, pp. 377-393.

[3] K. Lee, C., and Kim, H., S., "A Fuzzy Cognitive Map-Based Bi-Directional Inference Mechanism: An Application to Stock Investment Analysis," *Intelligent Systems in Accounting, Finance and Management*, vol. 6, 1997, pp. 41-57.

[4] M. S. Khan, and Quaddus, M., "Group Decision Support using Fuzzy Cognitive Maps for Causal Reasoning," *Group Decision and Negotiation Journal*, vol. 13, no. 5, 2004, pp. 463-480.

[5] Khan, M.S., Khor, S., and Chong, A. "Fuzzy Cognitive Maps with Genetic Algorithm for Goal-oriented Decision Support," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol.12, 2004, pp.31-42.

[6] D. Goldberg, *Genetic Algorithms in Search, Optimisation and Machine Learning*. Massachusetts: Addison Wesley, 1989.