A Tutoring Rule Selection Method for Case-based e-Learning by Multi-class Support Vector Machine

Daichi Hisakane, Minami Otsuki, Masaki Samejima and Norihisa Komoda
Graduate School of Information Science and Technology, Osaka University, 2-1, Yamadaoka, Suita, Osaka, Japan

Keywords: Intelligent Tutoring System, Multi-class SVM(Support Vector Machine).

Abstract: We develop an intelligent tutoring system on learners' answers to problems that are dealt with in case-based e-learning. A facilitator instantiates answers and tutoring advice as a tutoring rule preliminary, and the system automatically identifies an appropriate instantiated answer which corresponds to the input sentence of an answer from the learner. Although various kinds of tutoring rules are given on a certain problem, the instantiated answers are very similar to each other among tutoring rules, even if tutoring rules are different. So the input sentence is similar to the wrong instantiated answer of the tutoring rule, which makes it difficult to select the tutoring rule correctly. The proposed method selects the tutoring rule for the input sentence by machine learning of selecting the tutoring rules with the multi-class SVM(Support Vector Machine). The multi-class SVM, consisting of multiple binary classifiers, can output various tutoring rules identified as corresponding to one input sentence. In order to identify one correct tutoring rule, the proposed method introduces confidence on each identification result and integrates the results. The proposed method improves accuracy of selecting tutoring rules by 17% compared to the similarity-based selection method of tutoring rules.

1 INTRODUCTION

Case-based learning is an educational method to equip people with knowledge to solve problems(Hoffmann and Ritchie, 1997). In the case-based learning, an instructor selects a documented case that is useful for developing learners' problem-solving skill. The learners read the documented case and think what are problems and how to solve the problems. When the learners can not answer solutions to the problem correctly, the instructor does not give the correct answer but gives a hint to derive the correct answer in order to develop their thinking process(Savery, 2006).

However, because the instructor is lacking for the learner, the learners can not always receive advice from the instructor, which decreases their learning effect. If the learner can get appropriate advice without the instructor as well as the instructor gives, the instructor's workload can be reduced, and the learner can learn from the case-based learning by themselves. So, the intelligent tutoring system that gives advice automatically is valuable for both the instructor and the learner. Therefore, we propose an intelligent tutoring system that enables learners to be engaged in case-based learning without the instructor.

Intelligent Tutoring Systems(ITS) have been proposed to give advice to a learner(Corbett et al., 1997). In order to show advice, ITS have several kinds of tutoring rules that are defines as IF-THEN rules. The tutoring rules represent pairs of learners' answer and an instructor's advice to the answer. ITS need to select the appropriate tutoring rules by analyzing the answer from the learner based on the domain knowledge that is related to the learning contents. In this paper, we address tutoring rule selection for case-based e-Learning.

The rest of the paper is organized as follows. Section 2 describes the literature review of the intelligent tutoring system. Section 3 outlines case-based e-Learning systems and shows the research issues. Section 4 describes the selection method of tutoring rules. Section 5 shows the experimental results in applying the proposed method to the real data from learners. Section 6 deals with the conclusion derived from the experimental results.

2 LITERATURE REVIEW

Learner support in e-Learning is important because it influences learning outcomes(Aleven et al., 2003).

Hisakane D., Otsuki M., Samejima M. and Komoda N..
A Tutoring Rule Selection Method for Case-based e-Learning by Multi-class Support Vector Machine.
DOI: 10.5220/0005023501190125
In *Proceedings of the International Conference on Knowledge Management and Information Sharing* (KMIS-2014), pages 119-125
ISBN: 978-989-758-050-5

Online individual tutoring with human tutor via Internet is the most effective support. However learners always cannot receive such tutors' support unless the tutor stays for the learners. On the other hand, there is also an e-Learning system that sends learner's answers to a tutor in a remote area(Chen and Hsu, 2005). Checking the answers, the tutor gives the appropriate advice for the answer. The system has functions of displaying Q&A and pop up hint, and exchanging messages with other learners for developing learning ability.

