Computer-Aided Generation of Item Banks Based on Ontology and Bloom's Taxonomy

Ming-Hsiung Ying[1] and Heng-Li Yang[2]

[1] Department of MIS, Chung-Hua University, 707, Sec.2, WuFu Rd., HsinChu, Taiwan
[2] Department of MIS, National Cheng-Chi University, 64, Sec.2, Chihnan Rd., Taipei, Taiwan
mhying@chu.edu.tw, yanh@nccu.edu.tw

Abstract. Online learning and testing are important topics in information education. Students can take online tests to assess their achievement of learning goals. However, the test results should assign student scores and assess their achievement of knowledge and cognition levels. Teachers currently need to spend considerable time on producing and maintaining on-line testing items. This study applied ontology, Chinese semantic database, artificial intelligence and Bloom's taxonomy to propose a CAGIS E-learning system architecture to assist teachers in creating test items. As the result, the computer assisted teachers in producing a large number of test items quickly. These test items covered three types of knowledge and five dimensions of cognitive skills. The test items could meaningfully assess learning level meaningfully.

Keywords: Online Test, Test Item Bank, Bloom's Taxonomy, Ontology, Semantic Web.

1 Introduction and Related Works

Online learning and subsequent testing have been important topics in information education. Because education is intended to change students behaviors, teachers must use tests well to assess student achievements. Computer-based testing has numerous benefits, including data-rich test results, immediate test feedback, convenient test times and locations, and so on. [1].

In designing test items, teaching goals should be considered when designing test items. According to education testing theory, educational goals can be classified into three different levels: cognition field, emotional field and movement ability [2]. Types of instruction assessment can be grounded in types of knowledge. Three distinct knowledge types require assessment: declarative (knowing what/knowing about), procedural (knowing how), and conditional (knowing why and when) [3]. Bloom identified six levels within the cognitive domain, including knowledge, comprehension, application, analysis, synthesis and evaluation [4]. Anderson and Krathwohl [5] revised the original taxonomy of Bloom by combining both the cognitive process and knowledge dimensions. The revised Bloom's taxonomy comprises a two-dimensional table. One dimension identifies the knowledge (the kind of knowledge to be learned), while the other identifies the cognitive process (the process used to learn). The knowledge dimension comprises four levels: factual, conceptual, procedural, and

F. Li et al. (Eds.): ICWL 2008, LNCS 5145, pp. 157–166, 2008.

meta-cognitive. The cognitive process dimension comprises six levels: remember, understand, apply, analyze, evaluate, and create. This new expanded taxonomy can help instructional designers and teachers set meaningful learning objective, and provide the measurement tool for thinking.

Creating and maintaining the item bank is a time-consuming. When the item bank contains an insufficient number of items, the exposure frequencies of items may be too high and students may directly recall the answers [6]. Therefore, how to prepare sufficient items in the bank and efficiently generate items have become important research issues [7].

Deveszic [8] proposed developing Web-based educational applications with more theory and content-oriented intelligence. To increase the effectiveness of the testing system, numerous researchers have applied artificial intelligence, fuzzy theory and other techniques. If information techniques can be properly applied, numerous omplex issues can be solved, such as test item selection, item generation, scoring, explanation, and test feedback to enhance education and learning [9-15].

This study claims that computers can assist in aiding item generation in e-learning environments, if the material can be first stored based on knowledge ontological structure and semantic relation. An intelligent online learning system has been proposed to resolve the above problems.

2 Proposed System Architecture

To propose a system architecture for computer-aided tem bank generation, this study followed the following steps: (1) Conducting a pilot study to explore the difficulty faced by teachers in manually creating items, and analyzing the item types; (2) Developing course material knowledge and item structure ontologies, involving concept of Bloom's taxonomy; (3) Creating a knowledge base related to online course materials; (4) Developing a prototype for computer-aided generation of item system (CAGIS).

