31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Diagnosis Support using Fuzzy Cognitive Maps combined with Genetic Algorithms

Voula C. Georgopoulos, *Senior Member, IEEE* and Chrysotomos D. Stylios, *Member, IEEE*

***Abstract*— A new hybrid modeling methodology to support medical diagnosis decisions is developed here. It extends previous work on Competitive Fuzzy Cognitive Maps for Medical Diagnosis Support Systems by complementing them with Genetic Algorithms Methods for concept interaction. The synergy of these methodologies is accomplished by a new proposed algorithm that leads to more dependable Advanced Medical Diagnosis Support Systems that are suitable to handle situations where the decisions are not clearly distinct. The technique developed here is applied successfully to model and test a differential diagnosis problem from the speech pathology area for the diagnosis of language impairments.**

I. Introduction

FUZZY Cognitive Maps (FCMs) and Genetic Algorithms (GAs) are two computational intelligence techniques that have been successfully employed for developing intelligent systems, which have been applied effectively in many different application domains. These techniques incorporate knowledge, experience and useful information along with historical data and evolutionary patterns in order to adapt, handle and solve complex problems.

A Fuzzy Cognitive Map models any system as an ensemble of interconnected concepts and causal relations between them. FCMs are built on the experience and knowledge of experts. An FCM is a causal graphical descriptive model of the operation and behavior of any system; it is consisted of interrelated concepts [1]. FCMs are fuzzy signed directed graphs permitting feedback. The weighted edge w_{ij} from concept C_i to affected concept C_j corresponds to the influence of the first concept to the latter.

FCMs have been successfully used to develop Medical Decision Support System mainly for differential diagnosis [2]. However, there are complex cases where the available information is not adequate and the FCM-Diagnosis Support System cannot discriminate and reach a decision; thus, there is a need for a complementary mechanism to supplement the FCM-DSS. For these cases, it is proposed to exploit the high optimization characteristics of Genetic Algorithms, which will be applied to selected weighted edges w_{ij} of the FCM-DSS so that to lead to more confident decisions.

V.C.Georgopoulos is with the Department of Speech and Language Therapy, Technological Educational Institute of Patras, Koukouli 26334, Patras, Greece (e-mail: voula@teipat.gr).

C.D.Stylios is with the Department of Informatics and Telecommunications Technology, Technological Educational Institute of Epirus, 47100 Artas, Epirus, Greece (e-mail: stylios@teiep.gr)

This paper is organized in seven sections including this introduction. Sections II, III briefly present FCMs and GAs, respectively. Section IV introduces a novel synergist model of FCMs and GAs to develop an advanced Diagnosis Support System and Section V presents the proposed algorithm for the GA factor interaction to FCM DSS. Section VI shows how the algorithm is applied in a speech pathology problem and section VII offers conclusions.

II. Fuzzy Cognitive Maps for Medical Diagnosis Support

An advanced type of Fuzzy Cognitive Map for Medical Diagnosis is the Competitive Fuzzy Cognitive Map (CFCM) [3], which consists of two main types of concepts: diagnosis-concepts and factor-concepts. The causal relations between the two types of concepts influence the values of all concepts and determine the value of diagnosis concepts indicating final diagnosis.

Each decision concept represents a single diagnosis, which means that these concepts must be mutually exclusive if our intention is to infer always only one diagnosis. The factor-concepts can be considered as inputs of the diagnostic system including patient data, observed symptoms, patient records, experimental and laboratory tests etc.

III. Genetic Algorithms

Genetic algorithms (GA) were introduced by Holland [4] and are based on the genetic evolution of a species. Potential solutions to specific problems are encoded using simple data structures similar to chromosomes. The chromosomes make up an initial population where each individual chromosome is assigned a "fitness score" according to how good a solution to the problem it is. The higher the fitness score the more opportunity they have to produce “offspring” through mechanisms of mutation, selection and crossover as defined in [5] retaining some desirable properties from the parents. Genetic Algorithms have been successfully combined with to Fuzzy Cognitive Maps by Stach et al. [6-8] to optimize the learning process. GA are designed to evaluate existing potential solutions as well to generate new (improved) solutions to a problem for evaluation. Thus, GA can improve the quality of decision making.

IV. An Advanced Diagnostic Support System

Here, an advanced Diagnostic Support System (DSS) is proposed based on the synergy of FCMs and GA that

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

constitutes an advanced hybrid intelligent inferring methodology. When the CFCM has difficulty to infer a final decision with great certainty, then the GA is called to assist the CFCM, so that the hybrid DSS can propose a diagnosis.

Figure 1 diagrammatically shows the proposed Genetic Algorithm Factor Interaction Competitive Fuzzy Cognitive Map (GAFI-CFCM) Diagnostic Support Model. According to the methodology the patient data is input to the CFCM and the factor concepts take their initial values.. Patient information are experimental results, test results, physical examinations and other descriptions of symptoms and measurements of physical qualities. This input information can be described either in numerical values or in fuzzy linguistic values, which are then transformed into a numerical weight in the range [0,1], i.e. the allowable values for the CFCM concepts. The CFCM runs according to the algorithm described in [2] and when an equilibrium region is reached the CFCM ceases to interact. Then the values of the decision/diagnosis concepts are examined to determine if there is a distinct decision/diagnosis or not. A distinct outcome is inferred, if the value of a decision concept is surpassing the others by at least 10%. In this case the leading competitive node is the suggested diagnosis. Otherwise, when the percent difference between the two leading competitive nodes is less than 10%, then the comparison made in the "Distinct Outputs" box leads to a "NO" result, activating the GA component. The GA component is used to "manipulate" weights that correspond to between factor interactions and do not involve diagnosis concepts. These weights are usually the most difficult to assess by experts in medical diagnosis since on one hand, the interactions are complex and in many cases two-way while on the other hand, there can be a wide variation (even disagreement) between experts. In order to activate the GA component, the "reduced" weight matrix is formed from the original weight matrix W by removing the first m rows and columns which involve connections to diagnosis nodes:

$$\mathrm{Wr} = \{wr_{i,j}\}_{i,j=1,\dots,n-m} = \{w_{kl}\}_{k,l=m+1,\, m+2,\, ..n}$$

where: n is the total number of nodes

m is the number of diagnosis nodes.

Therefore, Wr contains only the weights between factor cause–effect relationships; where the column vector wr_j, corresponds to the causative effects on factor j from other factor concepts. Thus, each column vector wr_j, makes up a real-coded chromosome, which is allowed to mutate/reproduce. Whether each mutation/reproduction becomes part of the new population relies on whether a criterion is met as follows:

Each new weight matrix resulting from mutations/reproductions of Wr is then placed in the correct elements of initial weight matrix W forming a new matrix W^{new} *which is now used to run the CFCM again. If one of the two "leading" decision/diagnosis concepts is not the result of the new GAFI-CFCM or if again the 10% rule above is not met, then the new chromosomes are discarded. Otherwise, the new result (interim inferred decision) is recorded, "chromosome" weights are kept as part of the population and the algorithm is repeated for a large number of iterations. The final decision is the diagnosis with the highest probability as a result of the iterations. The detailed algorithm is discussed in the next section.*

V. Description of the GAFI-CFCM Algorithm

The Advanced Medical Decision Support System GAFI-CFCM based on the synergism of Genetic Algorithms and Competitive Fuzzy Cognitive Map (CFCM) techniques is implemented by the following proposed algorithm:

Step1. First apply the CFCM algorithm described in [2].

Step 2 When the CFCM algorithm drives decision concepts to reach steady state, measure the difference between the values of competitive decision/diagnosis concepts.

Step3 If the difference between the highest values of the decision/diagnosis concepts is more than 10%, THEN a decision/diagnosis is inferred, which is reflected in the concept with the highest value – go to Step 6. ELSE activate Genetic Algorithm.

Step 4 For only between factor interactions obtain a new reduced matrix Wr, where the weights wr_{ij} *are allow to change based on mutations, crossover, and inversion to form new "offspring" weights for between factor interactions.*

Step 5 Run the FCM based on the new values, if the two decision/diagnosis concepts having the highest values are not the same as before, OR IF the difference between the highest values of the decision/diagnosis concepts is less than 10%, discard the new weights and repeat Step 4. ELSE IF the difference between the highest values of the decision/diagnosis concepts is more than 10%, THEN an intermediate decision/diagnosis is inferred which is recorded. Repeat Step 4 – this step is iterated 1000 times.

Step 6. Final Decision/diagnosis is either the result of Step 3 OR, if GAFI-CFCM was activated, the intermediate inferred decision with the highest probability as a result of 1000 iterations of steps 4 and 5.

