Realization of the English Assisted Learning System Based on Rule Mining

Kun Li, Yufang Song, Yan Zhang & Chunyi Lou
Qinhuangdao Institute of Technology, Qinhuangdao, Hebei, China

ABSTRACT: This paper first makes a brief introduction on the research progress of artificial intelligence, then introduces the basic structure of the whole English assisted learning system from the angle of system functional requirements, and finally discusses the realization of functions of the English assisted learning system under the support of rule-based data mining, aiming at attracting more attentions.

Keywords: English assisted learning system; data mining; rule; system structure; realization

1 INTRODUCTION

In the field of artificial intelligence, the rule-based expert system is approved and favored by a large number of researchers. It has now developed as one of the branches with the most application prospect and potential in artificial intelligence field. It has pointed out in a great deal of studies that it is worth trying to combine the expert system theory and data mining theory. The internet platform provides teaching and research activities with opportunities of application so as to achieve twice the result with half the effort. Therefore, from the angle of English learning, this paper carries out a discussion and an analysis on the design and realization of English assisted learning system through the rule-based data mining technology

2 RESEARCH OVERVIEW

Artificial intelligence is one of the relatively new research directions currently, the application fields of which are becoming increasingly extensive. Relevant researches at home and abroad become more and more mature. At present, research findings in artificial intelligence at home and abroad mainly include the following parts: First, the understanding of natural language, namely robots' understanding of human language, which is mainly realized by search engine. Second, database retrieval that means the search of useful information with related tools and technologies in a high-volume database. Third, expert system, namely the expert system established based on rules. The fourth part is mechanical theorem proving. The fifth part is robotics. The sixth part is automatic programming and the seventh part is combination and scheduling, namely simplified decomposition of complex things and solutions of optimal process.

3 SYSTEM STRUCTURE

The application of English assisted learning system provides students with a targeted practice platform, where tests of knowledge are carried out with objectives. Meanwhile, teachers are able to analyze students' answers and performance levels by inputting examination data in the system so as to guide the improvement of teaching plan. From this perspective, the English assisted learning system can avoid the fact that system data costs too much time on the one hand and it is of great significance for improving students' learning autonomy and enthusiasm on the other hand. According to the requirements mentioned earlier, the basic procedure of the whole English assisted learning system is shown as below (Figure 1).

In the procedure shown in Figure 1, the whole English assisted learning system can be classified according to the different identities of users, which include four aspects:

(1) System administrators: The major responsibility of system administrators is to conduct comprehensive maintenance on users' information of the whole learning system.

(2) Domain experts: The major responsibility of domain experts is to maintain relevant information data of the knowledge base, which not only contains basic knowledge base but also covers question bank. The main object of the former is grammar while the latter includes not only question information but also related interpretation of questions.

(3) Teachers: The major responsibility of teachers is to be in charge of utilizing the whole English assisted learning system so as to give play to its values. In the assisted learning system, the question bank provided by the system can be classified in line with the difficulty level. In practical teaching activities, teachers can not only compose test papers and input questions flexibly according to the information of the system but
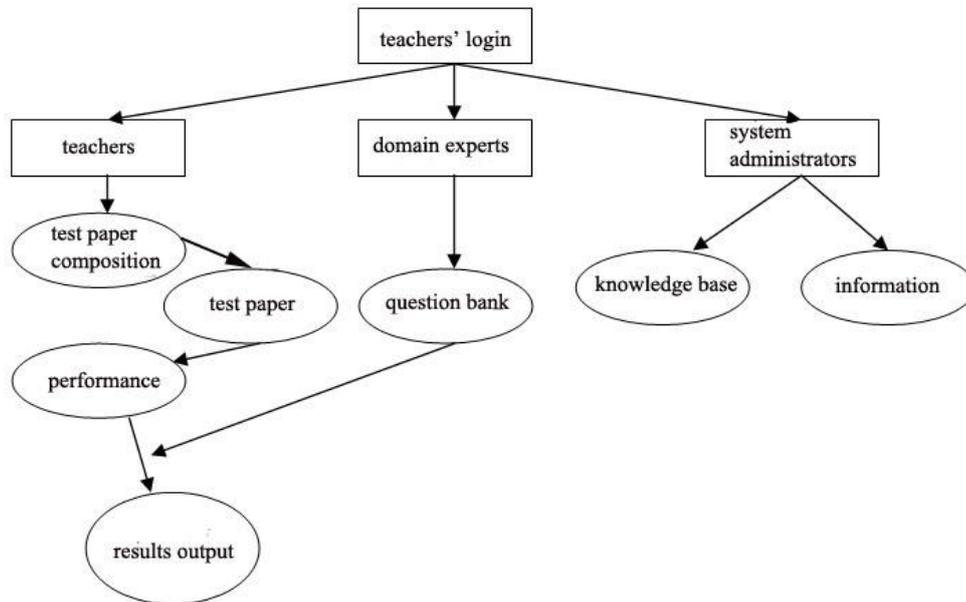
Figure 1. Basic flow of English assisted learning system

also carry out the compiling of test questions information. Examination scores are then uploaded to the learning system by teachers. According to system functions, analyses are carried out on classes and students' learning situation with the unit of a class or a student. Relevant improvement opinions and suggestions are finally provided for teaching plans.