However, a learner does not still obtain advice in a moment and this system does not reduce a tutor's burden, because the human tutor must think and send the advice for a learner's answer. The system has not solved the problems of the limitation of the tutors and obtaining advice to the learner in a moment.

Currently many studies on Intelligent Tutoring System (ITS) have been conducted. In ITS, a programmed tutor agent teaches learners instead of a human tutor. The tutor agent provides advice to learners based on the learner model in order to provide appropriate advice to each learner. The learner model consists of learner properties, degree of understanding and the degree of guesswork. ITS estimates the learner model based on time taken to solve and which answer is correct and so on(Sobue et al., 2004; Santos and Jorge, 2013). The learner property and the degree of understanding are estimated by analyzing the history of learning statistically. For example, the degree of understanding can be calculated as a rate of the correct answers to all the answers in case of that the problem has clear and deterministic answers as well as the multiple choice questions. In fact, e-Learning systems target not only the problem having deterministic answers but also the problems having non-deterministic answers:

- Mathematics
  ITS for mathematics have been developed in the past(Heffernan and Koedinger, 2000; Virvou and Sidiropoulos, 2013; Pholo and Ngwira, 2013) which enables the dialogue as well as a human tutor. By inputting questions with natural language, the learner can understand mathematical problems through dialogue with the system.
- Circuit design
  There has been the system that accepts inputs of free descriptions of circuits and gives advice on the circuit design(Dzikovska et al., 2010). The system simulates circuit action and displays hints according to student's input in circuit education.

As the above systems deal with non-deterministic answers, the case-based e-Learning system also deals with non-deterministic answers that are inputted as natural language by the learners. However, generally in case-based e-Learning, the learners input answers from various aspects. So, our case-based e-Learning system must equip the function to give advice for various answers.

3 CASE-BASED E-LEARNING SYSTEM

3.1 Outline of the Case-based E-Learning System

The GUI(Graphical User Interface) of this e-learning system is shown in Figure 1. Case description is shown on the upper-left text area in this window, and the learner thinks what are the problems and how to solve the problems from this case description. Then the learner inputs the problems and the solutions into the answer from at the bottom of the window. The advice corresponding to the answers is indicated in the upper-right text area in the window. The instructor of the case-based learning empirically judges advice appropriately corresponding to the learner's answer.

So we propose a rule-based system to give appropriate advice based on the instructor's experience. The instructor instantiates answers and appropriate advice to the answers. Based on the pair of the instantiated answers and the appropriate advice, we make a tutoring rule as the following IF-THEN rule: IF the input sentence corresponds to the instantiated answer THEN the system outputs the advice in the pair. For an example in Figure 1, IF the input sentence corresponds to "I don't know." THEN the system outputs "What happens when Project Manager(PM) makes members work late?"

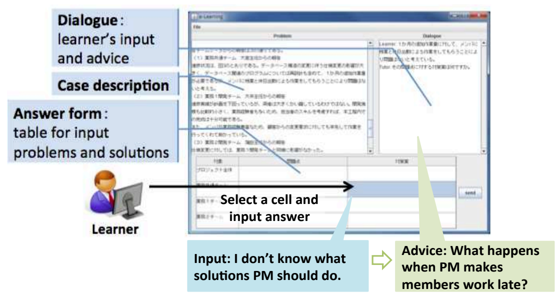

Figure 1: GUI of case-based e-learning system.

Figure 2 is an outline of our proposed system. The input of this system is an answer from a learner(hereafter called input sentence), and the output of this system is advice corresponding to the input sentence. The instructor preliminary registers tutoring rules in a data base(hereafter called tutoring rule DB). Due to various kinds of input sentences, the instructor sets multiple instantiated answers on each tutoring rule. For an example in Figure 2, the Rule B has instantiated answers of “Delivery delay”, “Stagnant work” and so on. In addition, the instructor can not instantiate answers that cover all the input sentences. In order to give advice for the input sentences that do not correspond to any instantiated answers, the system collects such input sentences from history of the learners’ inputs. And the collected past input sentences are registered into the tutoring rule of ‘Others’ in order to identify the input sentences that do not correspond to any tutoring rules.

