A Novel Approach on Designing Augmented Fuzzy Cognitive Maps Using Fuzzified Decision Trees

Elpiniki I. Papageorgiou

Department of Informatics & Computer Technology, Technological Educational Institute of Lamia, 3rd Km PEO Lamia-Athens, 35100 Lamia, Greece
epapageorgiou@teilam.gr

Abstract. This paper proposes a new methodology for designing Fuzzy Cognitive Maps using crisp decision trees that have been fuzzified. Fuzzy cognitive map is a knowledge-based technique that works as an artificial cognitive network inheriting the main aspects of cognitive maps and artificial neural networks. Decision trees, in the other hand, are well known intelligent techniques that extract rules from both symbolic and numeric data. Fuzzy theoretical techniques are used to fuzzify crisp decision trees in order to soften decision boundaries at decision nodes inherent in this type of trees. Comparisons between crisp decision trees and the fuzzified decision trees suggest that the later fuzzy tree is significantly more robust and produces a more balanced decision making. The approach proposed in this paper could incorporate any type of fuzzy decision trees. Through this methodology, new linguistic weights were determined in FCM model, thus producing augmented FCM tool. The framework is consisted of a new fuzzy algorithm to generate linguistic weights that describe the cause-effect relationships among the concepts of the FCM model, from induced fuzzy decision trees.

1 Introduction

Nowadays, the knowledge acquisition and representation constitutes a major knowledge engineering bottleneck. A large number of techniques in the field of artificial intelligence used to represent knowledge: production rules, decision trees, rule-based architectures semantic nets, frameworks, fuzzy logic, causal cognitive maps, among others. The decision trees gained popularity because of their conceptual transparency. The well-developed design methodology comes with efficient design techniques supporting their construction, cf. [1-3]. The decision trees generated by these methods were found useful in building knowledge-based expert systems. Due to the character of continuous attributes as well as various facets of uncertainty one has to take into consideration, there has been a visible trend to cope with the factor of fuzziness when carrying out learning from examples in the case of tree induction. In a nutshell, this trend gave rise to the name of fuzzy decision trees and has resulted in a series of development alternatives; cf. [4-6]. The incorporation of fuzzy sets [7-10] into decision trees enables us to combine the uncertainty handling and approximate reasoning capabilities of the former with the comprehensibility and ease of application of the latter. Fuzzy decision trees [10,11] assume that all domain attributes

Z. Cai et al. (Eds.): ISICA 2009, CCIS 51, pp. 266–275, 2009.

or linguistic variables have pre-defined fuzzy terms for each fuzzy attribute. Those could be determined in a data driven manner. The information gain measure, used for splitting a node, is modified for fuzzy representation and a pattern can have nonzero degree of matching to one or more leaves [12,13].

Fuzzy logic and causal cognitive maps, in the other hand, are some of the main topics of artificial intelligence on representation of knowledge and approximation of reasoning with uncertainty [14]. The choice of a particular technique is based on two main factors: the nature of the application and the user's skills. The fuzzy logic theory, based on representation of knowledge and approximation of reasoning with uncertainty, is very close to the expert's reasoning, and it is well known as artificial intelligence-based method, especially in the field of medical decision making. An outcome of this theory is fuzzy cognitive maps [15,16]. Fuzzy cognitive maps are diagrams used as causal representations between knowledge/data to represent events relations. They are modeling methods based on knowledge and experience for describing particular domains using concepts (variables, states, inputs, outputs) and the relationships between them. FCM can describe any system using a model having signed causality (that indicates positive or negative relationship), strengths of the causal relationships (that take fuzzy values), and causal links that are dynamic (i.e. the effect of a change in one concept/node affects other nodes, which in turn may affect other nodes).

Most decision tree induction methods used for extracting knowledge in classification problems do not deal with cognitive uncertainties such as vagueness and ambiguity associated with human thinking and perception. Fuzzy decision trees represent classification knowledge more naturally to the way of human thinking and are more robust in tolerating imprecise, conflict, and missing information.

In this work, a new algorithm for constructing fuzzy cognitive maps by using pre-generated fuzzy decision trees is proposed. The methodology is partly data driven and knowledge driven so some expert knowledge of the domain is required.

The fuzzy decision tree approach is used to implement the fuzzy algorithmic methodology in order to assign new linguistic weights among the FCM nodes as well as new paths between FCM nodes that enhance their structure and improve their operational ability to handle with complex modeling processes. This naturally enhances the representative power of FCMs with the knowledge component inherent in fuzzy decision trees rule induction.

2 Main aspects of fuzzy decision trees

Fuzzy decision trees are an extension of the classical artificial intelligence concept of decision trees. The main fundamental difference between fuzzy and crisp trees is that with fuzzy trees, gradual transitions exist between attribute values [7]. The reasoning process within the tree allows all rules to be fired to some degree, with the final crisp classification being the result of combining all membership grades. Recent approaches to developing such trees were through modifications to the ID3 algorithm [3,5,6,8,18]. Sison and Chong [3] proposed a fuzzy version of ID3 which automatically generated a fuzzy rule base for a plant controller from a set of input–

output data. Umano et al. [5] also proposed a new fuzzy ID3 algorithm. This algorithm generates an understandable fuzzy decision tree using fuzzy sets defined by the user. The fuzzy tree methodologies proposed by [3,5] require the data to have been pre-fuzzified before the fuzzy decision trees are induced.

More recent work by Janikow involves the induction of fuzzy decision trees directly from data sets by the FID algorithm [10,11]. The [10] takes a detailed introduction about the non fuzzy rules and the different kind of fuzzy rules.

In this point it is essential to refer that the data (real values) are partitioned into fuzzy sets by experts.

This approach consists on the following steps:

Step 1: A fuzzy clustering algorithm is used for input domain partition. The supervised method takes into account the class labels during the clustering. Therefore the resulted partitions, the fuzzy membership functions (fuzzy sets) represent not only the distribution of data, but the distribution of the classes too.

Step 2: During a pre-pruning method the resulted partitions could analyze and combine the unduly overlapped fuzzy sets.

Step 3: The results of the pre-pruning step are input parameters (beside data) for the tree induction algorithm. The applied tree induction method is the FID (Fuzzy Induction on Decision Tree) algorithm by C. Z. Janikow.

Step 4: The fuzzy ID3 is used to extract rules which are then used for generating fuzzy rule base.

Step 5: While the FID algorithm could generate larger and complex decision tree as it is necessary, therefore a post pruning method is applied. The rule which yields the maximal fulfillment degree in the least number of cases is deleted.

This method provides compact fuzzy rule base that can be used for building FCM-DSS.

2.1 Fuzzy Cognitive Mapping causal algebra

Fuzzy cognitive maps are an intelligent modeling methodology for complex decision systems, which originated from the combination of Fuzzy Logic and Neural Networks [14]. An FCM describes the behavior of an intelligent system in terms of concepts; each concept represents an entity, a state, a variable, or a characteristic of the system [15]. FCM nodes are named by such concepts forming the set of concepts $C = \{C1,C2, . . . ,Cn\}$. Arcs (*Cj*,*Ci*) are oriented and represent causal links between concepts; that is how concept *Cj* causes concept *Ci* Weights of arcs are associated with a weight value matrix *Wn·n*, where each element of the matrix *wji* taking values in [-1, . . .,1] . Kosko has developed a fuzzy causal algebra that describes the causal propagation and combination of concepts in an FCM. The algebra depends only on the partial ordering P, the range set of the fuzzy causal edge function e, and on general fuzzy-graph properties (e.g., path connectivity). Kosko notes that this algebra can be used on any digraph knowledge representation scheme.

A causal path from some concept node *Ci* to concept node *Cj*, say *Ci--~Ck1, Ckl--~… Ckn, Ckn --~Cj,* can be indicated by the sequence (i, *k, ... kn,j).* Then the indirect effect of *Ci* on *Cj* is the causality *C~I* imparts to *Cj* via the path *(i, kl . . . . . kn,j).* The total effect of *Ci* on *Cj* is the composite of all the indirect-effect causalities

C~ imparts to *Cj*. If there is only one causal path from *Ci* to *Cj,* the total effect C~ imparts to *Cj* reduces to the indirect effect.

The indeterminacy can be removed with a numeric weighting scheme. A fuzzy causal algebra, and hence FCMs, bypasses the knowledge acquisition processing tradeoff.