(4) Students: The major responsibility of students is to carry out self-training positively with the English assisted learning system as a platform. In addition, the system also supports query of recorded scores for students as well as passwords modification and reset.

4 SYSTEM IMPLEMENTATION

From the perspective of system structure design, the client can be divided into two parts regarding convenience and adaptability of the English assisted learning system application. One is the teachers' end and the other is students' end. The teachers' end can be further subdivided into three parts according to different user permissions and identities, namely system administrators, domain experts and teachers. After user' login through names and passwords, the system analyzes objective answers of students (Models for reference are data uploaded to the server terminal by teachers) through data mining on the database. There are numerous knowledge points in each objective question. Therefore, the function of the English assisted learning system that needs to be satisfied first is to integrate analyses on specialized knowledge, judge answers of students and provide teachers with suggestions on the

adjustment of teaching plan based on rules of expert database.

According to this requirement, the leading function of the English assisted learning system is the rule-based data mining. This is also one of the most important goals of the system development. Users of the rule-based data mining function are teachers, who are in charge of entering data of students' answers, systematic calling of internal database (namely knowledge points in options) and definition rules (namely the division of credibility), and outputting students' mastery of a certain knowledge point.

The English assisted learning system proposes two different ways of diagnosis in terms of various requirements of students and teaching. One is the data diagnosis established based on knowledge point analysis. The other is the data diagnosis established based on the analysis on students' mastery of knowledge points. The former infers students' mastery of knowledge through the analysis on a certain knowledge point. During this process, data mainly comes from students' scores uploaded by teachers and corresponding questions of a certain answer. Implicit knowledge points can be found out from these questions through rule mining. The system can judge students' mastery of knowledge points by analyzing the overall knowledge structure of a test paper and students' answering situation. The latter analyzes the examination situation of students in groups with the unit of a class or a department. Teachers are able to master teaching situations in this way, which plays a key role in adjusting teaching arrangement. Special attentions are required in that: the rule-based data

mining function supports analyses on knowledge points in questions. Summarize knowledge points that are easily confused through the optimized Apriori Algorithm and classify the difficulty of questions according to the result.

In the basic condition of introducing optimized Apriori algorithm, the theoretical procedure of the system function realization is: If the corresponding relation between questions and knowledge points are shown as follows (Table 1), operating steps, which must be followed in association rules mining of data set when the optimized Apriori algorithm is introduced, should be:

Table 1. Diagram of the corresponding relation between questions and knowledge points

Questions	Knowledge points
Q1	I1,I2,I3
Q2	I1,I2,I3,I5
Q3	I1,I3
Q4	I2,I3
Q5	I1,I3
Q6	I1,I2,I4
Q7	I2,I3
Q8	I2,I4
Q9	I1,I2,I5

Firstly, with the basic condition that the minimum support value is 2, scan each item set and give the statistics of the appearance of each knowledge point so as to obtain the support degree count of each item set and candidate item set C1. Compare candidate support degree counts according to the given minimum support degree, and carry on the next step for counts satisfying the condition of being larger than the minimum support degree while conduct pruning process for those that do not meet the condition so as to obtain the frequent item set L1. Then connect all the data in this set to form a new candidate item set through pair-wise combination, which is defined as C2. Conduct pruning process for item sets in C2 that do not satisfy the condition of being larger than the minimum support degree so as to form L2. The frequent item set L3 can be obtained by repeating the above steps. After the connecting process of L3, the obtained candidate item set C4 has only one element {Il, I2, I3, I5}. Besides, the support degree count after pruning process is 1 that fails to satisfy the condition of being larger than the minimum support degree and the obtained frequent item set L4 is an empty set. The final structure outputted after the association rules mining is a frequent item set L3.

Illustration: Altogether 30 students in a class take part in a unit evaluation test. The teacher selects 5 questions as analytical samples after the test. At first, the teacher login in the English assisted learning system to add the five questions in the system and then determine corresponding knowledge points and analy-

sis results of options. Questions are respectively:

1) It was essential that the application forms ( ) back before the contract.

A: were sent; B: would sent; C: be sent; D: must be sent;

2) The match was cancelled because most of the members ( ) a match without a standard court.

A: objected to having; B: were objected to have; C: were objected to having; D: objected to have;

3) The goals ( ) he had fought all his life no longer seemed important to him.

A: with which; B: for which; C: after which; D: at which;

4) I would appreciate ( ) it a secret.

A: you to keep; B: you keeping; C: that you keep; D: that you will keep;

5) The children went there to watch the iron tower ( ).

