AN INTELLIGENT SYSTEM FOR SOLO TAXONOMY

John Vrettaros (1,2), George Vouros (2), Athanasios Drigas (1)

1) National Centre of Scientific Research "Demokritos"- Institute of Informatics and Telecommunications - Net Media Lab, Agia Paraskevi, Athens, Greece
2) University of Aegean, Information and Communication Systems Engineering, Karlovassi Samos, Greece

{jvr, dr}@imm.demokritos.gr
http://imm.demokritos.gr/

Abstract: The modeling of diagnostic systems of taxonomies using fuzzy logic is presented in this paper. Specifically the taxonomies system solo is studied, which that can be applied in a wide range of fields of diagnostic science. The intelligent system that is developed based on the presented modeling can make easier the use of diagnostic systems in education since the test correction is extremely hard and demands experts that are not always available. Additionally, the rate of the extraction of results is a reason for using and distributing such tools (diagnostic systems) in the educational process. It is very useful for e-learning systems [1], [2], and distance diagnostics systems.

Key words: fuzzy system, solo, taxonomy, diagnostics system, distance education, e-learning system

1.　　INTRODUCTION

An intelligent system is based on an extended quantity of knowledge relevant to an area of problems. This knowledge is organized under the form of a set of rules, which allow inferencing of the system from the available data. This "knowledge-based" methodology that have been used for solving problems and generally designing systems, has constituted an evolutionary change in Artificial Intelligence, since it substituted the traditional form of a

program (data + algorithm = program) with a new architecture. This new architecture has as its core a knowledge base and an inference engine and is of the form:

Knowledge + Inferencing = System   (1)

The specific problem that has to be solved is the construction of an intelligent system, which will be able to evaluate and categorize students in different levels of knowledge based on some information that will be extracted from their answers. The results are based on a research that was made on some High School students and was about the broader field of Mathematics. The problems of categorizing student into different levels of knowledge, the study of transition in between knowledge levels, as well as the study of the semantic change, as it is mentioned, that takes place when students stop using an naïve (erroneous) model and start using a scientific (correct) model, are three of the most important problems in Cognitive Science. Many researchers have proposed methodologies for knowledge acquisition into different fields of science (Mathematics, physics, etc.) under the aid of computational systems and Artificial Intelligence models. The methods of computational intelligence present great interest from theoretical point of view, since they cope with complexity and uncertainty which are two of the most important problems of system theory that are strongly related to reality.[3]In the specific application the analysis begins with the processing of the answers from correctly formed and selected questionnaires, which are filled up by students. By this analysis some information is extracted that leads the categorization into levels in five different theme sections. Each section consists of four questions, and each question corresponds to one of the following knowledge levels: Single-Structural (S), Multi-Structural (M), Correlative (C), and Abstractive (A). [4], [5].

Table 1. Different fields of science

	(M)	(C)	(A)
ARITHMETIC	Quest. 1	Quest. 2	Quest. 3
ALGEBRA	Quest. 4	Quest. 5	Quest. 6
SPACE PERCEPTION	Quest. 7	Quest .8	Quest .9
APPLICATION	Quest. 10	Quest. 1	Quest. 12
PROBABILITIES	Quest. 13	Quest. 14	Quest. 15

At this point it should be mentioned that the question that corresponds to the abstractive level couldn't be answered by students of this age. Consequently, it can be said that each theme section has three questions. In addition, if none of the three questions of a theme section is answered by a student, the student is categorized to the pro - structural (P) level. The questionnaire analysis is illustrated in the following figure (fig.1).

LEVELS

		PS	SS	MS	R	S
	Arithmetics		Q.1	Q.2	Q.3	
	Algebra		Q.4	Q.5	Q.6	
Thematic categories	Space Conception		Q.7	Q.8	Q.9	
	Applications		Q.10	Q.11	Q.12	
	Probabilities and Data		Q.13	Q.14	Q.15	

Figure 1. Questionnaire Analysis.