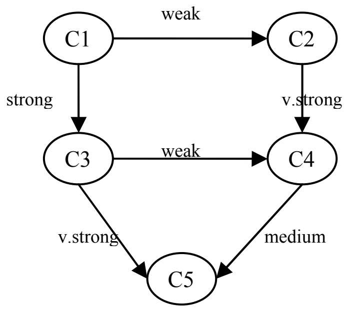

Fig. 1. A cognitive map with fuzzy labels at the edges

A simple fuzzy causal algebra is created by interpreting the indirect effect operator *I* as the minimum operator (min) and the total effect operator *T* as the maximum operator (max) on a partially ordered set P of causal values. Formally, let ~ be a causal concept space, and let e: ~ × ~ P be a fuzzy causal edge function, and assume that there are m-many causal paths from *Ci* to *Cj*: (i, k~ ..... k~, j) for 1 ~< r ~< m. Then let *Ir(Ci, Cj)* denote the indirect effect of concept *Ci* on concept *Cj* via the rth causal path, and let *T(i, Cj)* denote the total effect of *Ci* on *Cj* over all m causal paths. Then

I~(Ci,Cj)=min{e(Cp,Cp+,):(p,p+ 1) ~ (i,k~ . . . . . k,~,j) }

T(Ci,Cj)= max(Ir(Ci,Cj)) , where l <~r<~m

where p and p + 1 are contiguous left-to right path indices.

The indirect effect operation amounts to specifying the weakest causal link in a path and the total effect operation amounts to specifying the strongest of the weakest links. For example, suppose the causal values are given by P = *{none, weak, medium, strong, v.strong}* and the FCM is defined as in Figure 1. There are three causal paths from C1 to C5*:* (C1, C3, C5) , (C1, C3, C4, C5) , (C1, C2, C4, C5).

The three indirect effects of C1 on C5 are:

I1(C1,C5) = min {el3 , e35) = min *{strong, v.strong} = strong*

12 (C1 ,C5) = min{e13,e34,e45}= *weak,*

I3(C1,C5) = min{e12,e24,e45}= *weak.*

Hence, the total effect of C 1 on C5 is:

T(C1,C5) = max {I1,(C1,C5), I2(C1,C5), I3(C1,C5) }

= max {strong, weak, weak} = strong.

In words, C1 can be said to impart *strong* causality to C5.

3 Novel Approach on Designing Augmented Fuzzy Cognitive Maps

There is a necessity to develop a framework extracting fuzzy interconnections among attributes from available data using knowledge extraction techniques and then insert these fuzzy linguistic interconnections to restructure the fuzzy cognitive map model producing a new augmented FCM tool for medical decision making. The framework can incorporate any decision tree algorithm, but for the purpose of this work C4.5 has been chosen as it is a well-known and well-tested decision tree induction algorithm for classification problems [17]. As it has already been stated, the central idea of the proposed method is to combine a fuzzy decision tree to extract the available knowledge of data and to generate fuzzy linguistic weights. The resulted fuzzy relationships among leaf nodes are applied to restructure the FCM model. Among the different fuzzy inference techniques we selected for our approach the Zadeh's union and intersection parameters. The derived FCM model is subsequently trained using an unsupervised learning algorithm to achieve improved decision accuracy. The inference algorithm of FCMs remains the same and only the weight matrix multiplied with previous concept values was changed. Figure 2 illustrates the proposed framework with the corresponding steps and final decision.

The algorithmic approach for the restructure of FCM using fuzzy decision trees is consisting of the following steps:

Step 1: For all the M experts, set credibility weight $b_k = 1$

Step 2: Each of the M experts is asked to suggest and describe each of the N concepts that comprise the FCM.

Step 3: For all the ordered pair of concepts (Ci and Cj) each kth of the M experts is asked to make the following statement (using an if-then rule):

IF the value of concept *Ci* {increases, decreases,is stable} **THEN** causes value of concept *Cj* to {increase, decrease, nothing} **THUS** the influence of concept *Ci* on concept *Cj* is *T(influence)*

Through this step a number of linguistic weights have been assigned by experts.

Step 4: If quantitative data (numeric or symbolic) are available, the approach of using fuzzified crisp decision trees (presented in above section 2.1) is implemented into the data set to derive the available structure of fuzzy decision trees and the fuzzy labels in the branches Di.

Step 5: From the created fuzzy decision trees, a number of causal paths among the branches *i*, connecting leaf nodes *Di* to *Dj*, is determined. These causal paths transferred in FCM model as causal paths interconnecting concepts *Ci* to *Cj*, through a number of direct positive relationships.

