Constructing concept maps for adaptive learning systems based on data mining techniques

Shyi-Ming Chen *, Po-Jui Sue

Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, ROC

Abstract

In this paper, we propose a new method for automatically constructing concepts maps for adaptive learning systems based on data mining techniques. First, we calculate the counter values between any two questions, where the counter values indicate the answer-consistence between any two questions. Then, we consider four kinds of association rules between two questions to mine some information. Finally, we calculate the relevance degree between two concepts derived from the association rule to construct concept maps for adaptive learning systems. The proposed method can overcome the drawbacks of Chen and Bai’s (2010) and Lee et al.’s method (2009). It provides us with a useful way to construct concept maps for adaptive learning systems based on data mining techniques.

Keywords:
Adaptive learning systems
Association rules
Concept maps
Data mining
Questions-relationship map

1. Introduction


In this paper, we propose a new method for automatically constructing concepts maps for adaptive learning systems based on data mining techniques. First, we calculate the counter values between any two questions, where the counter values indicate the answer-consistence between any two questions. Then, we consider four kinds of association rules between two questions to mine some information. Finally, we calculate the relevance degree...
between two concepts derived from the association rule to construct concept maps for adaptive learning systems. The proposed method can overcome the drawbacks of Chen and Bai’s method (2010) and Lee et al’s method (2009). It provides us with a useful way to construct concept maps for adaptive learning systems based on data mining techniques.

2. A review of Chen and Bai’s method for constructing concept maps based on data mining techniques

Chen and Bai (2010) presented a method for automatically constructing concept maps based on data mining techniques. In the following, we briefly review Chen and Bai’s method (2010). Assume that there are n learners S1, S2, ..., Sn, m questions Q1, Q2, ..., Qm, and p concepts C1, C2, ..., Ck, then we can transform the test portfolio of the learners and the degree of relevance between test questions and concepts into the grade matrix G and the questions-concepts matrix QC, respectively. Let Qj denote the ith question, where 1 ≤ i ≤ m, and let Sk denote the jth learner, where 1 ≤ j ≤ n. Then, we can get the grade matrix G, shown as follows:

\[
G = \begin{bmatrix}
S_1 & S_2 & \ldots & S_n \\
Q_1 & g_{11} & \ldots & g_{1n} \\
Q_2 & g_{21} & \ldots & g_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
Q_m & g_{m1} & \ldots & g_{mn}
\end{bmatrix}
\]

where \(g_{ij} \in \{0, 1\}\), \(g_{ij} = 1\) denotes the learner \(S_k\) gets the right answer in question \(Q_j\), \(g_{ij} = 0\) denotes the learner \(S_k\) has a wrong answer in question \(Q_j\), where 1 ≤ i ≤ m and 1 ≤ j ≤ n. In the same way, we can construct the questions-concepts matrix QC, shown as follows:

\[
QC = \begin{bmatrix}
C_1 & C_2 & \ldots & C_k \\
Q_1 & q_1C_1 & \ldots & q_1C_k \\
Q_2 & q_2C_1 & \ldots & q_2C_k \\
\vdots & \vdots & \ddots & \vdots \\
Q_m & q_mC_1 & \ldots & q_mC_k
\end{bmatrix}
\]

where \(q_{ij} \in \{0, 1\}\), \(q_{ij} = 1\) denotes the degree of relevance of question \(Q_j\) with respect to concept \(C_i\) and 0 ≤ \(q_{ij} \leq 1\). In the following, we briefly review Chen and Bai’s (2010) as follows:

Step 1: Based on the grade matrix G and the Apriori algorithm (Agrawal & Srikant, 1994), mining two kinds of association rules between questions:

(1) if question \(Q_i\) is correctly learned by the learner, then question \(Q_i\) is also correctly learned by the learner.

(2) if question \(Q_i\) is incorrectly learned by the learner, then question \(Q_i\) is also incorrectly learned by the learner.

Construct the association rules from the large 1-itemset with respect to the other questions and calculate the confidence of the rules, where the confidence “\(\text{conf}(Q_i \rightarrow Q_j)\)” of an association rule “\(Q_i \rightarrow Q_j\)” is calculated as follows:

\[
\text{conf}(Q_i \rightarrow Q_j) = \frac{\text{sup}(Q_i, Q_j)}{\text{sup}(Q_i)},
\]

where \(Q_i\) is a question in the large 1-itemset, \(Q_j\) is a question in the test paper, “\(Q_i \rightarrow Q_j\)” denotes the association rule from \(Q_i\) to \(Q_j\), “\(\text{conf}(Q_i \rightarrow Q_j)\)” denotes the confidence of the association rule “\(Q_i \rightarrow Q_j\)”, “\(\text{sup}(Q_i, Q_j)\)” denotes the support of the 2-itemset \(\{Q_i, Q_j\}\), “\(\text{sup}(Q_i)\)” denotes the support of the large 1-itemset \(Q_i\), where 1 ≤ i ≤ m, 1 ≤ j ≤ m and \(i \neq j\).

