Information Sciences 349–350 (2016) 154–171

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On the convergence of sigmoid Fuzzy Cognitive Maps

Gonzalo Nápoles [a,b,*], Elpiniki Papageorgiou [a,c], Rafael Bello [b], Koen Vanhoof [a]

[a] Faculty of Business Economics, Hasselt University, Belgium
[b] Department of Computer Science, Universidad Central "Marta Abreu" de Las Villas, Carretera a Camajuaní km. 5 1/2, Santa Clara, Villa Clara, Cuba
[c] Department of Computer Engineering, Technological Education Institute of Central Greece, Greece

ARTICLE INFO

Article history:
Received 21 May 2015
Revised 16 February 2016
Accepted 21 February 2016
Available online 27 February 2016

Keywords:
Fuzzy Cognitive Maps
Pattern classification
Convergence

ABSTRACT

Fuzzy Cognitive Maps (FCM) are Recurrent Neural Networks that are used for modeling complex dynamical systems using causal relations. Similarly to other recurrent models, a FCM-based system repeatedly propagates an initial activation vector over the causal network until either the map converges to a fixed-point or a maximal number of cycles is reached. The former scenario leads to a hidden pattern, whereas the latter implies that cyclic or chaotic configurations may be produced. It should be highlighted that FCM equipped with discrete transfer functions never exhibit chaotic states, but this premise cannot be ensured for systems having continuous neurons. Recently, a few studies dealing with convergence on continuous FCM have been introduced. However, such methods are not suitable for FCM-based systems used in pattern classification environments. In this paper, we first review a new heuristic procedure called Stability based on sigmoid Functions, which allows to improve the convergence on sigmoid FCM, without modifying the weights configuration. Afterwards, we examine some drawbacks that affect the algorithm performance and introduce solutions to enhance its performance in pattern classification environments. Additionally, we formalize several definitions which were omitted in the original research. These solutions lead to accurate classifiers and prevent specific scenarios where the method may fail. Towards the end, we conduct numerical simulations across six FCM-based classifiers with unstable features in order to evaluate the proposed improvements in pattern classification environments.

1. Introduction

In recent years the FCM theory [17] has become a proper tool for designing knowledge-based systems with interpretable features. Essentially, a FCM is an information network where graph nodes represent objects, states, concepts or entities of the investigated system and comprise a precise meaning for the problem domain. Such concepts are equivalent to neurons in connectionist models and are connected by causal relationships that normally take values in the range $[-1, 1]$. The direction and intensity of causal relations involve the quantification of a fuzzy linguistic variable which is assigned by experts during the modeling phase [18]. These elements iteratively interact during the inference phase when updating the activation value of each neuron. More explicitly, a FCM exploits an activation vector by using a rule similar to the standard McCulloch-Pits

* Corresponding author at: Department of Computer Science, Universidad Central "Marta Abreu" de Las Villas, Carretera a Camajuaní km. 5 1/2, Santa Clara, Villa Clara, Cuba. Tel.: +53 42 275267, 53 42 281515.
E-mail address: gnapoles@uclv.edu.cu, napoles.gonzalo@gmail.com (G. Nápoles).

http://dx.doi.org/10.1016/j.ins.2016.02.040
0020-0255/

schema [25]. It implies that the activation degree of each map neuron is given by the value of the transformed weighted sum that this processing unit receives from connected neurons on the causal network. This activation value actually comprises an interpretable feature for the modeled system, that is, the higher the activation value of a concept, the stronger its influence (positive or negative) over the connected neural entities.

Although FCM are considered as neural systems, there are important differences regarding other types of artificial neural networks (ANN). Classical ANN models regularly perform like *black-boxes*, where both neurons and connections do not have any clear specific meaning for the problem itself. However, all the neurons of the FCM have a precise meaning for the physical system and correspond to specific variables. It should be highlighted that a FCM does not comprise hidden neurons since these entities could not be interpreted, and thus, the system becomes impossible to be modeled. This suggests that the representation capability of the FCM is superior compared to ANN models.

