Mining Concepts' Relationship Based on Numeric Grades

Mohammed Al-Sarem[1], Mostafa Bellafkih[2] and Mohammed Ramdani[3]

[1]**Département d'informatique, FST Mohammedia**
BP 146 Mohammedia 20650, Morocco

[2]**Institut National des Postes et Télécommunications (INPT)**
Rabat, Morocco

[3]**Département d'informatique, FST Mohammedia,**
BP 146 Mohammedia 20650, Morocco

Abstract

With vigorous development of the Internet, e-learning system has become more and more popular and many adaptive learning systems have been developed. In recent years, researchers have proposed various approaches for developing adaptive learning systems based on concept maps. Nevertheless, most of them deal only with binary grades of each test item, furthermore, the existing methods that based on fuzzy set theory do not take in consideration the conceptual weight of concept in each question that might be cause to construct incorrectly relationships or exaggerate the degree of relationship. To cope with this problem, this study proposes an innovative approach to automatically construct concept maps. Firstly, we use look ahead fuzzy association rule mining algorithm to mind some information about the relationships between questions, then we construct the questions-relationships mapping. After that, we calculate the relevance degree between concepts in each question to obtain the final concept maps.

Keywords: *Fuzzy Association Rules, Data Mining, Concept Maps, Analysis of Numerical Testing Scores.*

1. Introduction

With vigorous development of the Internet, e-learning system has become more and more popular and many adaptive learning systems have been developed. In recent years, researchers have proposed various approaches for developing adaptive learning systems based on concept maps. Nevertheless, most proposed methods [1-6, 8, 10, 11, and 12] deal only with binary grades of each test item. Furthermore, the existing methods that based on fuzzy set theory [15, 16] do not take into consideration the weight of concept in each question that might be cause to construct incorrectly relationships or exaggerate the degree of relationship. Therefore, in this paper, we propose an approach to automatically construct concept maps based on the result of analysis of numerical testing scores. Firstly, we use look ahead fuzzy association rule mining algorithm [16] to mind some information about the relationships between questions, then we construct the questions-relationship mapping [3]. After that, we calculate the relevance degree between concepts in each question to obtain the final concept maps.

The rest of the paper is organized as follows. In section 2, we discuss the related work. In Section 3 we briefly review the ahead fuzzy association rule mining algorithm. Proposed approach is given in Section 4. In Section 5 we use an example to illustrate the process of constructing concept maps. Section 6 concludes the whole paper and discusses the future work.

2. Related work

In 1965, Zadeh proposed the concept of a fuzzy set to describe imprecision that is characteristic of much of human reasoning. Since that time, the fuzzy set theory has found use in many applications, ranging from pattern recognition, control engineering to modeling human decision making [7]. Nowadays, fuzzy sets are increasingly used for automatically constructing concept maps. In [15], Tsai et al. proposed a Two-Phase fuzzy mining to find the embedded association rules from the historical learning records of students. In this work, we are interesting just the first phase where fuzzy set theory applied to transform the numeric testing records of students into symbolic. As known, the numeric testing data are hard to analyze by association rule mining approach [9]. Therefore, a look ahead fuzzy association rules [9, 15, and 16] is proposed to mine some rules, which are used to construct the concept maps and fed back to teachers for further analyzing [16].

3. The look ahead fuzzy association rules algorithm for mining relationships between questions

Let there are "n" learners $S_1, S_2, \ldots, S_n$ take a test of "m" questions $Q_1, Q_2, \ldots, Q_m$ and let matrix "G" is a grade matrix transformed from the test portfolio of the learners as follows:

$$G = \begin{array}{c c} & \begin{array}{cccc} S_1 & S_2 & \cdots & S_n \end{array} \\ \begin{array}{c} Q_1 \\ Q_2 \\ \vdots \\ Q_m \end{array} & \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} \end{bmatrix}_{n\times m} \end{array}$$

, where "g_{S_i,Q_j}" denotes the score of question "Q_j" of the learner "S_i", $g_{S_i,Q_j} \in [0, P]$, $1 \leq i \leq n$, $1 \leq j \leq m$ and "P" is the maximum value that can obtain by learners. In the first step, the test records of learners are fuzzified to transform the numeric grade data into symbolic. In the second step, we construct the association rules from each question in the large 1-itemset to all other questions (an item set is called a large-item set if its support value is greater or equal to the user-specified support threshold called min Support) where the support value of item set "x" in C_ℓ is calculated as follows:

$\text{support}(x) = \sum_{1}^{n} x_{nk}$, where $X \subseteq$ the ℓ-Candidate item set C_ℓ, "n" is number of learners and "k" is the fuzzification rule type.