The intelligent tutoring system judges whether the input sentence corresponds to the instantiated answer in each tutoring rule DB or not. Then the system selects a tutoring rule which includes the instantiated answer corresponding to the input sentence and shows advice corresponding to the tutoring rule. If the input sentence corresponds to the tutoring rule of ‘Others’, the system shows a message to induce the learner to change the input sentence by rephrasing the expression. The learner reads the shown advice and inputs the answer repeatedly until the learner inputs a completely correct solution to the problem. The system has to select the correct tutoring rule corresponding to the input sentence to show appropriate advice to learners.

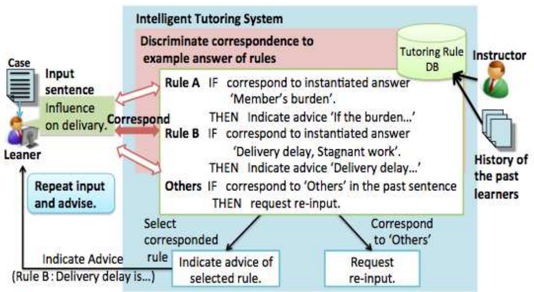

Figure 2: Case-based e-learning system.

3.2 Research Issue

Similar input sentences tend to correspond to the same advice in the tutoring rule. So a typical approach to judge the correspondence between 2 sentences is to measure the similarity between the sentences. Even if the sentences are the same in meaning, different similar words are often used in the input sentence and the instantiated answers. In order to judge the similarity between the sentences, it is necessary to judge the different similar words. For this problem, recognizing textual entailment (RTE) has been developed(Mark et al., 2011). RTE-based method selects the rule by the similarity between the input sentence and each instantiated answer by identifying the similar words with the conceptual dictionary WordNet(Jung and VanLehn, 2010).

Figure 3 shows the tutoring rule selection by similarity and its issue. The similarity is calculated by Jaccard coefficient(Broder, 1997). In comparing a sentence A to a sentence B, let a, b, c denote the number of words in sentence A, the number of words in sentence B, and the number of common words in both sentences. The Jaccard coefficient is decided by the following formula:

$$Jaccard\ coefficient = \frac{c}{a+b-c} \qquad (1)$$

After the similarities to all the instantiated answers are calculated, the tutoring rule that has most similar instantiated answer is selected for advice.

However, the instantiated answers in different tutoring rules for the same problem are similar to each other. As shown in Figure 3, the input sentence is similar to not only the instantiated answer in the tutoring rule corresponding to the input sentence but also one incorresponding to the input sentence. Therefore, the conventional method often selects the wrong tutoring rule.

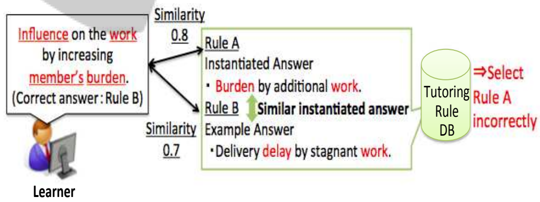

Figure 3: Selection of tutoring rule by similarity and an issue.

4 SELECTION METHOD OF TUTORING RULE

4.1 Outline of the Method

We focus on that the instantiated answers correctly correspond to the correct advice in the tutoring rule. So our proposed method learns how to select the tutoring rule based on the similarity between the instantiated answers as supervised data. As a learning classifier, we use 1 vs 1 multi-class SVM(Support Vector Machine) which indicates good performance of discriminating various input sentences for the correct tutoring rule(Brunner et al., 2012). 1 vs 1 multi-class SVM consists of 2-class SVM that are learned with supervised data on 2 classes, and integrates the tutoring rules identified as corresponding to the input sentence by each SVM.

