

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

**ANSI***net*

Asian Network for Scientific Information  
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

## A CBR System for Predicting Student Achievement

Dr. Stamos T. Karamouzis

Department of Computer Information Science, Loyola University, New Orleans,  
Campus Box 14, 6363 St. Charles Ave. New Orleans, LA 70185, USA

**Abstract:** Research in predicting student performance in an educational setting resulted in the development of an intelligent system that uses case-based reasoning in order to forecast student class performance. The system drew conclusions on the basis of similarities between a student's current class performance and the performance of other students that attended the same class. This study evaluated the intelligent system and presents the results achieved. A Turing-like comparison, where the system's performance was compared and contrasted with the prediction abilities of human instructors, placed the achieved results in perspective. Findings of the evaluation indicated that in comparison to humans the system outperformed non-expert instructors. Educators may develop similar systems that are customized to the structure of their own classes and are capable of assisting them in advising students on their class progress way before it is too late for the student.

**Key words:** Prognostication, case-base reasoning, expert system, student achievement, application in education

### INTRODUCTION

Advising students on their class performance and motivating them in order to continue or improve their performance is an integral part of every instruction. Study findings suggest that students in academic jeopardy who receive diagnostic and prescriptive information do better academically than students who receive no such treatment<sup>[1-3]</sup>. The advising task may be enhanced with the use of methodologies or tools that forecast the student's future class performance. Predicting future class performance is a difficult process and every attempt to automate this task must overcome a number of challenges. To address these challenges, Karamouzis<sup>[4]</sup> reports the development of a prototype electronic system, called FIG (FInal Grade). This study evaluates FIG's predictive power. Additionally, a Turing-like comparison, where the system's performance is compared and contrasted with the prediction abilities of human instructors, places the achieved results in perspective.

FIG's implementation employs a rapidly growing methodology for dealing with decision making, that involves the use of case-based techniques<sup>[5]</sup>. Case-based Reasoning (CBR) systems solve new problems by finding solved problems similar to the current problem and adapting their solutions to the current problem, taking into consideration any differences between the current and previously solved situations. Because CBR systems associate features of a problem with a previously derived solution to that problem, they are classified as associational reasoning systems.

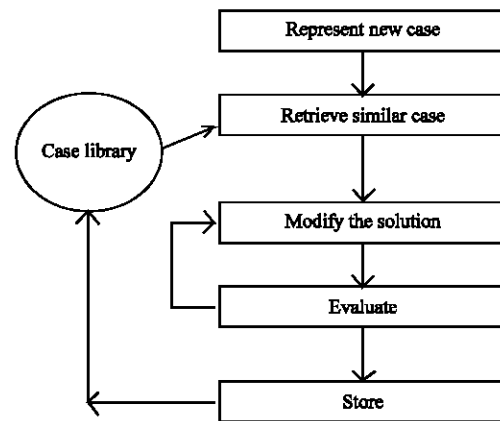


Fig. 1: Basic algorithm of case-based systems

The basic characteristic of a CBR system is its ability to represent and utilize a library of cases that at least coarsely cover the problems that come up in a particular domain. The basic CBR algorithm is represented in Fig. 1<sup>[6]</sup>. Various CBR applications and domains emphasize different parts of this algorithm. However the overall quality of CBR solutions depends on; a) The quality of cases that the systems case library has; b) The appropriate indexing of cases so they can be recalled under appropriate circumstances; c) The appropriate adaptation mechanisms for modifying and old solution to the demands of a new situation<sup>[7]</sup>.

**Quality of case library:** There are two basic modes of Case-based Reasoning: the statistical mode and the

expert system mode. In the statistical mode the case base is statistically representative of the population. Cases are retrieved by using nearest neighbor or decision trees are induced from the case base using heuristics that rely on the distribution of cases. In the expert systems mode the case base is stocked with a few well-chosen examples. These are chosen to be informative, but do not accurately represent the distribution of cases in the population.

**Indexing of cases:** Recalling cases from the case library is a massive search problem. In realistic domains the case library may contain thousands of cases which rules out employment of retrieval methods that rely on an exhaustive search. Instead library cases are indexed by appropriate features, thus making the retrieval process more selective. Selecting the appropriate set of indices is a crucial issue for successfully choosing the case library that matches the most with the current case. A variety of methods and techniques, including inductive learning and explanation-based techniques, are used to determine the case features that may be used as indices.