The confidence of each rule calculated as follows:

$$\text{conf}(X \rightarrow Y) = \frac{\text{Support}(X \cap Y)}{\text{Support}(X)} \tag{1}$$

, where Support $(X\cap Y)$ is the support value of every item set "x" in Candidate "C_ℓ" that can be evaluated as follows:

Support $(X\cap Y) = \sum_{1}^{n} \text{Min}(X, Y)$, where "n" is the number of learners, X, Y are two item sets in "C_ℓ" and $X \cap Y = \emptyset$, conf $(X\rightarrow Y)$ denotes the confidence of the association rule "$X\rightarrow Y$".

The following is the detail of the Look Ahead Fuzzy Association Rule Mining Algorithm [15, 16].

Symbol Definition:

α_ℓ: The minimum support threshold in the ℓ-large items.

C_ℓ: The ℓ-Candidate item set.

L_ℓ: The ℓ-large items

λ: The minimum confidence threshold.

Input: Data set after fuzzification, the minimum support threshold α and λ.

Output: The fuzzy association rules.

Procedure:

Step1: While $C_\ell \neq NULL$

1.1: Generate and insert the ℓ-item set into C_ℓ

1.2: $\alpha_\ell = \max\left(\frac{\alpha_1}{2}, \alpha_{\ell-1} - \frac{\alpha_1}{(\ell-1)\times c}\right)$, where $\ell > 1$ and "c" is constant.

1.3: $L_\ell = \{x \mid \text{support}(x) \geq \alpha_\ell, \text{for } x \in C_\ell\}$

1.4: $\ell = \ell + 1$

Step2: Generate the association rules according to the given λ in L_ℓ.

4: **Return** $LI = \cup_k LI_k$;

Based on the mentioned above algorithm, in the next section we will present a new method to automatically construct concept maps which provide us a useful way to construct concept maps for adaptive learning system and can overcome the drawbacks of most existing methods.

4. The proposed approach

Step1: Transform the test portfolio of the learners and the conceptual weight relationships into the matrix G and the matrix QC, respectively.

$$G = \begin{array}{c c} & \begin{array}{cccc} S_1 & S_2 & \cdots & S_n \end{array} \\ \begin{array}{c} Q_1 \\ Q_2 \\ \vdots \\ Q_m \end{array} & \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} \end{bmatrix}_{n\times m} \end{array}$$

, where "g_{S_i,Q_j}" denotes the score of question "Q_j" of the learner "S_i", $g_{S_i,Q_j} \in [0, P]$, "P" is the maximum value that can obtain by learners, $1 \leq i \leq n$, $1 \leq j \leq m$, "n" is number of learners and "m" is number of questions.

The questions-concepts matrix QC shown as follows:

$$QC = \begin{array}{c c} & \begin{array}{cccc} C_1 & C_2 & \cdots & C_p \end{array} \\ \begin{array}{c} Q_1 \\ Q_2 \\ \vdots \\ Q_m \end{array} & \begin{bmatrix} qc_{11} & qc_{12} & \cdots & qc_{1p} \\ qc_{21} & qc_{22} & \cdots & qc_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ qc_{m1} & qc_{m2} & \cdots & qc_{mp} \end{bmatrix}_{p\times m} \end{array}$$

, where "qc_{ij}" denotes the weighting or degree of relevance for each concept to each question integrated from multiple experts and $0 \leq qc_{ij} \leq 1$.

Step 2: Fuzzificate the grade matrix to transform numeric grade into symbolic data, then apply the mentioned above algorithm to find the association rules of test item.

Step 3: Based on the associated rules derived in Step 2, we construct two kinds of questions-relationship maps [3]:

1. For the association rules "$Q_i.L \rightarrow Q_j.L$", "$Q_i.H \rightarrow Q_j.L$" and "$Q_i.H \rightarrow Q_j.H$" (the related explanations of the

rule types are shown in Table 1), we build a relationship from question "Q_i" to question "Q_j".

2. For the rules "$Q_i, L \rightarrow Q_j, H$", we build a relationship from question "Q_j" to question "Q_i".

Let the confidence of an association rule be the confidence of the relationship between questions builds from it. For any two questions "Q_i" and "Q_j":

1. If the confidence of the questions-relationship is smaller than the minimum confidence "θ", delete the relationship between the questions "Q_i" and "Q_j" to get the completed questions-relationship map.
2. If there is more than one relationship between any two questions "Q_i" and "Q_j", we only keep the relationship with the maximum degree and delete the others.