2. SYSTEMS OF STUDENT LAYER SPECIFICATION TO THEME SECTION: CONTROVERSIAL ANSWERS

For the evaluation of these answers the following factors (that correspond to the factors that the teachers consider when they evaluate such controversial cases): 1) Difficulty of the specific theme that obviously affects its grading. 2) The number of the blank answers, meaning the number of unanswered questions for each student. This factor is considered since it affects the student's evaluation. For instance, let us consider the case where we want to grade a controversial answer (e.g. Answer 4 → WRONG, Answer 5 → CORRECT, Answer 6 → CORRECT,) and the student has a great number of unanswered questions. This means that the student probably doesn't answer randomly, but he/she answers after serious consideration of the question. We conclude that is most probable that the incorrect answer at question 4 is a careless mistake, since the correct answers at questions five and six (which are obviously more difficult than question 4) are not given by chance. Consequently the student can be categorized to the Correlative level for the corresponding theme section 3) The level of the student, meaning the general presentation of the student [6].

In general, it can be said that the selection of the layer for the controversial cases differs from one student to another. It is affected from the student's answers, the number of the questions that he/she left unanswered, and the level of the question. For the modelling of the controversial cases there have been designed and developed two fuzzy systems, which are analyzed in the following section. [7],[8].

2.1   Rigidity grading specification sub-system

The systems consist of three inputs and one output. The inputs are the factors that affect the grading of each controversial answer: the number of the unanswered questions, the level of the question and the level of the student. The output is just one: the rigidity (fig.2):
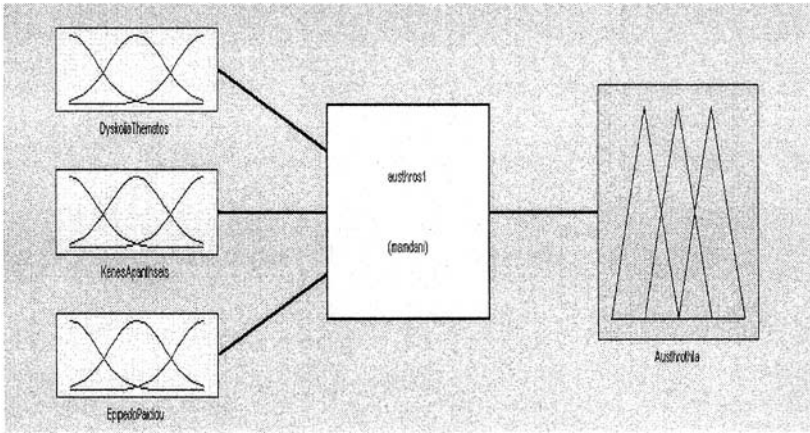

Figure 2. Inputs and Outputs of the System.

In this specific case the grades of every input are between two rates. The Difficulty of the Subject and the Number of the Blank Answers takes rates between 0 and 100, and the Child's Level between 0 and 3. The Difficulty of the Subject is calculated from the answers of the rest of the students. The X axis is normalized and takes rates obviously from 0 to 100. We can set a taxonomy on the definition level of the Difficulty of the Subject, for example we can say that if a rate of Difficulty belongs to the aggregation (0,30) then it is large, if it belongs to (30,65) then it is medium and finally if it belongs to (65,100) then it is small. This specific way of classic taxonomy inputs in principle a big uncertainty in some sections, e.g. close to 30, 65 and 100. That is if the rate of Difficulty is equal to 29 then the Difficulty consider

small but if it is equal to 30 then it is considered medium. To avoid this problem we define a fuzzy taxonomy (one per input) in the definition levels of every input A1, A2 and A3.

Every fuzzy taxonomy is class 3. A fuzzy taxonomy B, class 3, we define also in the output definition level, which as we mentioned is [9]. The fuzzy taxonomies A1, A2, A3 and B are linguistic images of the definitions levels therefore their elements are linguistic terms of the form "SMALL", "LARGE", "MEDIUM".

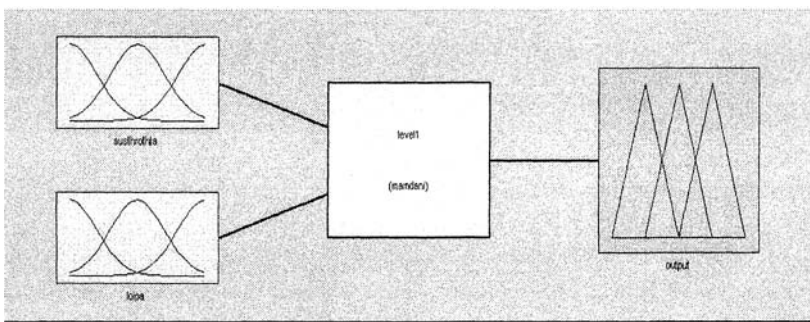## 2.2   Student level relevant to theme section specification sub-system

Figure 3. Inputs and outputs of the system

The second system specifies the level of the student at each theme section. The system has two inputs and one output. The first input is the rigidity and the second input is the three answers to that theme. The output is the number from 0 to 3 that corresponds to one of the four levels of knowledge (pro-structural, single-structural, multi-structural and correlative) for each theme section. The result for the controversial cases can be a decimal number. At fig. 3 the inputs and output of the system can be seen. [10]

The output takes rates from 0 to 3. 0-1 corresponds to single-structural, 1-2 in multi-structural and 2-3 in correlative.

The rules that join the inputs and output are the following:

1. If rigidity is "LARGE" then the Level is "SMALL".
2. If rigidity is "MEDIUM" then the Level is "MEDIUM".
3. If rigidity is "SMALL" then the Level is "LARGE".
4. If the Answers are "FEW" then the Level is "SMALL".
5. If the Answers are "ENOUGH" then the Level is "MEDIUM".

6. If the Answers are "PLENTY" then the Level is "LARGE".
An analytical representation of the rules sows in the Schematics. (fig 4)
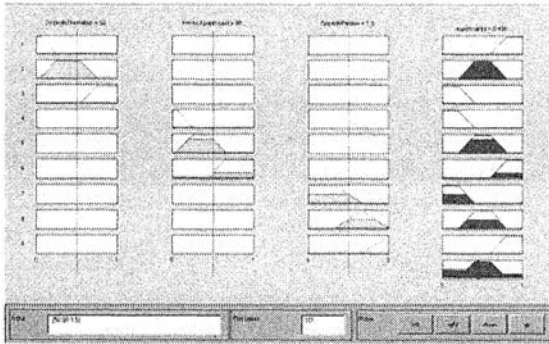

Figure 4.

3.   SPECIFICATION OF THE FINAL LEVEL

Up to this point there have been estimated the levels of knowledge for the students in five theme sections. Based on these levels the estimation of the final level will be implemented. The final level is a number from 0 to 3 that corresponds to one of the four levels of knowledge. At the previous sections we described the procedure of level estimation based on theme sections. The following procedure examines the students' answers based on the levels of knowledge and not the theme sections. This means that a grade of confidence should be estimated for each level. The grades of confidence will correspond to the percentage that each examined student belongs at each level. The grade of confidence is a number in the range from 0 to1.

The grades of confidence are three: one for the single-structural level, one for the multi-structural level, and one for the correlative level. For the pro-structural level no confidence number is estimated since its value is always set to 1, because there are no questions or answers and it is also the lowest level. Consequently the grade of confidence cannot be less than 1. Next, we estimate the final level by averaging the three available confidence grades. The averaging is done in accordance with the relevant level. Setting the confidence grade to 1 for the pro-structural level, 2 for the multi-structural level and 3 for the correlative level we then have:

$$\varepsilon = \frac{1C_1 + 2C_2 + 3C_3}{C_1 + C_2 + C_3} \tag{2}$$

Where ε is the final level. ε can also be a decimal number.

4.   CASE STUDY: THE SOLO PROGRAM

The SOLO program is the interface that contains a powerful Neural Network engine, which basically manages the data of the class and the students. It is very simple and easy to use, providing help support.

Below are stated some selections provided by the interface:

New Database:

This function provides to the user the possibility to create a new database. The window contains combo boxes and textboxes where the user inputs the variables. Using the add button the user inputs a new record to the database. With the delete button the user can delete the present record. By pressing the refresh button the user can refresh the database (for multi-user environment only). With the update button the user can post the database for the changes done, and with the exit button the user closes the window and returns to the main window of the application. By pressing the SaveDB the user saves the database to the hard-disk.
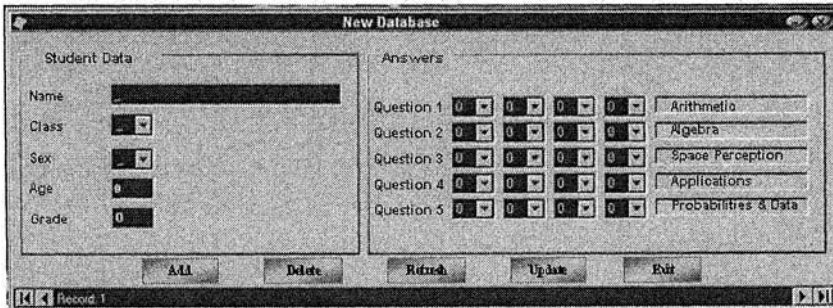

Figure 4.1. New Database Window

Open Database: This selection is used to load a database which is added to the main application's window, in the Combo box under the title Available Databases. By clicking there the user selects this database which is loaded in the form.
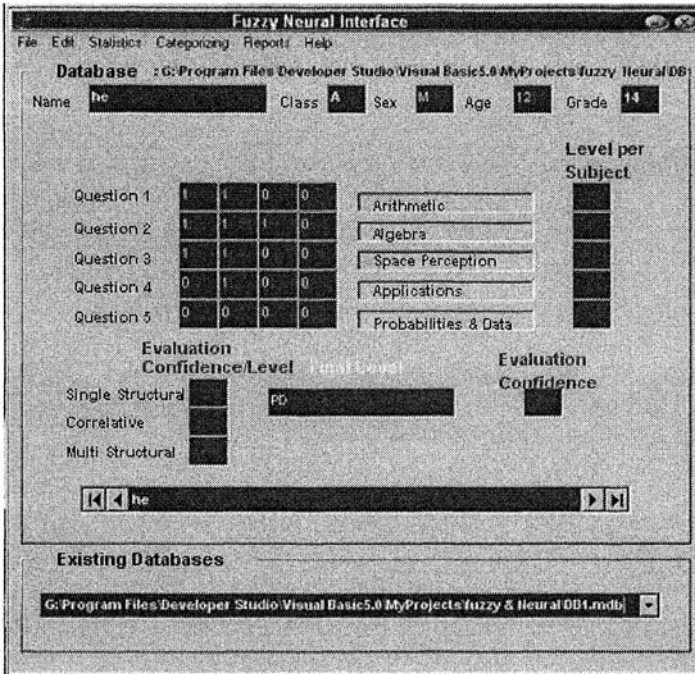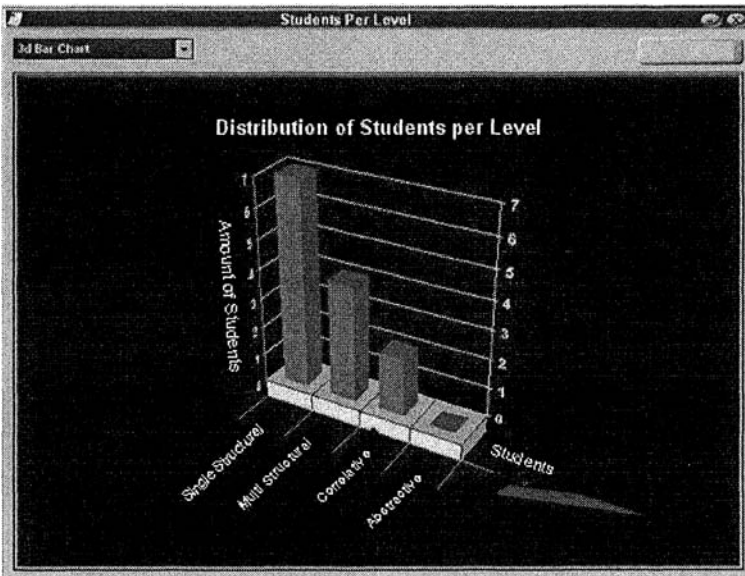
Figure 4.2. Open Database Window

Figure 4.3. Students Per Level Graph

With this last selection above, the user is provided in 2D or 3D graph the distribution on the students depending on the level the students are.