In integrating the identified rules by all SVM, the proposed method emphasizes on the identified tutoring rule that is considered to correspond to the input sentence more correctly. Furthermore, the method selects the tutoring rule using the instantiated answer which has similar confidence to the confidence of the input sentence.

Figure 4 shows the outline of the tutoring rule selection method. This method calculates the confidence of the identified rule by discriminating instantiated answers to each tutoring rule by applying multi-class SVM to the input sentence and the instantiated answer in the tutoring rule DB. When different tutoring rules have similar instantiated answers each other, the confidence of identified input sentence also increases on both of tutoring rules. So our method selects the tutoring rule corresponding to the instantiated answer by identifying the instantiated answer which has similar confidence to the confidence of the input sentence.

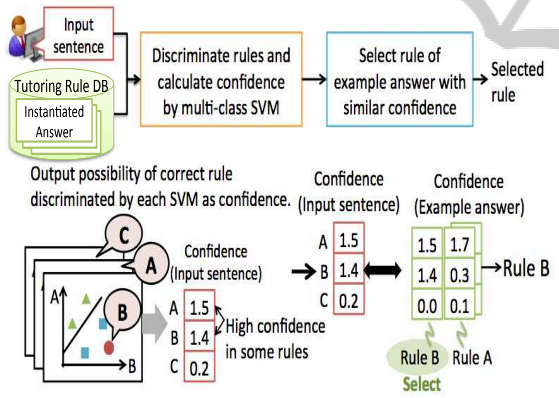

Figure 4: Outline of the method of selecting tutoring rule.

4.2 Confidence Estimation in Discriminating Input Sentence

For the sake of calculating the confidence of the discrimination with a criterion of the similarity, the proposed method learns how to select the tutoring rule by the similarity between words included in instantiated answers.

In learning from instantiated answers, the similarity between words in 2 instantiated answers is calculated with the conceptual dictionary WordNet(Bond et al., 2012; Isahara et al., 2008). As Figure 5 shows, the WordNet consists of associated conceptual groups 'synset'. Each synset represents a concept, and has higher and lower rank groups(Leacock and Chodorow, 1998).The similarity between the word w_m and w_n is calculated by the following formula:

$$sim(w_m, w_n) = \frac{\frac{L_m}{L_1} + \frac{L_n}{L_2}}{L_m + L_n} \qquad (2)$$

where L_1 and L_2 are the route length from the synset of w_m and w_n to the upper synset where links from synsets of w_m and w_n join together. And L_m and L_n are the route length from the synset of w_m and w_n to the uppermost synset linked from w_m and w_n.

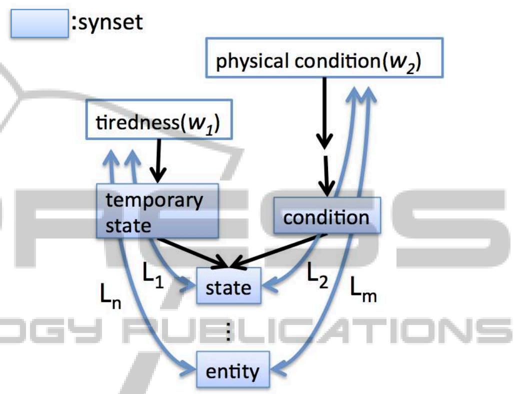

Figure 5: Conceptual groups in WordNet.

The feature vector is defined as similarities between the words. Let F denote the feature vector of the similarities s_{ij} between words w_i in the input sentence and w_j in instantiated answers in a tutoring rule.

$$F = \{s_{ij}\} \qquad \forall i \in I, j \in J \qquad (3)$$

where I and J are a set of words in the input sentence and a set of words in all the instantiated answers in a tutoring rule. The dimensionality is reduced by principal component analysis in order to use only the characteristic similarities(Ian, 2005).