Table 1: The explanations of rule types [16]

Rule	***Description of relationships***
$Q_i, L \rightarrow Q_j, L$	It is means that the related concepts in question "Q_i" are the prerequisite of concepts in "Q_j" and explain why getting low grade in question "Q_i" might imply getting low grade on "Q_j"
$Q_i, L \rightarrow Q_j, H$	It is means that the related concepts in question "Q_j" are the prerequisite of concepts in "Q_i" because "Q_i" may be not learned well resulting from "Q_j"
$Q_i, H \rightarrow Q_j, L$	It is means that the concepts of question "Q_i" are the prerequisite of concepts in "Q_j"
$Q_i, H \rightarrow Q_j, H$	It is means that the concepts of question "Q_i" are the prerequisite of concepts in "Q_j"

Step 4: Calculate the relevance degree between concepts. For all association rules ""$C_i \rightarrow C_j$"" obtained in Step 3, we calculate the relevant degree [6] "$rev(C_i \rightarrow C_j)_{Q_xQ_y}$" between concepts "$C_i$" and "$C_j$" from the relationship "$Q_x \rightarrow Q_y$", shown as follows:

$$rev(C_i \rightarrow C_j)_{Q_xQ_y} = W_{Q_xC_i} \times W_{Q_yC_j} \times conf(Q_x \rightarrow Q_y). \qquad (2)$$

Where "$rev(C_i \rightarrow C_j)_{Q_xQ_y}$" denotes the relevance degree of the relationship "$C_i \rightarrow C_j$" converted from the relationship "$Q_x \rightarrow Q_y$", $rev(C_i \rightarrow C_j)_{Q_xQ_y} \in [0,1]$, "$C_i$" denotes a concept appearing in the question "Q_x", "C_j" denotes a concept appearing in the question "Q_y", "$W_{Q_xC_i}$" denotes the weight of the concept "C_i" in the question "Q_x", "$W_{Q_yC_j}$" denotes the weight of the concept "C_j" in the question "Q_y", "$conf(Q_x \rightarrow Q_y)$" denotes the confidence of the relationship "$Q_x \rightarrow Q_y$", $x \neq y$, $1 \le x \le m$, $1 \le y \le m$ and $1 \le i \le p$. Furthermore, let "$conf(Q_x \rightarrow Q_y)$" be the confidence of the relationship "$C_i \rightarrow C_j$".

If there is more than one relationship between any two constructed concepts, then the relationship between the two concepts chosen as follows:

$$rev(C_i \rightarrow C_j) = Max(rev(C_i \rightarrow C_j)_{Q_xQ_y}). \qquad (3)$$

Step 5: Create a new concept-concept matrix C' based on the conceptual weights found in matrix QC, shown as follows:

$$C' = \begin{array}{c} \\ C_1 \\ C_2 \\ \vdots \\ C_n \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array} \\ \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \end{array}, b_{ij} = \begin{cases} 0, & i = j \\ \min(qc_i, qc_j), & i \neq j \end{cases}$$

, where "b_{ij}" denotes the minimum nonzero values in a concept C_i's column and C_j's column for all questions from matrix QC, $b_{ij} \in [0, 1]$ and $1 \le i, j \le n$.

Step 6: Construct the concepts-relationship map

Let $\mu = Min(b_{ij})$, where "μ" denotes the minimum nonzero value of the relevance degree in C' matrix and "μ" ∈ [0, 1]. For each relationship "$C_i \rightarrow C_j$", calculate the relative questions-concepts value between concepts "C_i" and "C_j" shown as follows:

$$Ret(C_i \rightarrow C_j) = \frac{N_i}{N_j}. \qquad (4)$$

Where "$Ret(C_i \rightarrow C_j)$" denotes the relative questions-concept value between concepts "C_i" and "C_j", N_i denotes the number of the questions which have concept C_i and N_j denotes the number of the questions which have concept C_j.

$$Ret(C_i \rightarrow C_j) \times b_{ij} \ge \mu \qquad (5)$$

If equation (5) is true and the relevance degree between any two concepts is greater than μ, add an edge from concept C_i to C_j into the concept map with the relevance degree of relationship "$C_i \rightarrow C_j$" to construct a co ncept map. Otherwise, delete it.

Step 7: If there is more than one relationship between concepts "C_i" and "C_j", we only keep the concept-relationship with the maximum relevance degree and delete the others.

5. Illustrated example

Assuming that there are ten learners take a test. Let in test sheet, there are five questions and the chart of their portfolio shown in Table 2.