**Adaptation:** In most instances it is unlikely that a library case is exactly the same with the new situation. The solution from the case that is most similar to the new situation must be modified in order to accommodate the current situation. The modification is done using rules; thus the selection of the appropriate rules is crucial for a successful modification. These rules can be much simpler than those required by a pure rule-based system but developing the rules requires deep domain knowledge about what kinds of factors make two cases differ.

## DOMAIN

As a prototype, FIG was designed, developed and tested for an introductory computer science class at a University. The class attracts roughly 350 students per semester. The students' interests are very diverse and range from Computer Science to History. Every week the students must attend two hours of class lecture and two hours of laboratory lecture. Most of the times the class lecturer and the laboratory lecturer are not the same instructor. Student attendance is not mandatory in either lecture.

Student performance is determined through eight homework assignments, two midterm examinations, a series of twelve laboratory assignments and a final examination. The homework assignments accounted for 35% of the student's final grade, each midterm examination for 10%, the 12 laboratories for 30% and the final examination for 15% of the student's final grade.

Although homework and laboratory assignments are similarly weighted their difficulty level is quite different.

Students find laboratory assignments easy while homework appears to them more difficult and the written examinations even more difficult. This, along with the very diverse student body, indicates that traditional forecasting methods such as regression or a rule-based reasoning system may not be the most appropriate forecasting method for this domain. On the other hand, the fairly large number of students indicates that a case-based reasoning approach may be more natural.

## MATERIALS AND METHODS

**Reasoning:** Believing that analogues may provide a way to predict results based on what has been true in the past, FIG's reasoning mechanism was developed along the lines of a traditional case-based reasoning system. It maintains a library of cases that represent the class performance of students along with the final grade given to them by their instructor. When the partially known class performance of a new student is presented to FIG, the reasoner searches its library, finds the library student that his or her performance is most similar with the class performance of the new student and finally predicts that the letter grade given to the student in the library will be the letter grade to be assigned to the new student after the completion of the class. At this stage there is no adaptation phase.

**Cases:** The subjects in this study consisted of 105 students that already attended the introductory college class. Each student constituted a case that contained the following information: class, date, the student's performance in each of the eight homework assignments, the student's performance in each of the three examinations, the student's performance in all of the laboratories and the student's final grade for this class. Cases that included all of this information comprised the space of library cases. Cases that were lacking the final grade, but contained at least a subset of the remaining information, were considered as test (input) cases. The choice of students that comprised the library cases and the students that comprised the test cases was random.

FIG's task is to find an appropriate value for the final grade field of an input case; therefore this field is considered the solution data for a particular case in this domain. The possible data values for the solution data are the characters A, B, C, D, F, W that represent grade levels. A is the highest possible grade, F is the failure grade and W denotes withdrawal from the class. The possible values for the examination and homework fields are integers that range between 0 and 100. The possible value for the laboratory field is an integer between 0 and 275 and the possible values for the class and date fields are any character string. Except for the class, date and final grade

fields, all other fields were used for indexing each library case. The following exemplify typical library and input cases:

Library case		Test case	
Feature	Value	Feature	Value
Class:	CS131	Homework 1:	90
Date:	Fall 1998	Homework 1:	76
Homework 1:	100	Homework 1:	61
Homework 2:	74	Examination 1:	61
Homework 3:	84	Examination 2:	75
Homework 4:	64	Laboratory:	250
Homework 5:	76	Final Examination	77
		Final Grade	C

**Matching:** During testing of FIG's predictive power the reasoner used two matching methodologies. Based on the first method in order to establish the similarity between a certain test case and a library case, FIG compared corresponding features one at a time; for example the first homework of the test case with the first homework of the library, the second homework with the second one and so on. Since each homework assignment was equally weighted, instead of comparing corresponding homework values, the second matching mechanism compared the mean homework values between a certain test case and a library case.

For example, let us assume that each case contains only two features that describe performance in two homework assignments. Additionally, we assume that the student represented by a library case achieved a score of 50 in the first homework assignment and a score of 90 in the second homework and that the student represented by a test case achieved a score of 90 in the first homework assignment and a score of 50 in the second homework assignment. In this example, if we employ the first matching mechanism the reasoner will establish that the two students are completely dissimilar. On the other hand, if we employ the second matching mechanism we have 100% similarity between the two cases.

**Development tool:** FIG was developed using ESTEEM™, a tool for building problem solving applications that use case-based reasoning. The cases used in FIG were constructed using ESTEEM's case editor. The system runs on an IBM compatible personal computer with Microsoft Windows.

**Evaluation procedure:** FIG's forecast power was evaluated by utilizing a hold-out set<sup>[8]</sup>. The available case space (105 cases) was partitioned into a training set of 85 library cases and a test set of 20 test (input) cases. Each test case contained six features. The first two of these

features are used to identify the particular class that a certain student is enrolled. The remaining four features denote the partially known class performance of the same student and they are used as indexing features. These four features are the student's grades in the first, second and third homework assignments and the student's grade in the first examination. Given this partially known class performance, i.e., given a test case, FIG's task was to predict the student's final grade in the class. This was done by using the case library in order to find the case or cases that are most similar to the test case. Similarity was determined by comparison of corresponding indexing features. Corresponding indexing features with identical numerical values receive a similarity count of 1 while corresponding features that the absolute value of their difference is greater than 10% receive a similarity count of 0. If the difference is less than 10% then the similarity count is a numerical value between 0 and 1. The sum of the similarity counts for each feature constitutes the degree of similarity between two cases; therefore the maximum possible match value between two cases is equal to the number of case features. For example, the previously shown library and test cases exhibit a similarity degree of 2.5 if compared with each other.

FIG was evaluated for predictability and matching confidence. Predictability is defined as the measure of FIG's ability to correctly predict the solution data in a set of test cases. A correct prediction is a letter grade that agrees with the grade that eventually was assigned to the student by his/her instructor at the end of the semester. For example, for a given test case where the correct solution data is the letter grade A if FIG retrieved four library cases as being the most similar cases and if three of them indicate that the final grade is an A and one of them that the final grade is a B then we say that FIG's prediction rate is 75%.

Matching confidence is a measure that is directly related to the degree of similarity between a test case and the most similar library case. It represents the degree of similarity as a percentage after taking in consideration that each index feature is weighted in terms of how important the feature is in establishing similarity between two cases. In general, the matching confidence indicates how certain is FIG in offering a prediction for a particular test case (student). For example, the previously shown library and test cases exhibit a 62.5% matching confidence since their degree of similarity is 2.5.

## RESULTS

FIG was tested with 20 cases and a library of 85 cases. Each test case included the student's performance on the first, second and third homework

Table 1: Instructors' prediction and confidence rates

Instructor	Prediction	Confidence
1	60.00	92.50
2	60.00	88.75
3	65.00	70.50
4	85.00	78.75
5	80.00	69.50
6	75.00	60.45
Mean	70.83	76.74

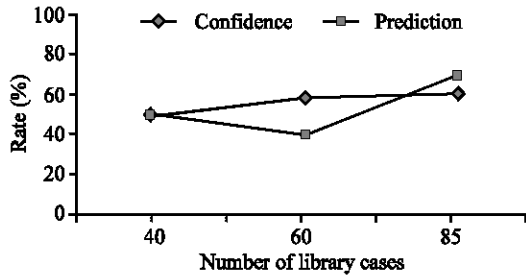


Fig. 2: Performance as the case library increases

assignments and the student's performance on the first examination, i.e., each test case consisted of 33.3% of the required information in order for someone to accurately determine the student's final grade. During testing FIG achieved a mean prediction rate of 70% and a mean matching confidence rate of 60.7%. Figure 2 demonstrates the system's progressive performance as the library case was augmented from 40 to 85 cases.

The present results were achieved when FIG employed the first matching methodology. A second test was performed where the second matching methodology was employed. During this second test the mean prediction rate was 72.5% and the mean matching confidence rate was 96.4%.

**Human instructors:** In order to put the achieved results into perspective a turing-like comparison was performed by comparing the system's performance with the prediction abilities of human instructors. Six experienced college instructors with several years of teaching experience agreed to participate in this study. Each of the instructors was given the same 20 test cases that FIG was tested. The instructors' task was to predict each student's final grade and to declare their confidence in their prediction as a percentage value. The instructors' mean prediction rate was 70.83% and their mean confidence rate was 76.74%. Table 1 displays each instructor's prediction and confidence rates.