VI. Example from Speech and Language Pathology

Specific Language Impairment (SLI) is a language disorder that cannot be easily diagnosed because it has similar characteristics to other language disorders and there is no widely accepted method of identifying children with SLI [9]. Findings in the literature have shown that severe cases of dyslexia and mild cases of autism are disorders, whose diagnoses often have been confused with the diagnosis of SLI [10].

Fundamental factors that may appear in all three disorders with different frequency and severity in most cases were identified in [2]. The considered factors are either causative factors and/or symptoms of the disorders and are the following: 1) Reduced Lexical Abilities, 2) Problems in Syntax, 3) Problems in Grammatical Morphology, 4) Impaired or Limited Phonological Development, 5) Impaired

Use of Pragmatics, 6) Reading Difficulties, 7) Echolalia, 8) Reduced Ability of Verbal Language Comprehension, 9) Difference between Verbal - Nonverbal IQ, 10) Heredity, 11) Impaired Sociability, 12) Impaired Mobility, 13) Attention Distraction, 14) Reduced Arithmetic Ability, and 15) Limited Use of Symbolic Play.

The factors within each disorder were taken into consideration in a comparative way in the development of the model. The significance of each factor as a diagnostic criterion is defined using fuzzy variables: a) very-very important, b) very important, c) important, d) medium, e) not very important, and f) minimally important. These criteria were represented in the Competitive Fuzzy Cognitive Map Differential Diagnosis Model that was developed as the fuzzy weight with which each factor influences every one of the three diagnoses. Interactions between factor-concepts have also been found and are presented in Table I [2].

Table I. Relationships between factor-concepts

Influencing Factor	**Influenced Factors**
Reduced Lexical Abilities	▪ *Reading Difficulties* ▪ *Reduced Arithmetic Ability* ▪ *Reduced Ability of Verbal Language Comprehension* ▪ *Impaired Sociability*
Problems in Grammatical Morphology	*Reading Difficulties*
Impaired Mobility	▪ *Reading Difficulties* ▪ *Reduced Arithmetic Ability*
Attention Distraction	▪ *Reading Difficulties* ▪ *Reduced Arithmetic Ability*
Problems in Syntax	▪ *Reading Difficulties* ▪ *Problems in Grammatical Morphology*
Impaired or Limited Phonological Development	*Reading Difficulties*
Problems in Grammatical Morphology	*Problems in Syntax*

As an example of the GAFI-CFCM, we consider an input case (from a database of patients from a private speech and language clinic) taking the initial fuzzy values for the factors listed above as follows: Factors 1,2,3,4, and 9 are HIGH, factor 6 is VERY HIGH and factors 12 and 14 are MEDIUM. No values were reported for factors 5, 7, 8, 10, 13 and 15. These values are based on the patient's history and test results and it is often typical for a patient's history to be incomplete or test results be inconclusive or not available.

If we use the input information of Table I in the CFCM model, after 4 simulation iterations, equilibrium is reached where decision concepts take the values:

$$SLI= 0.9281 \quad Dyslexia=0.9645 \quad Autism=0.7935$$

It is apparent, that two of the three possible diagnoses (SLI and Dyslexia) have high values and their difference of 3.8% is less than 10%. Thus, the CFCM cannot support the diagnosis of Dyslexia, which has the highest value, without additional information.

Then, we test the same input data for the GAFI-CFCM model. According to the proposed algorithm, for the same input case, the distinction between SLI and Dyslexia decisions is "No" leading to the activation of the GA component of the Diagnosis Support System.

The GAFI-CFCM through mutations/reproductions of the between factor weights of the CFCM results in a probability of Dyslexia $p= 0.9578$ for the iterations of the algorithm where there is a difference greater than 10% between the leading two decision concepts. This leads to a more confident diagnosis, since due to the GAFI-CFCM none of the outcomes were 'Autism' and 'SLI' has a probability of $p=0.0422$ for the iterations where there is a "clear" decision. Thus, it is obvious that the concept of 'Dyslexia' dominates over the values of 'SLI' and 'Autism' concepts and thus, the diagnosis of Dyslexia is proposed for this case. It should be noted, that this diagnosis was independently confirmed by two specialized clinicians.

This can also be seen in Figure 2 which shows the result of 1000 iterations of the GAFI-CFCM for the input example of Table I, as well as the detail for the first 100 iterations.