Table 2: The chart of test portfolio

Question	Learners									
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Q_1	12	12	12	2	2	2	20	10	18	20
Q_2	18	14	16	8	8	10	5	6	5	3
Q_3	20	18	14	12	12	8	5	6	5	4
Q_4	20	3	4	6	2	2	4	1	1	0
Q_5	7	7	7	20	12	20	10	18	15	15

If we convert the chart of test portfolio shown in Table 2 to the grade matrix G, we get the matrix as shown as follows:

$$G = \begin{bmatrix} 12 & 12 & 12 & 2 & 2 & 2 & 20 & 10 & 18 & 20 \\ 18 & 14 & 16 & 8 & 8 & 10 & 5 & 6 & 5 & 3 \\ 20 & 18 & 14 & 12 & 12 & 8 & 5 & 6 & 5 & 4 \\ 20 & 3 & 4 & 6 & 2 & 2 & 4 & 1 & 1 & 0 \\ 7 & 7 & 7 & 20 & 12 & 20 & 10 & 18 & 15 & 15 \end{bmatrix}$$

Furthermore, assuming also that there are five experts participate in integrating process of corresponding weighting of each concept respect to each question. If we integrate the opinions of experts according to [13] to the chart of the conceptual weight relationships in test questions as shown in Table 3, then convert it to the questions-concept mapping matrix, shown as follows:

Table 3: The chart of the conceptual weight relationships in test questions

Question	Concepts				
	C_1	C_2	C_3	C_4	C_5
Q_1	1	0	0	0	0
Q_2	0.2	0.7	0	0.1	0
Q_3	0	0	0	0	1
Q_4	0	1	0	0	0
Q_5	0.6	0	0.4	0	0

$$QC = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.2 & 0.7 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 \end{bmatrix}$$

[Step 2] As we had the G-matrix, we applied fuzzy concept to transform the numeric score data into the symbolic data. Let the three membership functions are proposed as shown in fig. 1. In the fuzzification result, "LOW", "MIDDLE" and "HIGH" denote "a learner S_i has a low grade in question "Q_j", "a learner "S_i" has a middle grade in question "Q_j" and a learner "S_i" has a high grade in question "Q_j" respectively.

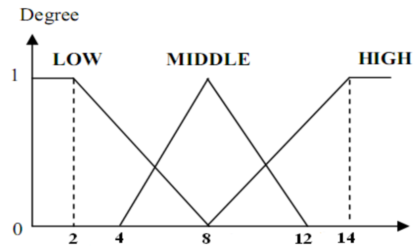

Fig.1 The given membership functions of each quiz's grade

$$d_L\left(G_{S_i,Q_j}\right) = \begin{cases} 0, & if\ G_{S_i,Q_j} \geq 8 \\ \dfrac{(8 - G_{S_i,Q_j})}{6}, & if\ 2 \leq G_{S_i,Q_j} < 8 \\ 1, & if\ G_{S_i,Q_j} < 2 \end{cases}$$

$$d_M\left(G_{S_i,Q_j}\right) = \begin{cases} 0, & if\ G_{S_i,Q_j} \leq 4\ or\ G_{S_i,Q_j} \geq 12 \\ \dfrac{(G_{S_i,Q_j} - 4)}{4}, & if\ 4 < G_{S_i,Q_j} \leq 8 \\ \dfrac{(12 - G_{S_i,Q_j})}{4}, & if\ G_{S_i,Q_j} > 8 \end{cases}$$

$$d_H\left(G_{S_i,Q_j}\right) = \begin{cases} 0, & if\ G_{S_i,Q_j} \leq 8 \\ \dfrac{(G_{S_i,Q_j} - 8)}{6}, & if\ 8 < G_{S_i,Q_j} \leq 14 \\ 1, & if\ G_{S_i,Q_j} > 14 \end{cases}$$

By given membership functions, the results of fuzzification of score data of learners are listed in Fig. 2.

learners	Q_1			Q_2			Q_3			Q_4			Q_5		
	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H
S_1	0	0	0.67	0	0	1	0	0	1	0	0	1	0.16	0.75	0
S_2	0	0	0.67	0	0	1	0	0	1	0.83	0	0	0.16	0.75	0
S_3	0	0	0.67	0	0	1	0	0	1	0.67	0	0	0.16	0.75	0
S_4	1	0	0	0	1	0	0	0	0.67	0.33	0.5	0	0	0	1
S_5	1	0	0	0	1	0	0	0	0.67	1	0	0	0	0	0.67
S_6	1	0	0	0	0.5	0.33	0	1	0	1	0	0	0	0	1
S_7	0	0	1	0.5	0.25	0	0.5	0.25	0	0.67	0	0	0	0.5	0.33
S_8	0	0.5	0.33	0.33	0.5	0	0.33	0.5	0	1	0	0	0	0	1
S_9	0	0	1	0.5	0.25	0	0.5	0.25	0	1	0	0	0	0	1
S_{10}	1	0	0	0.83	0	0	0.67	0	0	1	0	0	0	0	1

Fig.2 The fuzzification of learners' testing records

As a result of Look Ahead fuzzy association rules, we found four association rules types, LL, L-H, H-H, and H-L [15, 16], where the support value of item set "x" in Candidate_1 is calculated by adding learners' fuzzification results of the item set up [9] and Large_1 item set is generated by deleting item set that is less than α_I=2.1 as shown in Fig. 3.

C_1

1-Itemset	Support
Q_1,L	4
Q_1,H	4.33
Q_2,L	2.16
Q_2,H	3.33
Q_3,L	2
Q_3,H	4.34
Q_4,L	7.5
Q_4,H	1
Q_5,L	0.48
Q_5,H	6

α_1=2.1 $\Rightarrow$

L1

1-Itemset	Support
Q_1,L	4
Q_1,H	4.33
Q_2,L	2.16
Q_2,H	3.33
Q_3,H	4.34
Q_4,L	7.5
Q_5,H	6

Fig.3 The obtained 1-itemset

Then, by combining each item with the large 1-itemset, we can obtain the support value of every item set in 2-itemset as shown below:
$support(X \cap Y) = \sum_1^n Min(X,Y)$, where n is the number of learners, X and Y are two item sets in C_i
For example, Support (Q_1,L$\cap Q_2$,L) = Min(0,0)+ Min(0,0)+ Min(0,0)+ Min(1,0)+ Min(1,0)+ Min(1,0)+ Min(0,0.5)+ Min(0,0.33)+ Min(0,0.5)+ Min(1,0.83)=0.83

C_2

1-Itemset	Support
$Q_1,L \cap Q_2,L$	0.83
$Q_1,L \cap Q_2,H$	0.33
$Q_1,L \cap Q_3,H$	1.34
$Q_1,L \cap Q_4,L$	3.33
$Q_1,L \cap Q_5,H$	3.67
$Q_1,H \cap Q_2,L$	1.33
$Q_1,H \cap Q_2,H$	2
$Q_1,H \cap Q_3,H$	2
$Q_1,H \cap Q_4,L$	3.34
$Q_1,H \cap Q_5,H$	1.66
$Q_2,L \cap Q_3,H$	0
$Q_2,L \cap Q_4,L$	2.16
$Q_2,L \cap Q_5,H$	1.99
$Q_2,H \cap Q_3,H$	3
$Q_2,H \cap Q_4,L$	1.83
$Q_2,H \cap Q_5,H$	0.33
$Q_3,H \cap Q_4,L$	2.5
$Q_3,H \cap Q_5,H$	1.34
$Q_4,L \cap Q_5,H$	5.33

α_1=1.8 $\Rightarrow$

L2

1-Itemset	Support
$Q_1,L \cap Q_4,L$	3.33
$Q_1,L \cap Q_5,H$	3.67
$Q_1,H \cap Q_2,H$	2
$Q_1,H \cap Q_3,H$	2
$Q_1,H \cap Q_4,L$	3.34
$Q_2,L \cap Q_4,L$	2.16
$Q_2,L \cap Q_5,H$	1.99
$Q_2,H \cap Q_3,H$	3
$Q_2,H \cap Q_4,L$	1.83
$Q_3,H \cap Q_4,L$	2.5
$Q_4,L \cap Q_5,H$	5.33

Fig.4 The obtained 2-itemset

According to Eq. (1), the mined rules and their corresponding confidence are shown in Table 4.

Table 4: The confidence of mind rules

Rules	***Confidence***	***Rules***	***Confidence***
$Q_1,L \rightarrow Q_4,L$	0.83	$Q_3,H \rightarrow Q_2,H$	0.69
$Q_1,L \rightarrow Q_5,H$	0.92	$Q_3,H \rightarrow Q_4,L$	0.58
$Q_1,H \rightarrow Q_2,H$	0.46	$Q_4,L \rightarrow Q_1,L$	0.44
$Q_1,H \rightarrow Q_3,H$	0.46	$Q_4,L \rightarrow Q_1,H$	0.45
$Q_1,H \rightarrow Q_4,L$	0.77	$Q_4,L \rightarrow Q_2,L$	0.29
$Q_2,L \rightarrow Q_4,L$	1	$Q_4,L \rightarrow Q_2,H$	0.24
$Q_2,L \rightarrow Q_5,H$	0.92	$Q_4,L \rightarrow Q_3,H$	0.33
$Q_2,H \rightarrow Q_1,H$	0.60	$Q_4,L \rightarrow Q_5,H$	0.71
$Q_2,H \rightarrow Q_3,H$	0.90	$Q_5,H \rightarrow Q_1,L$	0.61
$Q_2,H \rightarrow Q_4,L$	0.55	$Q_5,H \rightarrow Q_2,L$	0.33
$Q_3,H \rightarrow Q_1,H$	0.46	$Q_5,H \rightarrow Q_4,L$	0.89

Assume that, the minimum confidence given by user 0.75, and then the mined rules and corresponding prerequisite relationship of questions are shown in Table 5.
[Step 3] Based on the rules obtained in Step 2, we construct the questions-relationship map, as shown in Fig. 5.

Table 5: Mined prerequisite relationship of questions

Rule types	***Mined rules***	***Prerequisite relationship***	***Confidence***
L-L	$Q_1,L \rightarrow Q_4,L$	$Q_1 \rightarrow Q_4$	0.83
	$Q_2,L \rightarrow Q_4,L$	$Q_2 \rightarrow Q_4$	1
L-H	$Q_1,L \rightarrow Q_5,H$	$Q_5 \rightarrow Q_1$	0.92
	$Q_2,L \rightarrow Q_5,H$	$Q_5 \rightarrow Q_2$	0.92
H-L	$Q_1,H \rightarrow Q_4,L$	$Q_1 \rightarrow Q_4$	0.77
	$Q_5,H \rightarrow Q_4,L$	$Q_5 \rightarrow Q_4$	0.89
H-H	$Q_2,H \rightarrow Q_3,H$	$Q_2 \rightarrow Q_3$	0.90

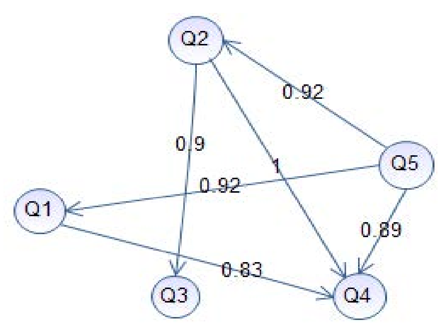

Fig.5 The questions-relationship map

[Step 4] Based on the questions-relationship map obtained in Step 3 and Eq. (2), we can calculate the relevance degree between any two concepts, shown as follows:

i. Relevance degree for the relationship $C_1 \rightarrow C_2$:
$Q_1 \rightarrow Q_4$: 1 ×1×0.83=0.83
$Q_2 \rightarrow Q_4$: 0.2× 1 ×1=0.2

$Q_5 \rightarrow Q_2$: 0.6 ×0.7×0.92=0.386
$Q_5 \rightarrow Q_4$: 0.6× 1 ×0.89=0.534
Based on Eq. (3) the maximum value among these relevance degrees is:
$rev(C_1 \rightarrow C_2)=Max(0.83,0.2,0.386,0.534)=0.83$
$conf(C_1 \rightarrow C_2)=conf(Q_1 \rightarrow Q_4)=0.83$

ii. Relevance degree for the relationship $C_1 \rightarrow C_4$:
$Q_5 \rightarrow Q_2$: 0.6 ×0.1×0.92=0.0552
Then the relevance degree for the relationship $C_1 \rightarrow C_4$= 0.0552 with confidence =0.92

iii. Relevance degree for the relationship $C_1 \rightarrow C_5$:
$Q_2 \rightarrow Q_3$: 0.2 ×1×0.92=0.184
Then the relevance degree for the relationship $C_1 \rightarrow C_5$= 0.184 with confidence =0.92

iv. Relevance degree for the relationship $C_2 \rightarrow C_5$:
$Q_2 \rightarrow Q_3$: 0.7 ×1×0.9=0.63
Then the relevance degree for the relationship $C_2 \rightarrow C_5$= 0.63 with confidence =0.9

v. Relevance degree for the relationship $C_3 \rightarrow C_1$:
$Q_5 \rightarrow Q_1$: 0.4 ×1×0.92=0.368
$Q_5 \rightarrow Q_2$: 0.4× 0.2 ×0.92=0.074
Based on Eq. (3) the maximum value among these relevance degrees is:
$rev(C_3 \rightarrow C_1)=Max(0.368,0.074)=0.368$
$conf(C_3 \rightarrow C_1)=conf(Q5 \rightarrow Q1)=0.92$

vi. Relevance degree for the relationship $C_3 \rightarrow C_2$:
$Q_5 \rightarrow Q_2$: 0.4 ×0.7×0.92=0.26
$Q_5 \rightarrow Q_4$: 0.4× 1 ×0.89=0.356
Based on Eq. (3) the maximum value among these relevance degrees is:
$rev(C_3 \rightarrow C_2)=Max(0.26,0.356)=0.356$
$conf(C_3 \rightarrow C_2)=conf(Q5 \rightarrow Q4)=0.89$