Although all of the instructors had considerable teaching experience the first three were considered non-experts since they never taught the same class that data was extracted from. The remaining three were considered experts because in the past they taught this

particular class. Considering the expert instructors their mean prediction rate was 80% while their mean confidence rate was 69.6%. The non-expert instructors' mean prediction rate was 61.7% and their mean confidence rate was 83.9%.

## DISCUSSION

In summary, based on the belief that prior experience may provide a way to predict results grounded on what has been true in the past, an intelligent system was developed that uses case-based reasoning in order to forecast a student's class performance. The system draws conclusions on the basis of similarities between a student's current class performance and the performance of other students that attended the same class. The reasoner maintains a library of cases and operates within the traditional case-based reasoning phases of input, matching, retrieval and storage.

An evaluation test was conducted in order to identify the system's forecast power. The test involved twenty test cases that were presented to the system one at a time. The system was evaluated for predictability and matching confidence. With a progressively increasing number of library cases the system's predictably rate increased from 50 to 72%. The increasing trend was expected since the more cases the library contains, the better the chances that the system will find a case that is very similar to the input case. Since this domain requires that the case base is statistically representative of the population, it is expected that a larger library of cases will substantially improve the system's predictability rate. In comparison to humans, FIG performed at the same level (72.5% versus 70.8% on all instructors) as human instructors. When the instructors were partitioned into an expert and non-expert groups we observe that the system outperformed non-expert instructors while expert instructors outperformed it.

FIG is a prototype system and as such was developed and tested in order to investigate, in our problem domain, the potential utility of reasoning based on past history. The same predictive approach may be used by any educator. Similar systems may be developed for practically any traditional class that requires that students are evaluated throughout the class on various assignments. The choice to evaluate FIG with cases that were 33% complete, 4 indexing features out of possible 12, was random. Educators may choose to use the predictive power of such a system at any time when there is more or less information available for each student. It is expected that the more information available to each input case the better the chances for finding better matches, thus making better predictions.

One judging the potential utility of this predictive approach must consider the time required to create such a system versus the potential payoff. Findings from the development of FIG determine that the total time required for the development of such a system is relatively low. This is because: a) The developing tool reduces the time required to build the reasoning portions of the system to a minimum, b) The volume of information required to create each case is relatively small since for each class students are evaluated only on a small number of assignments and c) The nature of the domain allows for a gradual buildup of the case library. On the other hand as it is determined with FIG, the relatively high prediction rate (72.5%) along with the relatively limited amount of information (33%) needed to reach a conclusion suggest that the potential payoff is high, especially when it is used by non-expert instructors in identifying marginal students before the end of the class. In conclusion, the potential utility of this predictive approach is very promising. All of the above factors support the expectation that any educator may develop a similar system that is customized to the structure of his/her own classes and be capable of assisting in advising students on their class progress way before it is too late for the student.

#### REFERENCES

1. Tait, H. and N. Entwistle, 1996. Identifying students at risk through ineffective study strategies. *Higher Edu.*, 31: 97-116.
2. Easton, S. and D. Van Laar, 1995. Experiences of lecturers helping distressed students in higher education. *Br. J. Guidance and Counseling*, 23: 173-78.
3. Boyd, V., 1987. Diagnostic and prescriptive interviews with transfer students in academic jeopardy. Research Report No. 16-87, Counseling Center, Maryland Univ., College Park, MA.
4. Karamouzis, S., 2000. Using analogues to predict student's performance. In: *Proc. IASTED's Artificial Intelligence and Soft Comp.*, Banf, Canada, pp: 355-360.
5. Kolodner, J., 1993. *Case-based reasoning*. Morgan Kaufmann Publishers, San Mateo, CA.
6. Riesbeck, C.K. and R.S. Schank, 1989. *Inside Case-based Reasoning*. Erlbaum, Northvale, NJ.
7. Jensen, D. and F. Provost, 1999. Evaluating machine learning and knowledge discovery. Tutorial Notes of the 16th Natl. Conf. Artificial Intelligence, Orlando, Florida.
8. Kolodner, J., 1992. An introduction to case-based reasoning. *Artificial Intelligence Rev.*, 6: 3-34.