Step 2: Based on the association rules obtained in Step 1, construct two kinds of questions-relationship maps. For the association rules that question \(Q_i\) is correctly learned by the learner and question \(Q_j\) is also correctly learned by the learner, build the relationship from question \(Q_i\) to question \(Q_j\) in the “correct-to-correct questions-relationship map” associated with the confidence. For the association rules that the learner failed question \(Q_i\) and failed question \(Q_j\), build the relationship from question \(Q_i\) to question \(Q_j\) in the “failure-to-failure questions-relationship map” associated with the confidence.

Step 3: Convert the two kinds of question-relationship maps obtained in Step 2 into the concept-relationship table. Calculate the relevance degree \(\text{rev}(C_i \rightarrow C_j)\) between concepts \(C_i\) and \(C_j\) from the relationship \(Q_i \rightarrow Q_j\), shown as follows:

\[
\text{rev}(C_i \rightarrow C_j) = \min\left(W_{Q_i,C_i}, W_{Q_j,C_j}\right) \times \text{conf}(Q_i \rightarrow Q_j),
\]

where “\(\text{rev}(C_i \rightarrow C_j)\)” denotes the relevance degree of the relationship “\(C_i \rightarrow C_j\)” derived from the relationship “\(Q_i \rightarrow Q_j\)”, \(\text{rev}(C_i \rightarrow C_j) \in [0, 1]\). \(C_i\) denotes a concept appearing in the question \(Q_i\), \(C_j\) denotes a concept appearing in the question \(Q_j\), \(W_{Q_i,C_i}\) denotes the weight of concept \(C_i\) in question \(Q_i\), \(W_{Q_j,C_j}\) denotes the weight of concept \(C_j\) in question \(Q_j\), “\(\text{conf}(Q_i \rightarrow Q_j)\)” denotes the confidence of the relationship “\(Q_i \rightarrow Q_j\)”, 1 ≤ i ≤ p, 1 ≤ j ≤ p, \(i \neq j\), 1 ≤ x ≤ m, 1 ≤ y ≤ m and \(x \neq y\). If there are more than one relationship between any two concepts, then choose the relationship which has the maximum relevance degree between them.

Step 4: Combine the concepts-relationship tables obtained in Step 3 into the combined concepts-relationship table, described as follows:

(1) If the relationship “\(C_i \rightarrow C_j\)” only exists in one of concepts-relationship tables, then put it to the combined concepts-relationship table.

(2) If the relationship “\(C_i \rightarrow C_j\)” exists in both concepts-relationship tables, then calculate the difference degree \(\frac{|\text{rev}(C_i \rightarrow C_j) - \text{rev}(C_i \rightarrow C_j)|}{\max(\text{rev}(C_i \rightarrow C_j))\times \min(\text{rev}(C_i \rightarrow C_j))}\), where “\(\text{rev}(C_i \rightarrow C_j)\)” denotes the relevance degree of the relationship \(C_i \rightarrow C_j\) in the “correct-to-correct concepts-relationship table”, “\(\text{rev}(C_i \rightarrow C_j)\)” denotes the relevance degree of the relationship \(C_i \rightarrow C_j\) in the “failure-to-failure concepts-relationship table”, 1 ≤ i ≤ p and 1 ≤ j ≤ p. If the difference degree is larger than the threshold value \(\lambda\), where 0 ≤ \(\lambda\) ≤ 1, then delete the relationship \(C_i \rightarrow C_j\). Otherwise, choose the relationship which has the largest relevance degree between the concepts \(C_i\) and \(C_j\).

However, Chen and Bai’s method (2010) has the drawback that it does not correctly construct concept maps in some situations. In the following, we use an example to show the drawback of Chen and Bai’s method. Assume that there are four learners S1, S2, S3, S4, three questions Q1, Q2, Q3 and two concepts C1, C2. Assume that the test portfolio of the learners is represented by a matrix A, shown as follows:

\[
A = \begin{bmatrix}
Q_1 & 1 & 1 & 0 & 0 \\
Q_2 & 1 & 1 & 1 & 0 \\
Q_3 & 1 & 1 & 0 & 0
\end{bmatrix}
\]
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات