Selecting Pedagogical Protocols Using SOM

F. Salgueiro, Z. Cataldi, P. Britos, E. Sierra and R. García-Martínez
Intelligent Systems Laboratory. School of Engineering. University of Buenos Aires
Educational Informatics Laboratory. School of Engineering. University of Buenos Aires.
Software & Knowledge Engineering Center. Graduate School. Buenos Aires Institute of Technology
Intelligent Systems in Engineering Group. School of Engineering. University of Comahue
rgm@itba.edu.ar
(Paper received on August 11, 2006, accepted on September 26, 2006)

Abstract. During the first courses of the Computer Engineering undergraduate Program at the University of Buenos Aires, the number of human tutors in Programming Area is usually not enough: the students/tutors ratio is very high and there is a great heterogeneity in the acquired knowledge and background of students. The main idea behind this paper is to describe a system that could emulate a human tutor. Thus, the tutor will be able to provide student with a high level of flexibility for the selection of the most adequate tutorial type. This could be a feasible solution to the stated problem. But a tutorial system should not only emulate the human tutor but also it should be designed from an epistemological perspective concerning what teaching Programming really means. This is specially important in an Engineering course due to the required profile and identity of the future engineer. The stated solution implements a series of artificial neural networks to determine if there is a relationship between the given initial population of students learning preferences and the different tutoring types. A series of experiences were carried out in order to validate the current model.

1. Introduction

The main objective of the tutor module of an Intelligent Tutoring System is to present the new knowledge to the student in the best way possible. To achieve this, our research group [Salgueiro *et al.*, 2005; Costa *et al.*, 2005] has designed a series of sub modules and interfaces to avoid the normal overlap that usually appears in the modules of an Intelligent Tutoring System. In the tutor module, the main sub module contains the pedagogical protocols, which is made up of two basic components: the profile analyzer and the database of pedagogical protocols available in the system. The system has a database of pedagogical protocols. Its use will be subordinated to the availability of the contents in the knowledge module, although the lesson always can be generated for some of the available protocols. In order to collect data about the way in which each student learns, lists of learning styles will be used as well as tools for data collection. It has been determined the validity and trustworthiness of this instrument through its application by various researchers from the date of its creation [Felder, 1988; Figueroa, 2004] up to now. Starting from data provided by each student, his or her learning style will be determined. Afterwards, in a second step, the learning style will be linked to the

© H. Sossa and R. Barrón (Eds.)
Special Issue in Neural Networks and Associative Memories
Research in Computing Science 21, 2006, pp. 205-214

pedagogical protocol. The Felder list [Felder 1988] is a validated tool that allows obtaining solid data from students. After giving a questionnaire to the students, we will try to get data records on different sets by using the tools provided by Artificial Intelligence (AI). such as Neural Networks (NN) in order to obtain a relationship between the preferences of the students and pedagogical protocols. From a statistically significant sample of students for which the lists of complete learning styles had been taken. will try to see if the learning styles can be grouped according to the education techniques or pedagogical protocols. This will allow correlating the preference of the student with the most suitable pedagogical protocol in the system. As the selection of the pedagogical protocol is one of the elements to determine, it is desired to group the students in families with common characteristics.

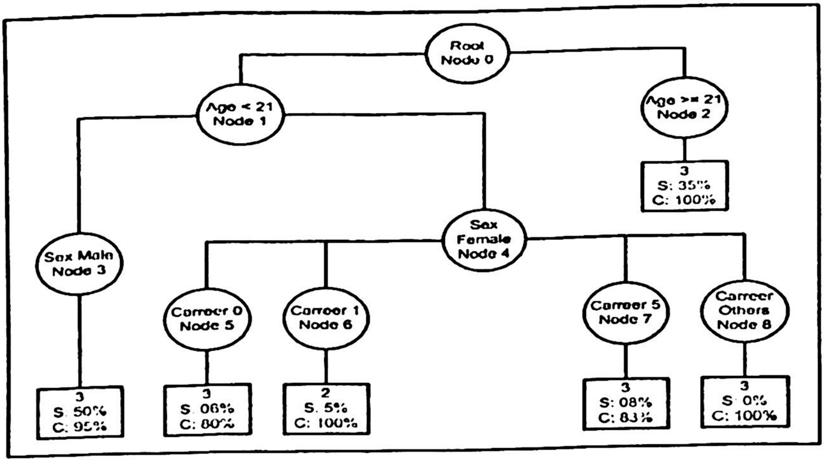

Fig. 1. Tree generated by C4.5 algorithm using SOM output as C4.5 Input.