The distance from hyperplane on the multi-class SVM is generally used as the confidence in discriminating tutoring rules. Figure 6 shows the confidence calculation of discriminating learners answer. SVMs as binary classifiers learn with instantiated answers for 2 kinds of tutoring rules that are selected from the tutoring rule DB. This learning process is applied to all combination of the tutoring rules, and 1 vs 1 multi-class SVM is created. Then the confidence is estimated by classifying input sentences and instantiated answers.

The bottom of Figure 6 shows the flow of discriminating input sentence to each tutoring rule. Each SVM is applied to the feature vector of the similarity whose dimensionality is reduced by principal component analysis. SVM has high confidence of discriminated input sentence when the identified result is far away from the hyperplane. So the confidence is defined as the distance from the hyperplane and the feature vector of the confidence is defined as the total distance on SVMs that discriminates the input sentence to the tutoring rule. For example, a classifier for tutoring rules of A and B, and a classifier for tutoring rules of A and C discriminate the input sentence to A with confidence of 0.2 and 0.3. The total confidence 0.5 is used to select tutoring rule A. The feature vector of the confidence of all instantiated answers is also calculated by the same way of calculating feature vector of confidence of the input sentence.

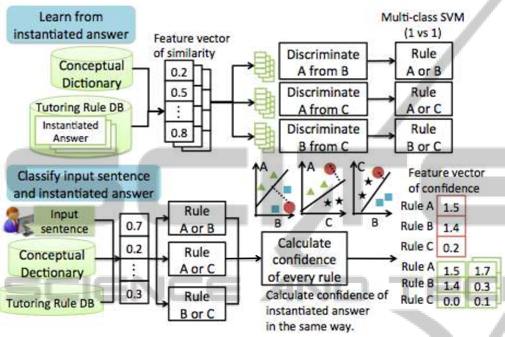

Figure 6: Confidence calculation of discriminating learner's answer.

4.3 Tutoring Rule Selection by Classifying Confidence

The tutoring rule is selected by the similarity of feature vectors of the confidence between the input sentence and the confidence of the instantiated answer. Figure 7 shows the outline of selecting tutoring rules. As Figure 7 shows, distance is small between the feature vectors of confidences on the instantiated answers that correspond to the same tutoring rule. The feature vectors of instantiated answers consist clusters that correspond to the tutoring rule.

Therefore, we can formulate the selection problem of tutoring rules as classifying the feature vector of input sentence to the cluster of the instantiated answers. So our method selects the tutoring rule as a major instantiated answer neighbor the input sentence by k-NN method on the confidence(Duda and Heart, 1973). K-NN method needs to set the number of the neighborhood k. The number of the instantiated answers is set based on the number of instantiated answers r. The value of k is 1 at a minimum, and the number of the instantiated answers r at a maximum. Therefore, the range of k is set as follows:

$$1 \leq k \leq r \tag{4}$$

Because it is possible to calculate the rate of the number of instantiated answers assigned to the correct cluster of tutoring rule as precision, our method selects k that indicates maximum precision rate within $1 \leq k \leq r$. If some k have the maximum precision rate, the largest k is used for selecting tutoring rules. In Figure 7, the value of k is set as $k = 3$. Then the tutoring rule is selected as a major instantiated answer in k instantiated answers.

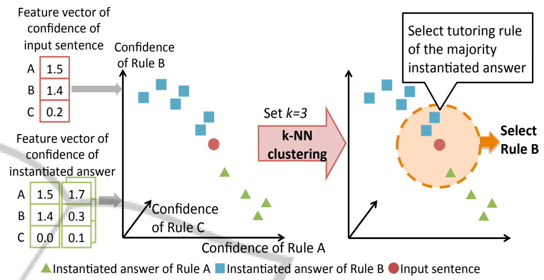

Figure 7: Tutoring rule selection by classifying confidence.

5 EXPERIMENT

5.1 Outline of the Experiment

In this experiment, we have developed the case-based e-learning system and collected answers for a documented case about project management from 20 students majoring in Information Science. The number of collected answers were 82 sentences. The document of this case describes project manager's decision-making on the project delay due to the additional development required by the customer. The instructor preliminary set 4 problems shown in the following. The number of tutoring rules, instantiated answers, and input sentences in each problem are shown in Table 1.