A: be erected; B: to erect; C: being erected; D: erecting;

Standard answers of the above five questions are respectively: C; A; B; B; A. Corresponding knowledge points are defined as vector y, meanings of which are: vector y1 stands for conjunctions, vector y2 stands for subject-predicate consistency, vector y3 stands for verbs, vector y4 stands for sentence patterns, and vector y5 stands for adjectives.

Table 2. Diagram of students' answers

Question No.	A	B	C	D
1	5	5	16	4
2	15	13	1	1
3	8	16	4	2
4	2	16	2	10
5	13	6	4	7

After the examination, answers of students can be provided in statistical data shown in the following table (Table 2) according to uploaded data results.

In the English assisted learning system, the above data can be stored in a database after organization. Judge the correctness of answers first and then match these answers with defined rules in the system knowledge base so as to get the value of u, namely quantitatively processing on knowledge points vector y. The corresponding quantized value of the knowledge point yn is:

$$U_{yn} = x/n;$$

Where, yn stands for the corresponding knowledge point vector of a certain question; x is the frequency level of students' answers of the specific knowledge point yn; n is the integration of all questions.

The quantized value u can be further processed with weighting on this basis. The weighted vector is defined as W and the basic principle that must be followed in the weighting process is: As for samples need to be tested, the corresponding weighted value of

a certain knowledge point is the result of weighting and superposition of credibility (defined as C). C is determined by rules in the system according to experience of domain experts. In the meantime, the value of weighted vectors decreases progressively in the unit of 2^n. According to the above processing, vectors analysis table of students' answers after processing is shown in the following table (Table 3).

Table 3. Vectors analysis table of students' answers

Question No.	1	2	3	4	5
Weighted values of u	16	8	21	7.5	3
Conclusion credibility	0.003	0.01	0	0.02	0.03

It can be concluded from the above analysis: If all the students choose options that are "mastered knowledge points" and the answer is a determined answer, the weighted value is the minimum, which is 0; if all the students choose options that are "not mastered knowledge points" and the answer is a determined answer, the weighted value is the maximum, which is 30.

With the combination of the above analysis results, the analysis on this unit test results outputted by the system is given in the following table (Table 4).

Table 4. Test analysis results of output class

Question No.	Mastery condition	Conclusion credibility
1	Very good	Relatively credible
2	Very good	Generally credible
3	Very good	Credible
4	Normal	Generally credible
5	Very good	Relatively credible

5 CONCLUSION

This paper mainly carries out an analysis and a study on the design and realization of the English assisted learning system, introduces the expert system that has the greatest potential in artificial intelligence field, optimizes data mining with Apriori algorithm on the basis of rules, and conducts undetermined reasoning on students' answers to enrich and improve system functions. However, as for objective problems like small scale of database, researches and development still need to be further carried out on the premise of guaranteeing network stability so as to provide paperless teaching with platform support.

REFERENCES

[1] Cai, J.X. & Ruan, X.G. 2011. The application of OCPA bionic autonomous learning system in balance control of robot postures, *Pattern Recognition and Artificial Intelligence*, 24(1): 138-146.

[2] Xie, M.F. & Sun, X. 2012. A study on the remote personalized network learning system model based on ontology knowledge management, *China Audio-Visual education*, (11): 47-53.

[3] Li, Y.F., Wang, M.J. & Wang, J.J., et al. 2012. Mobile learning system and relevant learning modes, *Open Education Research*, 18(1): 152-158.

[4] Wang, C.H., Zhong, Y.J. & Zhang, Y.H., etc. 2014. A research on the design of intelligent learning system of physics in junior high school, *China Audio-Visual education*, (10): 90-95.

[5] Wang, Y., Gao, Y. & Guo, H., etc. 2011. Website construction and teaching application of the integration of English learns' corpus and autonomous learning system, *China University Teaching*, (12): 62-64.

[6] Xiao, W.Q. & Zeng, L.R. 2012. Cloud Service vocabulary learning system based on word frequency and forgetting curve, *Foreign Languages and Literature (Journal of Sichuan International Studies University)*, (4): 119-124.

[7] Fernando Aparicio, Manuel De Buenaga. & Margarita Rubio, et al. 2012. An intelligent information access system assisting a case based learning methodology evaluated in higher education with medical students, *Computers & Education*, 58(4): 1282-1295.

[8] Annabel Latham, Keeley Crockett. & David McLean, et al. 2012. A conversational intelligent tutoring system to automatically predict learning styles, *Computers & Education*, 59(1): 95-109.

[9] OEzcan OEzyurt, Hacer OEzyurt. & Adnan Baki, et al. 2013. Design and development of an innovative individualized adaptive and intelligent e-learning system for teaching-learning of probability unit: Details of UZWEBMAT, *Expert Systems with Application*, 40(8): 2914-2940.

[10] Annabel Latham, Keeley Crockett. & David McLean, et al. 2014. An adaptation algorithm for an intelligent natural language tutoring system, *Computers & Education*, 71(2): 97-110.