This can be achieved using the Self Organizing Maps (SOM) neural networks (also known with the name Kohonen [2001] maps) that make a "determined clustering" or grouping action according to common characteristic of the original set of individuals. Once obtained the resulting groups of SOM network an induction algorithm will be used to find the rules that characterize each one of these groups. In this case the algorithms to be used will belong to the family of Top-Down Induction Trees (TDIT) algorithms. Although several algorithms exist that make these functions, a very complete one is Quinlan's C4.5 [Quinlan, 1993], an extension of algorithm ID3 (Induction Decision Trees) also proposed by Quinlan [Quinlan, 1987]. Its objective is to generate a decision tree and the inference rules that characterize this tree. In this particular case. the C4.5 will take as input the data of the students already clustered by SOM and the output will be the rules describing each cluster.

Once obtaining the smallest amount of rules by pruning the tree to avoid overfitting, we move to another stage of the analysis in which, by means of an inference process, we found the relation between the SOM clusters and the pedagogical protocols available. In order to carry out the inference, additional data concerning the performance of students with different protocols of education in the courses under study were used. In Figure 2 the scheme of the solution can be seen: it represents the process of selection in a global form, where we start from a student population for which we have their preferences concerning learning styles through the lists of Felder. We form groups of students by using SOM. A table is generated using the previously classified students, using all the attributes that describe them and the cluster predicted by SOM. Then C4.5 algorithm is used to generate the rules that best describe each cluster, relating a particular cluster not only with all its attributes, as in the table of classified students, but also with a set of rules.

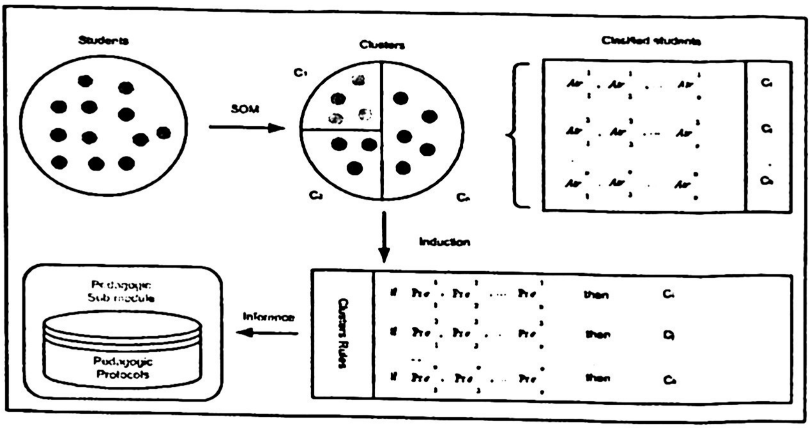

Fig. 2. Basic scheme of the solution.