Step 6: Using the fuzzy causal algebra, an indirect effect operator I used as the minimum operator (min) on an ordered set P of causal values. The simple fuzzy causal algebra is created by interpreting the indirect effect operator I as the minimum operator (min) on the set P of fuzzy values, corresponding to the above designed causal paths among the FCM concepts. Then the max operator T is applied to the resulted effect operators I, and a new linguistic weight produced among *Ci* and *Cj*. The overall linguistic weight is the sum of the path products. Thus a new linguistic weight is assigned between the concepts *Ci* and *Cj*.

Step 7: IF for one interconnection between the concepts *Ci* and *Cj*, more than 3M/4 different linguistic weights are suggested THEN ask experts to reassign weights for this particular interconnection and go to step 3.

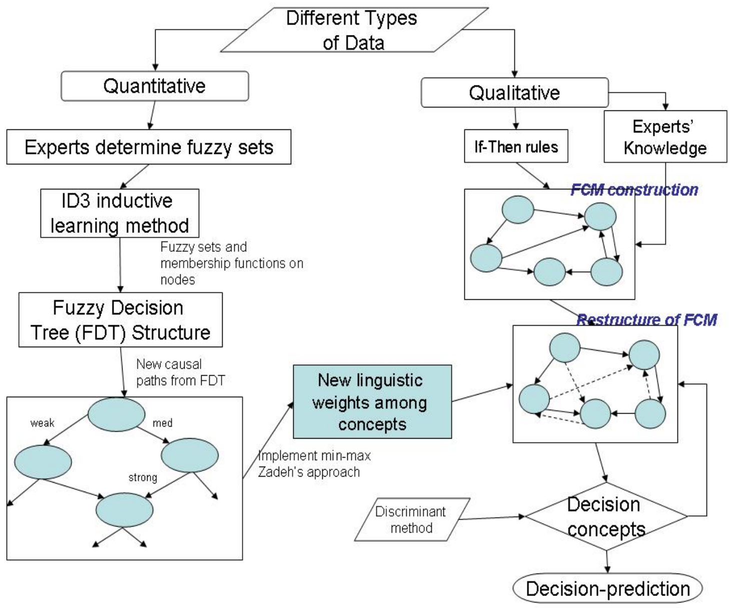

Fig. 2. Framework for constructing augmented FCMs by complementary use of fuzzy decision trees

Step 8: Aggregate all the linguistic weights proposed for every interconnection using the SUM method where the membership function μ suggested by kth expert is multiplied by the corresponding credibility weight b_k. Use the COG defuzzification method to calculate the numerical weight Wij for every interconnection.

Step 9: IF there is an ordered concept pair not examined go to step 3,

ELSE construct the weight matrix W whose are the defuzzified weights Wij

END.

Using the above algorithm, someone could use fuzzy decision trees to pass available knowledge into FCM reconstructed them by paths. Experts construct fuzzy sets and fuzzy membership functions for each problem and these fuzzy sets are used into the fuzzy decision tree algorithm due to compatibility reasons. This happens in the case of FCMs to derive the respective linguistic variables and then make the necessary comparisons.

The causal paths of the leaf nodes used to determine new causal paths in the FCM model. Thus the FCM model was augmented as new direct and indirect relationships among concepts determined.

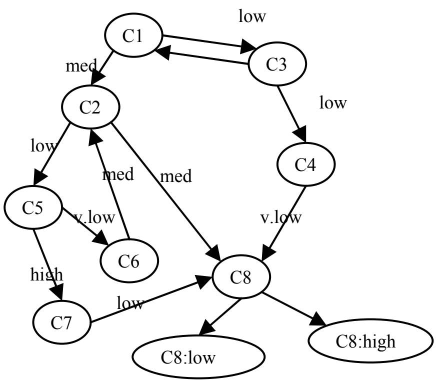

Fig. 3. Example FCM model with initial linguistic labels on interconnections (weights)

4 An illustrative generic example

An illustrative example, of FCM model with eight concepts and eleven interconnections among concepts, with fuzzy labels at the edges of connections, is presented in Figure 3. This initial FCM will be restructured using the proposed methodology and the available knowledge from fuzzified decision trees. Only for implementation reason, we consider that using the fuzzified decision trees on the available data which have been fuzzified, the following tree is produced in Figure 4.

The produced tree has a number of three paths for C1 to C8, two paths for C2 to C8, and one path of each one of the other concepts to C8, thus defining new interconnections and/or update the initial ones of the FCM model.

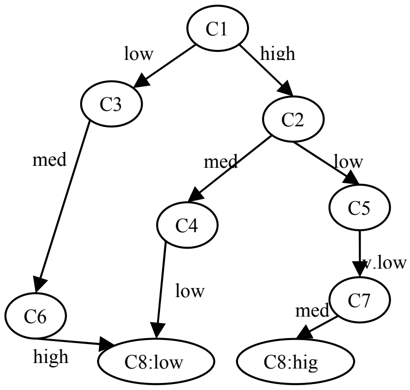

Fig. 4. Example Fuzzy decision tree induced from the data showing membership grades at each branch

Here, the causal effect of C1 to C8 is determined by taking the minimum of the attached labels of the individual paths. Let I1, I2 and I3 denote the effect of C1 to C8 through the paths 1 to 3 respectively, and e_{ij} be the label attached with edge from node i^{th} to node j^{th}. Then, to determine the total effect of C1 to C8, we take the maximum of paths I1 through I3 causal paths.

Path 1 from C1 to C8: c1→c3→ c6→ c8

I1(C1 to C8)=min(low, med, high)=low

Path 2 from C1 to C8:c1→c2→ c5→ c7 →c8

I2(C1 to C8)=min(high, low, v.low, med)=v.low

Path 3 from C1 to C8: c1→c2→ c4 →c8

I3(C1 to C8)=min(high, med, low)=low

Thus total effect of C1 to C8, denoted by T(C1,C8) is computed below:

T(c1,c8)=max{I1,I2,I3}=max{low,v. low,low}=*low*

In words, c1 imparts *low* causality to c8.

To determine the total effect of C2 to C8, we take the maximum of paths I4 through I5.

Path 4 from C2 to C8: c2→ c5→ c7 →c8

I4(C2 to C8)=min(low, v.low, med)=v. low

Path 5 from C2 to C8: c2→ c4 →c8

I5(C2 to C8)=min(med, low)=low

Thus total effect of C2 to C8, denoted by T(C2,C8) is: T(c2,c8)=max{I4,I5}=max{low, v. low}=*low*

In words, c2 imparts *low* causality to c8.

Path 6 from C6 to C8: c6 →c8: I6=high

To determine the total effect of C6 to C8, we take the maximum of path I6.

In words, c6 imparts *high* causality to c8.

Path 7 from C4 to C8: c4 →c8: I7=low

The total effect of C4 to C8 is determined by taking the maximum of path I7.

In words, c4 imparts *low* causality to c8.

To determine the total effect of C5 to C8, we take the maximum of path I8.

Path 8 from C5 to C8: C5→ C7 →C8: I8(C5 to C8)=min(v.low, med)=v.low

Thus total effect of C5 to C8, denoted by T(C5,C8) is computed: T(C5,C8)=max{I8}=v. low

In words, C5 imparts *v.low* causality to C8.

Summarizing, new causal paths describing the interconnections among concepts as well as some interconnections have updated their initial values due to the above paths.

After the implementation of the investigating methodology, the FCM model was restructured and a new FCM model was produced illustrated in Figure 5. Where each branch has fuzzy labels, fuzzy values derived from corresponding fuzzy sets as they have been initially prescribed by experts.

Some of the important points in the proposed approach are:

- Each attribute is represented by a fuzzy set.
- All branches will fire to some degree.
- Multiple input-single output fuzzy if-then rules.
- Each case passes through the tree fires all rules to some degree.

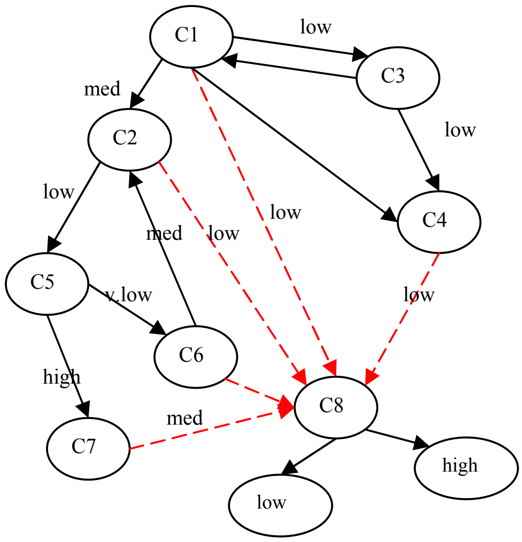

Fig. 5. The new restructured FCM model using the proposed framework

Some of the limitations of the proposed approach are the way of data fuzzification which has to be done automatically from data and without the experts' intervention.