vii. Relevance degree for the relationship $C_3 \rightarrow C_4$:
$Q_5 \rightarrow Q_2$: 0.4 ×0.1×0.92=0.036
Then the relevance degree for the relationship $C_3 \rightarrow C_4$= 0.036 with confidence =0.92

viii. Relevance degree for the relationship $C_4 \rightarrow C_2$:
$Q_2 \rightarrow Q_4$: 0.1 ×1×1=0.1
Then the relevance degree for the relationship $C_4 \rightarrow C_2$= 0.1 with confidence =1

ix. Relevance degree for the relationship $C_4 \rightarrow C_5$:
$Q_2 \rightarrow Q_3$: 0.1 ×1×0.9=0.09
Then the relevance degree for the relationship $C_4 \rightarrow C_5$= 0.09 with confidence =0.9

[Step 5] Based on matrix QC, we create a new concept-concept matrix C' and find the minimum nonzero value, as shown follows:

$$C' = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{array} \begin{array}{c} \begin{array}{ccccc} C_1 & C_2 & C_3 & C_4 & C_5 \end{array} \\ \begin{bmatrix} 0.0 & 0.2 & 0.2 & 0.1 & 0.2 \\ 0.2 & 0.0 & 0.4 & 0.1 & 0.7 \\ 0.2 & 0.4 & 0.0 & 0.1 & 0.4 \\ 0.1 & 0.1 & 0.1 & 0.0 & 0.1 \\ 0.2 & 0.7 & 0.4 & 0.1 & 0.0 \end{bmatrix} \end{array}, \mu=0.1$$

[Step 6] For each relationship "$C_i \rightarrow C_j$" obtained in Step 4 we calculate the relative questions-concepts value between concepts "C_i" and "C_j" as shown in Table 6.

[Step 7] Finally, Based on questions-concept status shown in Table 6, we add an arrow from concept C_i to concept C_j associated with the relevance degree and then we get the complete concepts-relationship map, as shown in Fig. 6.

Table 6: The status of concepts-relationship

Concept-relationship	*Relationship status*
$C_1 \rightarrow C_2$	Keep
$C_1 \rightarrow C_4$	Delete
$C_1 \rightarrow C_5$	Keep
$C_2 \rightarrow C_5$	Keep
$C_3 \rightarrow C_1$	Delete
$C_3 \rightarrow C_2$	Keep
$C_3 \rightarrow C_4$	Delete
$C_4 \rightarrow C_2$	Delete
$C_4 \rightarrow C_5$	Keep

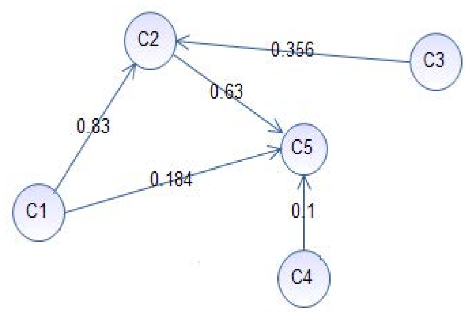

Fig.6 The complete concepts-relationship map

6. Conclusions

This study proposed an innovative approach to automatically construct concept maps. Most proposed methods [1-6, 8, 10, 11, and 12] deal only with binary grades of each test item; however, in a real assessment environment, a short answer also uses to assess the learning understanding. Thus, many studies [15, 16] apply fuzzy set theory to construct concept maps, but they do not take in consideration the weight of concept in each question and this might be a cause to construct incorrect relationships or exaggerate the degree of relationship. Therefore, to cope with this problem, the study proposes a different approach to construct concept maps based on analyses students' grades.

Firstly, we use look ahead fuzzy association rule mining algorithm [16] to mind some information about the relationships between questions, then we construct the questions-relationships mapping [3]. After that, we calculate the relevance degree between concepts in each question to obtain the final concept maps.
In prospect, we intend to integrate this approach with more efficient fuzzy association rules algorithm.

References

[1] S.M. Bai, and S.M. Chen, "A New Method for Learning Barriers Diagnosis Based on Fuzzy Rules", In: Proceedings of the Seventh International Conference on M achine Learning and Cybernetics, 2008, pp.3090 – 3095

[2] S.M. Bai, and S. M. Chen, "A New Method for Automatically Constructing Concept Maps Based on D ata Mining Techniques", In: Proceedings of the Seventh International Conference on M achine Learning and Cybernetics, 2008, pp. 3078 - 3083.

[3] S.M., Chen, and S. M. Bai, "Using data mining techniques to automatically construct concept maps for adaptive learning systems", Expert Systems with Applications, Vol. 37, 2010, pp. 4496–4503.

[4] S.M. Chen, and P.J. Sue, "A New Method to Construct Concept Maps for Adaptive Learning Systems", In: Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, 2010, pp. 2489 - 2494

[5] H.C. Chu, G.J. Hwang, and Y.M. Huang, "An Enhanced Learning Diagnosis Model Based on Concept-Effect Relationships with Multiple Knowledge Levels", Innovations in Education and Teaching International, Vol. 47, No. 1, 2010, pp. 53 - 67.

[6] H.C. Chu, G.J. Hwang, Judy C. R. Tseng, and G.H. Hwang, "A Computerized Approach to Diagnosing Student Learning Problems in Health Education", Asian Journal of Health and Information Sciences, Vol. 1, No. 1, 2006, pp. 43-60.

[7] E. C. Woodcoc, and G. Sucharita, "Fuzzy Set Theory and Thematic Maps: Accuracy Assessment and Area Estimation", Int. J. geographical information science, Vol. 14, No. 2, 2000, pp.153-172.

[8] K. Günel, and R. Aşliyan, "Determining Difficulty of Questions in Intelligent Tutoring Systems", The Turkish Online Journal of Educational Technology – TOJET ISSN: 1303-6521, Vol. 8, N.3, 2009, pp. 14-21.

[9] C.-L. Hung, and Y.W. Hung, "A Practical Approach for Constructing an Adaptive Tutoring Model Based on Concept Map", In: VECIMS 2009 - International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems Hong Kong, China, 2009, pp. 298 - 303.

[10] G.J. Hwang, C.L. Hsiao, and Judy C. R. Tseng, "A Computer-Assisted Approach to Diagnosing Student Learning Problems in Science Courses", Journal of Information Science and Engineering, Vol.19, 2003, pp.229-248.

[11] G.J. Hwang, Judy C. R. Tseng, and G.H. Hwang, "Diagnosing Student Learning Problems Based on Historical Assessment Records", Innovations in Education and Teaching International, Vol.45, No.1, 2008, pp. 77 – 89.

[12] C. H. Lee, G. G. Lee, and Y. H. Leu, "Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning", Expert Systems with Application, Vol.36, No. 2 , 2009, pp. 1675-1684.

[13] P. Panjaburee, G. J. Hwang, W. Triampo, and B.Y. Shih, "a multi-expert approach for developing testing and diagnostic systems based on t he concept-effect model", Computers & Education, Vol.55, 2010, pp. 527-540.

[14] K. Tatsuoka, "Rule space: An approach for dealing with misconceptions based on item response theory", Journal of Educational Measurement, Vol. 20, No. 4, 1983, pp. 345-354.

[15] C. J.Tsai, S. S.Tseng, and C.Y. Lin, "A two-phase fuzzy mining and learning algorithm for adaptive learning environment", In: Proceedings of the International Conference on C omputational Science (ICCS'01), Lecture Notes in Computer Science (LNCS 2074) CA, USA, Vol. 2 , 2001, pp. 429–438.

[16] S.S. Tseng, .S. P.C. Sue, J.M. Su, J.F. Weng, and W.N. Tsai,"A new approach for constructing the concept map", Computers & Education, Vol. 49, 2007, pp. 691–707.

Mohammed Al-Sarem received: The Research Master's in Computer Engineering and Information Technology from Volgograd State Technical University, Russia in 2007, and a BSc. in Computer Science in 2005 from the same university. He is conducting, since 2010, doctoral research in Computer Science in University of Hassan II, Mohammadia, Morocco, at e-learning systems: Models and methods of providing intellectual adaptive learning guidance in e-learning environment.

Mostafa Bellafkih received: PhD thesis in Computer Science from the University of Paris 6, France, in June 1994 and Doctorate Es Science in Computer Science (option networks) from the University of Mohammed V in Rabat, Morocco, in May 2001. His research interests include the network management, knowledge management, A.I., Data mining and Database. He is Professor in The National Institute of Posts and Telecommunications (INPT) in Rabat, Morocco since 1995.

Mohammed Ramdani received: The PhD thesis in Computer Science from the University of Paris 6, France, in February 1994 and Habilitation in Computer Science from the University of Paris 6, France, in June 2001. His research interests include the, knowledge management, A.I., Data mining and D atabase. He is Professor in Mohammedia Faculty of Sciences and Technologies (FSTM), Morocco since 1995.