VII. Conclusion

Here, a novel hybrid intelligent methodology was proposed, specifically, the GAFI-CFCM combining GA and FCMs. The proposed model was successfully applied to a simple problem proving that a sufficient Diagnostic Support Model was developed, which processes the information about a patient in such a way that out of three possible diagnoses, there was leading to the diagnosis of the most probable disorder is reached. This hybrid technique is based on two complementary techniques FCMs and GAs in order to overcome the situation where one technique does not infer a unique decision/diagnosis. This promising approach will be tested on other decision support problems where the possible outcomes are more than three increasing the complexity of the problem.

References

[1] B. Kosko, "Fuzzy cognitive maps," International Journal of Man-Machine Studies, Vol. 24, pp.65-75, 1986.

[2] V. C. Georgopoulos, G. A. Malandraki, and C.D. Stylios, "A fuzzy cognitive map approach to differential diagnosis of specific language impairment," *Journal of Artificial Intelligence in Medicine*, Vol. 29, 2003, pp. 261–278.

[3] V. C. Georgopoulos and C. D. Stylios. "Augmented fuzzy cognitive maps supplemented with case based reasoning for advanced medical decision support," *Soft Computing for Information Processing and Analysis Enhancing the Power of the Information Technology*. M. Nikravesh, L. A Zadeh, J. Kacprzyk (Eds), 2005, pp.391-406.

[4] J. H. Holland, "Adaptation in natural and artificial systems", University of Michigan press; 1975.

[5] D. E. Goldberg, "Genetic algorithms: search, optimization and machine learning", Addison-Wesley; 1989.

[6] W. Stach, L. Kurgan, W. Pedrycz, and M,Reformat, "Genetic Learning of Fuzzy Cognitive Maps," *Fuzzy Sets and Systems*, vol. 153, no.3, 2005, pp.371-401.

[7] W. Stach, L. Kurgan, and W. Pedrycz,"A Survey of Fuzzy Cognitive Map Learning Methods," In: Grzegorzewski, P., Krawczak, M., Zadrozny, S., (Eds.), Issues in Soft Computing: Theory and Applications, 2005, pp.71-84.

[8] W. Stach, L. Kurgan and W. Pedrycz, "Parallel Learning of Large Fuzzy Cognitive Maps," Proceedings of the International Joint Conference on Neural Networks (IJCNN 2007), pp.1584-1589.

[9] E. Krasswski and E. Plante,"IQ variability in children with SLI: implications for use of cognitive referencing in determining SLI," *J. of Communication Disorders*. 1997, Vol.30, pp.1-9.

[10] L. B. Leonard, *Children with specific language impairment*, MIT Press, Cambridge 2000.