- **Problem 1.** How to solve the problem of dealing with change request from the customer without manager's authorization.
- **Problem 2.** How to solve the problem of remaining a difficult work later during the project duration.
- **Problem 3.** How to solve the problem in dealing with additional work.
- **Problem 4.** How to solve the problem to speed-up development by working overtime and on a day off.

In the proposed method, the principal component analysis for reducing dimension is implemented by Weka (Hall et al., 2009). And multi-class SVM for learning instantiated answer is implemented by SVM-Light (Kanungo and Joachims, 1999).

Table 1: The number of tutoring rules, instantiated answers, and input sentences.

	Problem 1	Problem 2	Problem 3	Problem 4
Tutoring rules	4	5	5	7
instantiated answers	21	26	43	82
Input sentences	15	20	20	27

5.2 Evaluation of the Proposed Method

For evaluating the proposed method, we compared the accuracy rate of the tutoring rule selections, by the maximum similarity between input sentence and instantiated answer, by the maximum confidence calculated by multi-class SVM, and by the proposed method. The accuracy rate is the rate of selecting correct tutoring rules from input sentences. Together with the learners, we judge whether the selected tutoring rule is correct or not. In applying the proposed method to give advice to a certain learner, the proposed method uses the answer from the other learner as instantiated answers. Figure 8 shows the accuracy rate of the whole problems and each problem by each method.

The accuracy rate of tutoring rule selection by maximum similarity was 63%(the number of correct selection was 52), by confidence calculated by multi-class SVM was 71%(the number of correct selection was 58), and by the proposed method was 80%(the number of correct selection was 66). The proposed method improves the accuracy rate by 17%(the number of correct selection was 14) compared to the method by the similarity.

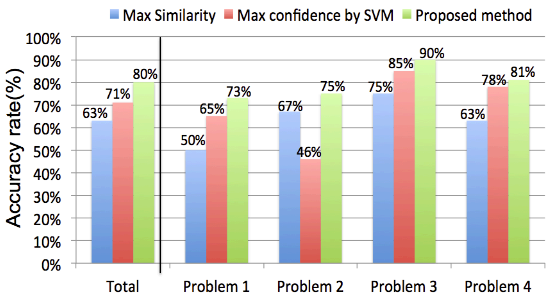

Figure 8: Accuracy rate.

According to Figure 8, the proposed method particularly improves the accuracy rate or Problem1 and 4 compared to the method of selecting tutoring rule by similarity. In Problem 1 and 4, the proposed method can select the correct advice from the input sentence from which the selection method by similarity selects the wrong tutoring rule because many input sentences have high confidence in some tutoring rules. For example, Problem 4 has 7 tutoring rules, and 2 of the 7 tutoring rules include the following words:

- Tutoring rule 1Stagnant work
- Tutoring rule 2Delivery delay

Although the input sentence 'Stagnant work caused by tired people fatigue leads to the delivery delay.' corresponds to Tutoring rule 2, it is judged to correspond to Tutoring rule 1 by the similarity-based selection method. The input sentence has a meaning of the Tutoring rule 1 and 2, but the proposed method correctly selects Tutoring rule 2.

6 CONCLUSION

In this paper, we proposed the intelligent tutoring system that gives an appropriate advice corresponding to the answer from a learner. This system needs to select the correct tutoring rule corresponding to the input sentence from the tutoring rules which consists of the instantiated rules and advices registered by the instructor.

Then we proposed the selection method of tutoring rules by the confidence of discriminating learner's answer. The proposed method selects the tutoring rule of the input sentence by multi-class SVM learning the tutoring rule selection based on the similarity of instantiated answer. The tutoring rule is selected by the confidence as distance from the hyperplane of each SVM. And, the proposed method integrated the discriminant results by the confidence.