Unlike to feed-forward artificial neural networks [2,25,44], FCM are recurrent neural models since they support backward connections and sometimes form cycles in the causal graph, which denote the underlying relation among variables. These backward relations (called feedback) enable the network to storage a previous state memory in order to compute the outputs of the current state and maintain a sort of recurrence to the past processing [9]. During the inference phase the updating rule is repeated until (i) the system converges to a fixed-point attractor or (ii) a maximal number of cycles is reached. The former implies that a hidden pattern was discovered [19] while the latter suggests that the outputs are cyclic or completely chaotic. Observe that a FCM will produce a state vector at each discrete-time step (iteration step) that comprises the activation degree of all neurons.

If the FCM is able to converge in a fixed-point attractor, then the map will produce the same output for each iteration step, and thus the activation degree of neurons will remain without change (or the changes are infinitesimal). On the other hand, a cyclic FCM produces dissimilar responses with the exception of a few states that are periodically produced. The last possible scenario is related to chaotic configurations on which the system produces different state vectors, and thus a stable pattern cannot be concluded. The fixed-point attractors are desirable since the responses computed by the modeled system become consistent. On the contrary, non-stable FCM produce inconsistent outcomes, being impossible to made suitable decisions. In these scenarios the reasoning (updating) rule stops once a maximal number of iterations is reached, and so the state vector is calculated from the last response. However, this output is partially unreliable due to the lack of convergence.

The non-stable conditions are mostly related with the causal weight matrix that describes the modeled system. More explicitly, a perfectly symmetric weight matrix implies the existence of a large number of positive cycles in the modeled system. Such cycles provide the system with positive feedback loops that amplify any initial change and thus lead to exponential growth or decline [49]. On the other hand, antisymmetric causal weight matrixes imply the existence of strong negative cycles with odd number of connections, providing the FCM with negative feedback loops that counteract any initial stimulus. Therefore, after a time period equal to the length of the cycle the neuron to which the initial stimulus was introduced will receive an influence that has an opposite sign from the initial change. This leads the system to periodic behavior and the creation of limit cycles.

A mentioned before, the responses in non-stable FCM-based models become inconsistent and making a precise decision is not possible. In such models we cannot simply modify the weight matrix to reach convergence without altering the physical representation of the system, which is often established by experts in a given domain. On the other hand, we cannot codify all scenarios to perfectly symmetric weight matrices because it implies that all pairs of neurons are equally influenced, and this statement is not true for all application domains. Even when no expert is involved during the system modeling (i.e. parameters required to design the causal network are automatically learned from historical data) the FCM may decrease its ability to recognize new patterns. That is why more efficient strategies to improve the convergence of FCM-based classifiers are still required.

Another element closely related to the convergence on FCM-based systems is the nonlinear transfer function $f(.)$ that is used to keep the activation value of neurons in the activation interval. The most frequently used variants are the sign function (which produces binary outputs), the trivalent function (which produces ternary outputs) and the sigmoid function [5]. The discrete transformation functions (i.e. bivalent and trivalent) produce a finite number of different states. This happens because FCM are deterministic and so, if it reaches a state to which it has been previously, the system will enter a closed orbit which will always repeat itself [48]. For example, a binary FCM having M neurons will produce at most 2^M different states where these states are located in the corners of the $[0, 1]^M$ hypercube. This suggests that binary FCM can converge to a fixed-point attractor or can show cyclic patterns (with an exponential period in the worst scenario) but never produce chaotic responses. A trivalent FCM will produce at most 3^M different states situated at the corners, at the middle of the edges, at the center of the sides and at the center of the $[-1, 1]^M$ hypercube. On the contrary, FCM equipped with continuous transfer functions can exhibit chaotic behaviors since the FCM could produce infinite different states freely distributed in the space defined by the $[-1, 1]^M$ hypercube.

Recently Nápoles et al. [33] proposed a heuristic algorithm for improving the convergence of sigmoid FCM, without varying the weights defined by experts or automatically computed from historical data using a learning algorithm. The procedure essentially modifies the original system by using “excitable neurons” allowing the improvement of the system convergence. The excitation (or inhibition) effect at each iteration step is achieved by using a custom transfer function $f_i(x)$ for each neuron C_i instead of using the same function for all neurons [34]. This heuristic procedure decreases the variability on the system responses for consecutive cycle steps, without altering the capability of the FCM-based model for recognizing new patterns. From numerical simulations, the authors observed four scenarios: (i) the convergence rate on stable maps was