2. Inference of the pedagogic protocol

In this stage we try to relate the groups generated by SOM to the pedagogical protocols by training a Backpropagation type neural network. In order to find the relationship between the learning style and the pedagogical protocol that best fits each group, the basic protocols described by Perkins [Perkins, 1995] in Theory One were used: [a] The *didactic or skillful instruction*: It satisfies a need that arises within the framework of the instruction in order to expand the repertoire of student knowledge [b] *The training*: It satisfies the need to make sure the student will have an effective practice, [c] *Socratic education*: This type of instruction is applied to provide educational aid to the student to include/understand certain concepts by himself and to give him the opportunity to investigate and learn how to do it, all by himself. Therefore the research is oriented towards the search of the relationship between the predilection of students concerning learning style and the pedagogical protocols used by the human tutors (professors). In order to accomplish this. the grades of the partial evaluations are used. Two courses (A and B) will be taken belonging to the Basic area of Programming. The only fundamental difference between both of them was centered in the form of education, that is to say, in the pedagogical protocol used in the classes. From this frame of reference. two courses were evaluated according to the control variables raised by García [1995]. The variables raised for the reference courses are the following ones: [a] Similar contents of the courses. [b] Similar schedules. [c] Similar bibliography used for references. [c] Random entrance of the students, without preference defined to some course. [d] Similar previous formation of the assistants and instructors in charge of practical works. [e] Similar didactic tools an [f] Way in which the class is dictated, where each one of the tutors presents the classes based on the pedagogical protocol that turns out more natural to carry out to him. The possible options are defined in Theory One and that are analyzed in this investigation, independently of the needs or preferences of individualized students. From the analysis made by García, it is observed for this study where the only variable that changes is the one denominated "way of class dictation", that is to say, the pedagogical protocol used for each course. In order to carry out the inference, the following hypothesis will be considered: It is possible to relate the learning styles to the pedagogical protocols. Two more particular hypotheses arise from this main one: (a) The composition of styles of learning (needs and preferences of students) of each student determine the style of education (or pedagogical protocol) (b) Those students for whom the education style does not agree with their preference, show difficulties in the approval of the taught subjects. From the second hypothesis it is given off that for the approved students, the protocol preferred by most of them will have to be the one that agrees with the used one in class by the tutor, whereas for the failed ones, the protocol must be inverted for most of them. In order to validate this affirmation a network of Backpropagation type was trained with the following characteristics: [1] the approved students of the course with professor who dictates in Socratic style and the most of the failed ones of the course with professor who dictates in skillful way and the network is trained considering as output the Socratic protocol. [2] the approved students of the course with professor who dictates in skillful style plus the failed ones of the course with professor who dictates in Socratic way and the network is trained considering the output exit as skillful protocol. In order to suppress the "data noise" the training is carried out in the previously indicated way due to the fact that the groups that are outside the analysis contribute to increase the data noise (those students that approved with any protocol which will be considered "indifferent" and those that failed by lack of study or other reasons) and hope that the error of the tool is minor than the percentage of elements that are outside the analysis. Therefore, each generated cluster will be analyzed in the following way:

– approved students	Correct protocol	→ majority class
	Indifferent	→ minority class
– failed students	Inverted protocol	→ majority class
	Lack of study	→ minority class

Now we look to relate the forms of education and the learning styles. In Figure 3 it is observed. taking as base an example on where two pedagogical protocols exist and two preferences in the set of students, that the students whose preference agrees with the form or style of education do not have problems to approve. Two subgroups (in red) of students exist whose preference does not agree with the education form and are those that fail, since they are "bad located".