In the experiment, we adopt the case of project management and got 82 sentences inputed by about 20 person who have experienced case-based Learning. The proposed method improves the accuracy rate by 17% compared to the selecting method by the maximum similarity and by 9% compared to the selection method by maximum confidence calculated by multi-class SVM.

REFERENCES

Aleven, V., Stahl, E., Schworm, S., Fischer, F., and Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73(3):277–320.

Bond, F., Baldwin, T., Fothergill, R., and Uchimoto, K. (2012). Japanese semcor: A sense-tagged corpus of japanese. *In Global WordNet Association*.

Broder, A. Z. (1997). On the resemblance and containment of documents. In *Proccedings of Compression and Complexity of Sequences*, pages 21–29.

Brunner, C., Fischer, A., Luig, K., and Thies, T. (2012). Pairwise support vector machines and their application to large scale problems. *Journal of Machine Learning Research*, 13:2279–2292.

Chen, C.-M. and Hsu, S.-H. (2005). Applying interactive mobile teaching agent to support e-learning platform for learning performance promotion. In *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies*, ICALT '05, pages 205–206, Washington, DC, USA. IEEE Computer Society.

Corbett, A. T., Koedinger, K. R., and Anderson, J. R. (1997). *Handbook of humancomputer interaction*. North Holland.

Duda, R. O. and Heart, P. E. (1973). *Pattern classification and scene analysis*. John Wiley & Sons.

Dzikovska, M. O., Bental, D., Moore, J. D., Steinhauser, N. B., Campbell, G. E., Farrow, E., and Callaway, C. B. (2010). Intelligent tutoring with natural language support in the beetle ii system. *Sustaining TEL: From Innovation to Learning and Practice*, pages 620–625.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18.

Heffernan, N. T. and Koedinger, K. R. (2000). Intelligent tutoring systems are missing the tutor: Building a more strategic dialog-based tutor. In *Proceedings of the 2000 AAAI Fall Symposium on Building Dialogue Systems for Tutorial Applications*, pages 14–19.

Hoffmann, B. and Ritchie, D. (1997). Using multimedia to overcome the problems with problem based learning. *Instructional Science*, 25(2):97–115.

Ian, J. (2005). *Principal component analysis*. John Wiley & Sons.

Isahara, H., Bond, F., Uchimoto, K., Utiyama, M., and Kanzaki, K. (2008). Development of the japanese wordnet. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*.

Jung, S. Y. and VanLehn, K. (2010). Developing an intelligent tutoring system using natural language for knowledge representation. *Lecture Notes in Computer Science*, 6095:355–358.

Kanungo, T. and Joachims, T. (1999). *Advances in Kernel Methods*. MIT Press.

Leacock, C. and Chodorow, M. (1998). Combining local context and wordnet similarity for word sense identification. *WordNet: An electronic lexical database*, 49(2):265–283.

Mark, S., Vydiswaran, V. V., and Roth, D. (2011). *Multilingual Natural Language Applications: From Theory to Practice*. IBM Press.

Pholo, D. and Ngwira, S. (2013). Integrating explicit problem-solving teaching into activemath, an intelligent tutoring system. In *International Conference on Adaptive Science and Technology (ICAST)*, pages 1–8.

Santos, G. S. and Jorge, J. (2013). Interoperable intelligent tutoring systems as open educational resources. *The IEEE Transactions on Learning Technologies*, 6(3):271–282.

Savery, J. R. (2006). Overview of problem-based learning: Definitions and distinctions. *Problem-based Learning*, 1(1).

Sobue, Y., Matsumoto, T., and Onishi, N. (2004). Construction of the statistical student model using learning history on web-based learning system. *IEICE Technical Report (Institute of Electronics, Information and Communication Engineers)*, 103(697):89–94.

Virvou, M. and Sidiropoulos, S.-C. (2013). An intelligent tutoring system over a social network for mathematics learning. In *Information, Intelligence, Systems and Applications (IISA)*, pages 1–4.