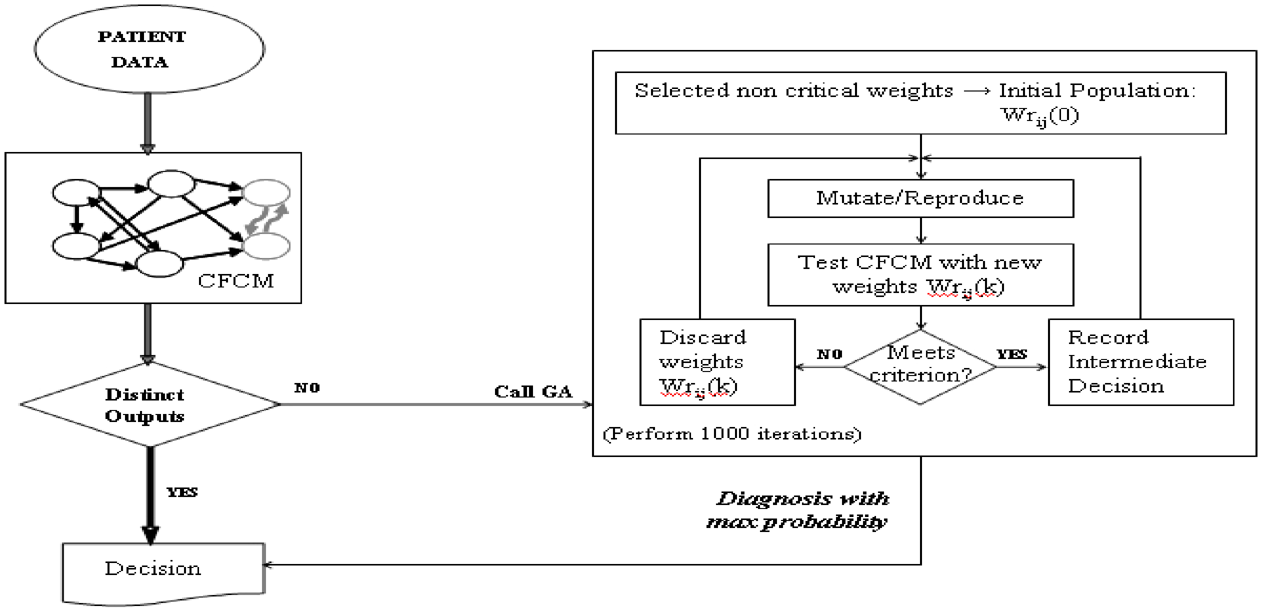

Figure 1. GA factor interaction Competitive FCM.

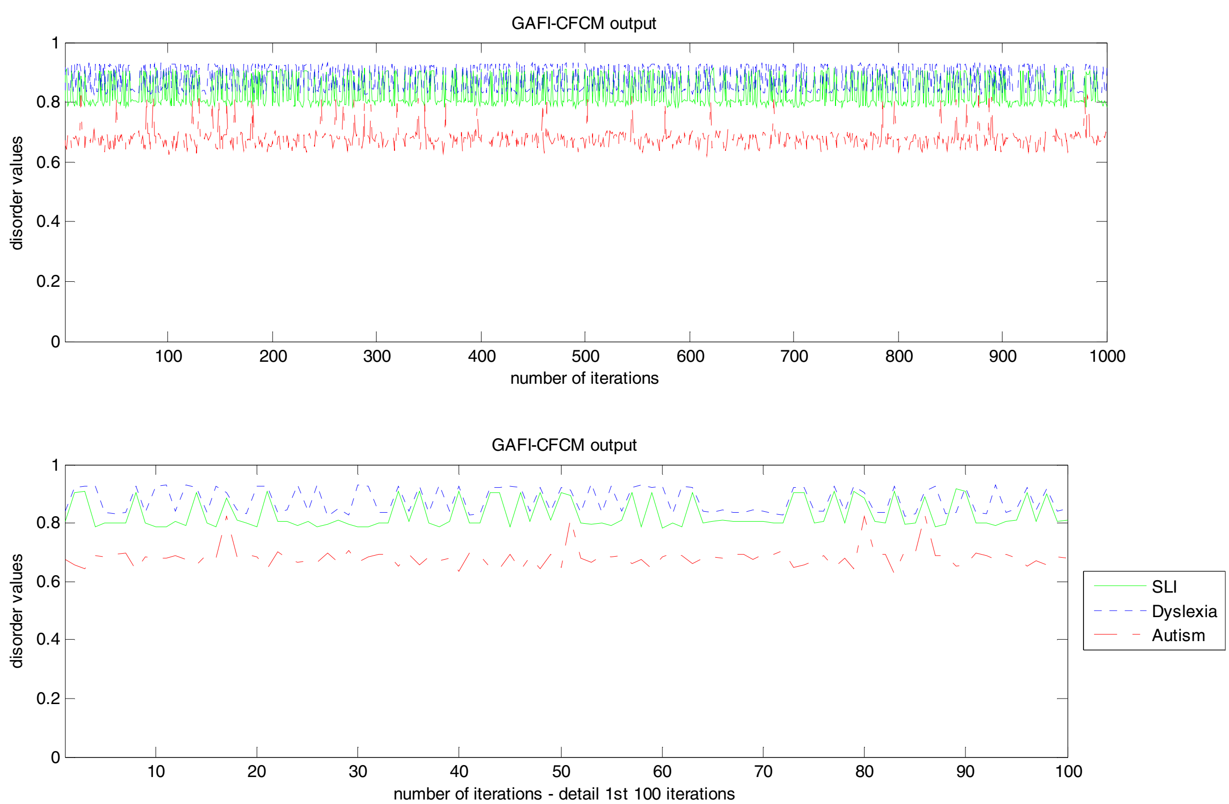

Fig. 2 Result of the GA-Factor Interaction CFCM after 1000 iterations as well as detail showing the first 100 iterations.