The proposed algorithm and the methodology for constructing FCMs using complementary the fuzzy decision trees as knowledge extraction methods can be used for decision making tasks especially in the medical field. In medical decision making there is enough knowledge hidden in data and the experts-physicians have difficulty to recognize and suggest this knowledge. Thus though the complementary use of fuzzy decision trees as knowledge extraction algorithm and the knowledge representation model of FCMs, an advanced decision making tool with sufficient accuracy and interpretability can be produced. This tool keeps the advantages of FCMs and FDTs coming to a promising task.

5 Conclusion

In this study, it was shown the role of the fuzzy decision tree framework in the design and analysis of augmented fuzzy cognitive maps. We discussed the role of the fuzzy decision tree in the determination of fuzzy linguistic weights and causal paths of FCM. It was stressed that this technique takes advantage of the available experimental data. We proposed a detailed design algorithm producing augmented FCMs that offer a comprehensive interpretation of the cognitive model. In particular, the formalism of fuzzified crisp decision trees helped us come up with endowing the medical decision making results with meaningful models.

Acknowledgment

The research was supported in part by the European Commission's Seventh Framework Information Society Technologies (IST) Programme, Unit ICT for Health, project DEBUGIT (no. 217139).

References

1. J. Quinlan, Induction of Decision Trees, Machine Learning, vol. 1, Kluwer Academic Press, Dordrecht, 1986 pp. 81–106.
2. S. Sestino, T. Dillon, Using single-layered neural networks for the extraction of conjunctive rules and hierarchical classifications, J. Appl. Intell. 1 (1991) 157–173.
3. L. Sison, E. Chong, Fuzzy modeling by induction and pruning of decision trees, IEEE Symposium on Intelligent Control U.S.A., 1994, pp. 166–171.
4. S.Mitra, K.M. Konwar, K.P. Sankar, "Fuzzy decision tree, linguistic rules and fuzzy knowledge-based network: generation and evaluation", IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev. 32 (4) (2002) 328–339.
5. M. Umano, H. Okamoto, I. Hatono, H. Tamura, Generation of fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis by gas in oil, Japan–U.S.A. Symposium, 1994, pp. 1445–1450.
6. C.W. Olaru, A complete fuzzy decision tree technique, Fuzzy Sets and Systems 138 (2003) 221–254.
7. W. Pedrycz, A. Sosnowski, Designing decision trees with the use of fuzzy granulation, IEEE Trans. Syst. Man Cybern. A 30, (2000) 151–159.
8. K. Crockett, Z. Bandar, D. Mclean, J. O'Shea, On constructing a fuzzy inference framework using crisp decision trees, Fuzzy Sets and Systems 157 (2006) 2809 – 2832.
9. H. Ishibuchi, K. Nozaki, N. Yamamoto, N. Tanaka, Selecting fuzzy if–then rules for classification problems using genetic algorithms, IEEE Trans. Fuzzy Systems 3 (3) (1995) 260–270.
10. C.Z. Janikow, Fuzzy decision trees: issues and methods, IEEE Trans. Systems Man and Cybernetics, 28 (1) (1998) 1–14.
11. C.Z. Janikow, Fuzzy partitioning with FID3.1, Proceedings of the 18th International Conference of the North American Fuzzy Information Society, 1999, pp. 467–471.
12. Y. Yuan, M.J. Shaw, Induction of fuzzy decision trees, Fuzzy Sets Systems 69 (1995) 125–139.
13. R. Weber, Fuzzy ID3: a class of methods for automatic knowledge acquisition, Second International Conference on Fuzzy Logic and Neural Networks, Iizuka, Japan, 1992, pp. 265–268.
14. B. Kosko, 1992. Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice-Hall, New Jersey.
15. Kosko, B.: Fuzzy Cognitive Maps, International Journal of Man-Machine Studies, vol. 24, pp. 65-75, 1986.
16. Papageorgiou E.I., Stylios C.D., Groumpos P.: An Integrated Two-Level Hierarchical Decision Making System based on Fuzzy Cognitive Maps (FCMs), IEEE Trans Biomed Engin, 2003; 50(12): 1326-1339.
17. J.R. Quinlan, Is C5.0 better than C4.5, Available: http://www.rulequest.com/see5-comparison.html, 2002.
18. I. Hayashi, T. Maeda, A. Bastian, L.C. Jain, Generation of fuzzy decision trees by fuzzy ID3 with adjusting mechanism of and/or operators, in: Proc. Int. Conf. Fuzzy Syst., 1998, 681– 685.