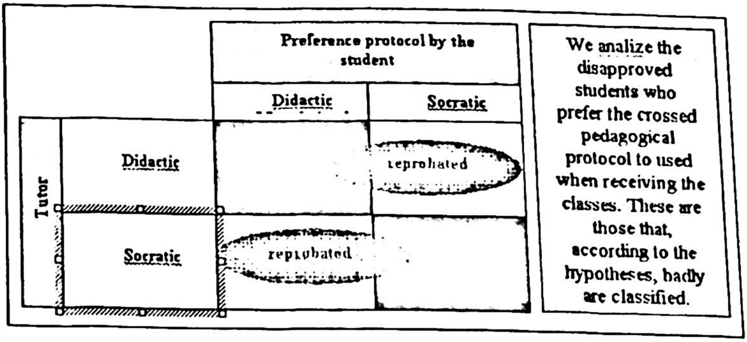

Fig. 3. Inference general scheme

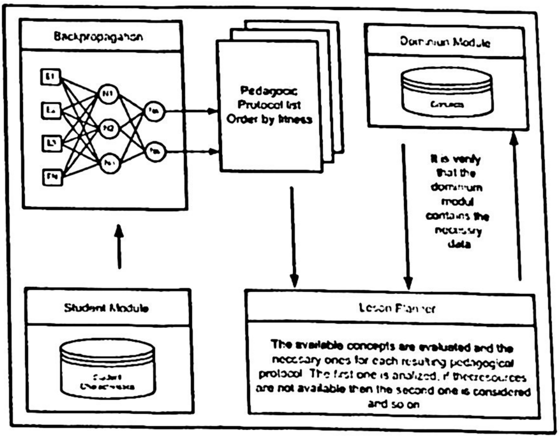

Fig. 4. Modified solution structure

Following the hypothesis: failed students who do not belong to the main cluster predicted by SOM must have a different preference concerning a pedagogical protocol (inverted in this case) from the one the professor used when they attended the classes. On the other hand, Figure 3 displays the idea of the hypothesis, where the failed students, who do not belong to the main cluster, must have a preference of pedagogical protocol different from the one received in the classes. Therefore, if the data provided by the educational system is analyzed in both courses with respect to the categorizations made by system SOM, it is possible to be obtain which is the percentage of students who would be badly located in the courses, and this is demonstrated through the bad results reprobated obtained in the evaluations. So that the obtained results are satisfactory, the Backpropagation neural network must have a classification error smaller than the percentage of elements that were left outside of the analysis. This way this tool will be useful for the classification of the preferences of education (pedagogical protocol) of students according to their styles of learning. This way, a submodule gives a ranking of best suitable pedagogical protocol, in descendent order with respect to the preference for the selected student. Soon, the only thing that is required is to cross all the pedagogical protocols including in the system. The basic scheme of the solution can be seen in Figure 4 in terms of the student groups and the selection of possible protocols, where the Backpropagation network provides a ranking of aptitude of the pedagogical protocols available in the system.

3. Experimental results

The experimental results will allow to validate the proposed solution and will establish the steps in order to repeat the experiences with another group of students or to apply the method in other knowledge domains different from the Basic Programming. The fundamental steps for the experimental design are described in Table 1 where it is started with the taking of data of the students (to lists of learning styles) and it is used them like entrance for the training of a neural network of SOM type to generate the different groups. Soon the rules identify what describes these groups by means of the C4.5 algorithm.

Table 1. Steps for the experimental design

Step	Input	Action	Output
1	Data recollection from students	Use Felder tool on students	Result of the Felder tool.
2	Felder tool result	SOM Training	Students Clusters
3	Cluster + Felder tool results.	Use C4.5 algorithm	Rules describing each generated cluster and the corresponding decision tree.
4	Academic performance	Academic data grid	Academic grid
5	Result of the Felder tool + Academic grid + Clusters	Analysis of the cluster and determination of reprobated students.	Reprobated Student List for each cluster.
6	Result of the Felder tool.	Backpropagation training	Determination of the training error and the data out of analysis. Find the relation between learning style and pedagogic protocol.

4. Validation of the population size

In order to determine the minimum number of elements in the sample we use Hernández Sampieri [Hernández Sampieri, 2001] for the calculation. An initial of 800 student's population has S2 variance of the sample of *n* student that can determine in terms of the probability p where:

$V = 0.03$ (5.1)

$V^2 = (0.03)^2 = 0.0009$ (5.2)

The number of samples without any adjustments will be:

$n = (S^2 / {}^2) = 0.09 / 0.0009 = 100$ *students.* (5.3)

Adjusting in order of the real *N* population:

$n = (n' / (1 - n' / N)) = 100 / (1 - 100 / 800) = 89$ Students. (5.4)

The generalization error is below 3%, with which it is possible to say that the sample size is representative for all the students of the courses. Now we are ready to train the SOM network using the data collected from the Felder tool. Most of the parameters of SOM network arise through an iterative process, where the network trains and the results are analyzed. If the results are satisfactory (that is to say, the training error is then sufficiently small), the parameters are modified slightly to try to improve them still more. If the results are little satisfactory they are compared with previous set and they are modified in a higher value.

Table 2. Parameters used for SOM with which the data of the students were classified.

Parameter	Value
Observations	121
Variables	47
Artificial Neurons	10
Cicles	1000
Aleatority	Yes
Learning Parameter	
Initial	0.9
Final	0.1
Decay Function	Exp
Gaussian boundary parameter	
Initial	99,0%
Final	01,0%
Decay Function	Exp

It is possible to indicate that for obtaining the final values for training the neural network they have been proven more than one hundred combinations, obtaining the best results with the list of parameters that is observed on Table 2.

The amount of clusters: If the amount of clusters is very high, it may occur that it does not exists a correlation between so many pedagogical protocols and clusters, since it is started from the hypothesis that 3 pedagogical protocols exist (the proposed by Theory One). The number of clusters that is expected to get will be annotated between two and three. Summary of the results that the training of the SOM networks gives is in Table 3, where the elements of each cluster generated are totalized and the respective percentage are indicated.

Table 3. Summary of resulting elements after applying SOM to the input data.

	Cluster 1	Cluster 2
Data with all the attributes	6 (5.00%)	114 (95.00%)

The result is within the awaited amount of clusters and therefore the experimental data, they agree in the amount of clusters generated. As all the data are categorical, the generated rules will be equalities and it will not be found any range for them (for example: the continuous data). In order to find the attributes with greater gain of information, it is required to use the first N passages of the C4.5 Algorithm. In this case, the first nine were taken and the rules appear in Table 4.

Table 4. Resulting rules to cross the tree generated by the C4.5 Algorithm

Rule	Antecedent	Consequent
Rule 1	If "Normally they consider me: Extrovert"	Then Cluster 2
Rule 2	If "Normally they don't consider me Reserved neither Extroverted"	Then Cluster 1
Rule 3	If "I Remember easily: Something that I have thought much"	Then Cluster 2
Rule 4	If "I don't remember easily something than I have thought much or something that I did"	Then Cluster 1
Rule 5	If "I learn: To a normal rate, methodically. If I make an effort, it profit".	Then Cluster 2
Rule 6	If "I do not learn to a normal rate, not methodically neither disordered"	Then Cluster 1
Rule 7	If "When I think about which I did yesterday, most of the times I think about: Images"	Then Cluster 2
Rule 8	If "When I think about which I did yesterday, most of the times I think about: Words"	Then Cluster 2
Rule 9	If "When I think about which I did yesterday, most of the times I don't think about words neither images"	Then Cluster 1

Oates [Oates *et al.*, 1997] among others has analyzed several algorithms of "pruning" to adjust the size of the rules generated from a great number of observations. Oates has found that as the size of the initial observations is increased, the size of the rules increases in linear form. This increase in the amount of rules antecedents does not significantly increase the precision in the classification of the rules. Continuing this way we get a result, for this case in individual, as proposed in the works of Quinlan [Quinlan, 1987] and Oates [Oates *et al.*, 1997] previously mentioned, offered the additional advantage when using the second tree (with less levels and minor amount of nodes), the Intelligent Tutorial System requires minor amount of information to select the pedagogical protocol of the student and with easier access information (it is simpler to know the answers of some key questions in the list that the answers to all the questionnaire). Training this way it is managed to suppress the "noise" that contributes the groups that are outside the analysis. In Table 5 the results of the students discriminated by courses can be seen, counting total students, students failed classified as belonging to the cluster in opposition to the one of the majority and the percentage that relates the failed and approved students that in addition are bad classified. For this experience the network of the Backpropagation type trained and a ranking (scale) of pedagogical protocols most adapted for a particular situation was obtained, in order to give flexibility to the module that stores the contents.

Table 5. Summary of percentage obtained for the analysis of students, discriminated by courses.

Observed Characteristic	Course A	Course B
Total of Students (For this study)	47	53
Students who reprobated the partial evaluation and were in a course with different pedagogical protocol	30	0
Students who approved the partial evaluation were in a course with different pedagogical protocol (inverted)	10	33
Approved students (no mattering about the protocol)	7	20
Reprobated students respect to the approved ones, within the subgroup of badly classified	75%	0%

For the training of the Backpropagation network 67% of the data (qualifications) were used randomly whereas 33% of the remaining data were used to validate the generated model. After more than 100 training of 1000 cycles each one, where it was carried out in order to diminish the error in the resulting network, it was reached the conclusion that the optimal values for the parameters of the network are those that are seen on Table 6.

Table 6. Neural Net Training Results.