5.   CONCLUSION

The developed system was applied on 100 high school and senior high school students, and it was tested on mathematics. The correction results obtained by the system were compared to the results obtained by the cognitive science expert. The system's results were found to be very close to the expert results, as it can be seen on the following table (fig. 5).[11]

Concluding, we can say that the diagnostic tools are trustworthy tools for the educators' cooperation and contribution.
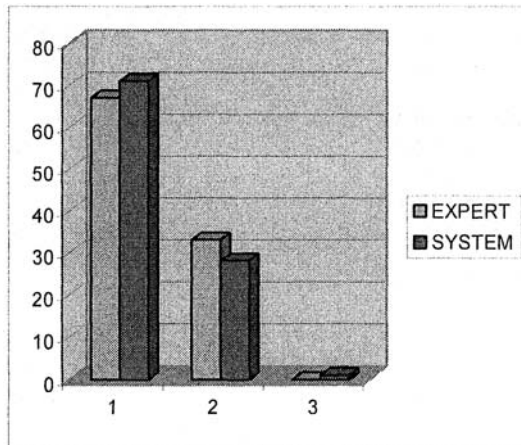

Figure 5. System - Expert Results Comparison. (blue color corresponds to the expert and red color corresponds to system).

REFERENCES

1.   A.S.Drigas, J.Vrettaros, D.Kouremenos, "Teleeducation and e-learning services for teaching English as a second language to Deaf People, whose first language is the Sign Language", WSEAS transactions on Information Science and Applications, Issue 3, Volume 1, September 2004.

2. A.S.Drigas, J.Vrettaros, L. Stavrou, D.Kouremenos, "E-learning Environment for Deaf people in the E-Commerce and New Technologies Sector", 6th WSEAS International Conference on E-Activities, Rethymno, October 20, 2004.

3. Maeda, S. and Murakami, S. (1993) The Use of a Fuzzy Decision-Making Method in a Large-Scale Computer System Choice Problem, Fuzzy Sets and Systems, 54, 235-249.

4. Biggs, J. B. and Collis, K. F. "Evaluating the Quality of Learning: the SOLO Taxonomy", New York: Academic Press, 1982.

5. Biggs, J. B. and Collis, K. F. Multimodal learning and the quality of intelligent behavior. In H. H. Rowe (Ed.) Intelligence: Reconceptualization and Measurement. Hillsdale, N. J.: Lawrence Erb.freemanm Associates and Hawthorn, Vic.: ACER, pp. 57-76, 1991.

6. Imrie, B.W. (1995) Assessment for Learning: Quality and Taxonomies, Assessment and Evaluation in Higher Education, 20, 2, 175-189.

7. J.Vrettaros, M.Grigoriadou, "Design of an hybrid architecture for fuzzy connectionist expert system and its application to approximate student model". CATE 96, The first international conference on computers and advanced technologies in education, Cairo, 1996.

8. Earl Cox, "The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, & Maintaining Fuzzy Systems", 1999.

9. Gogus, O. and Boucher, T.O. (1997) A Consistency Test for Rational Weights in Multi-Criterion Decision Analysis with Fuzzy Pairwise Comparisons, Fuzzy Sets and Systems, 86, 129-138.

10. Zadeh, L.F. (1997) Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic, Fuzzy Sets and Systems, 90, 111-127.

11. Drigas A, Kouremenos S, Vrettos S, Vrettaros J, Kouremenos D (Feb 2004), An expert system for job matching of the unemployed, Pergamon-Elsevier Science LTD, Oxford, IDS Number:763WW, ISSN: 0957-4174