Characteristic	Value
% Error (Training group)	3.75%
% Error (Validation group)	2.00%
Network characteristics	
Input neuron	13
First hidden layer neurons	20
Second hidden layer neurons	20
Output neurons	2

This training is valid since the error of the tool (3.75% for the set of training and 2.00 % for the validation set) is minor than the error of the elements that were outside the analysis, which represents the students who did not approve because lack of study, although the pedagogical protocol agreed with the preference of the student (who is 25%). Therefore it is possible to conclude that: [a] course B is related to cluster 1: since the errors induced by elements of cluster 2 within the course are in a 75% or in other words, the network classifies to 75% of the students failed in the course and [b] course A is related to cluster 2: since another possible allocation in this case does not exist and in addition the percentage error of classification and reprobation is of 0%. The obtained results agree with the affirmations of Perkins, where the Backpropagation network predicts that most of the failed students must have received classes using another pedagogical protocol. Socratic protocol is related with Cluster 2 and Magistral protocol is related with Cluster 1. This way the same turn out of the inferential step is obtained. In order to incorporate the experimental results to the design of the tutorial module of Intelligent Tutoring System (ITS), certain conclusions can be established. One concludes, that controls a module of the tutor able to categorize to the students according to its characteristics, within some of the pedagogical protocols available in the system. for the case in study, controlled data of 2 pedagogical protocols (Magistral and Socratic) and in this case is possible to be categorized automatically to the students within each one of them, according to its preferences to improve the results of a pedagogical session.

5. Conclusions

When validating the model against the real data. as much for the data triangulation as the training of the neural networks that support the model. it was found that the data adapt very satisfactorily to the test conditions, becoming thus. not only a theoretical tool being worth to guide the students in the learning process. but also in a validated instrument. In practice, that allows implantations of an Intelligent Tutorial System able to generate measurable and useful satisfactory results in real environments. It is fulfilled then the primary objective of this work which is to provide an additional tool for the human tutors, who can relegate some of their tasks that, either by lack of time or resources, cannot fulfill in a satisfactory way the student request, whereas it provides a secondary support for the students who try to complement their knowledge or to regulate its own rate of learning. Then, it is provided to the field of the Intelligent Tutorial Systems a new tool, to facilitate the selection of the suitable pedagogical protocol, resulting this in a gain, not only for the performance of the STI itself, but also for the student, who is the fundamental human component that makes the system useful and offers identity to them. Thus it was tried to make a contribution to improve the academic performance of different students and therefore their quality of life.

References

Costa, G.; Salgueiro, F., Cataldi, Z., García-Martínez, R. y Lage, F. (2005). Intelligent Systems for student modelling. Proc. GCETE'2005, Global Congress on Engineering and Technology Education CD. march 13-15. Brazil.

Felder R. & Silverman L. (1988). Learning Styles and Teaching Styles in Engineering Education. Engr. Education, Volumen 78, número 7, p. 674-681.

Figueroa, N.; Lage, F.; Cataldi, Z.; Denazis, J. (2003). Evaluation of the experiences for improvement of the process of learning in initial subjct of the computer science careeer. Proc. Int. Conf. on Eng. and Computer Educ. ICECE 2003, March16-19. San Paulo. Brazil.

Figueroa, N.; Cataldi Z.; Salgueiro F. A.; Costa G.; Méndez P.; Lage F. J. (2004). The styles of learning and the university desertion in Computer science Engineering. CACIC 2004. Arg.

García, M. (1995). El paradigma de Pensamiento del profesor. Editorial CEAC, Barcelona.

Hernandez Sampieri, C. (2001). Metodología de la investigación. Mc Graw Hil. México.

Kohonen, T. (2001). Self-Organizing Maps, third edition. Springer.

Quinlan, J. (1987). Simplifying Decision Trees. Simplifying Decision Trees. International Journal of Man-Machine Studies 27(3): 221-234.

Quinlan, J. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann.

Salgueiro, F., Costa, G., Cataldi, Z., García-Martinez, R. y Lage, F. (2005). Intelligent Systems for tutor modelling. Proc. Global Congress on Eng. and Tech. Education. march 13-15

Oates, T. (1997). The Effects of Training Set Size on Decision Tree Complexity. Proc. 14th International Conference on Machine Learning.

Perkins, D. (1995). Smart schools. The Free Press. A division of Simon & Schuster, Inc.