2.1 A Pilot Study Exploring the Difficulty of Manual Item Creation

Fifteen university teachers from 11 different universities - who had taught "management information system" courses, participated in the pilot study. These teachers were given two weeks to create test items from specific chapters of a textbook. It was required that the test items should include four types: true-false, multiple-choice, multiple-response, and fill-in-the-blank. No upper limited constrained the quantity of test items. Finally, the teachers produced 440 items manually, with the average time taken to complete the task being 4.3 hours. After deleting the duplicate items, there are 386 items left and shown in Table 1. The knowledge types of those items included “factual, conceptual, procedural” knowledge, and their cognitive levels included: “remember, understand, analyze, and evaluate”. The specific chapters are no suitable knowledge content to generate the item of "apply" level. Some teachers indicated that it would be very difficult to generate the "create" level items using true-false, multiple-choice, multiple-response, and fill-in-the-blank question type.

Table 1. Number of Items with Bloom's Taxonomy Produced by Teachers Manually

Knowledge Dimensions	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Total
Factual	192 (49.7%)	25 (6.5%)		56 (14.5%)	3 (0.8%)	276 (71.5%)
Conceptual	59 (15.3%)	27 (7.0%)		12 (3.1%)	0 (0%)	98 (25.4%)
Procedural	9 (2.3%)	0 (0%)		3 (0.8%)	0 (0%)	12 (3.1%)
Total	260 (67.3%)	52 (13.5%)	0 (0%)	73 (18.4%)	3 (0.8%)	386 (100%)

2.2 Course Material Knowledge Ontology

Since the meta-cognitive knowledge of Bloom's Taxonomy is not included in the regular teaching material or test [5,16], it was not considered in this study. To store knowledge content of course materials, and to consider the dimensions of Bloom's factual, conceptual, and procedural knowledge, this study developed a knowledge ontology, as shown in Fig. 1. This knowledge ontology was developed by content analysis of specific chapters from the above textbook, and includes the concepts of WordNet, revised Bloom's Taxonomy, Dublin Core, Semantic Header, and so on.

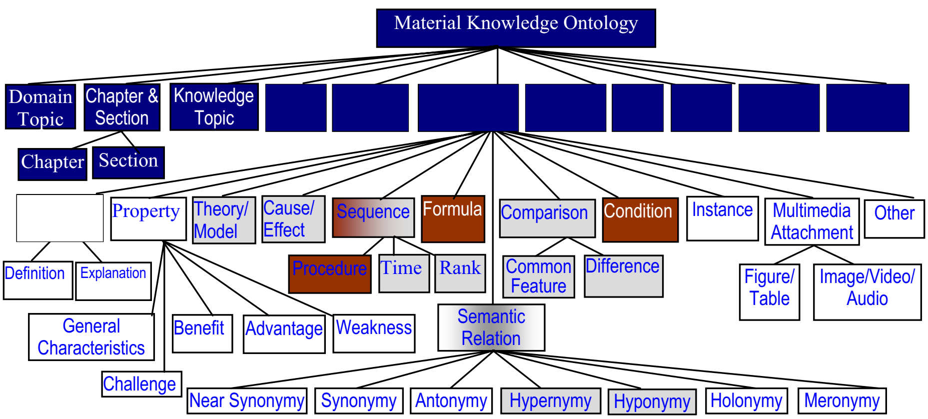

Fig. 1. Course Material Knowledge Ontology

Figure 1 uses the "Knowledge Content" to store the real course material content, and comprises 12 subclasses of knowledge, which are used to store knowledge concepts such as "What", "Why", "When" and "How". For example, sequence relation knowledge includes procedure (the procedural step, used to express the concept of "How"), time (the time sequence), rank (specific attribute rank). Hypernymy knowledge records a relation similar to generalization, is-a relation, is-a-kind-of. Meronymy knowledge records a relation similar to component-of.

The proposed course material knowledge ontology covers the knowledge dimension of Taxonomy of Bloom, as detailed below.

- **Factual Knowledge:**
  - Knowledge of terminology including technical vocabulary and musical symbols. In Fig. 1, such type of knowledge is stored through "Description" and "Multimedia Attachment".
  - Knowledge of specific details and elements: major natural resources and reliable sources of information. In Fig. 1, such type of knowledge is stored through "Description", "Property", "Instance", "Holonymy", "Meronymy", "Near Synonymy", "Synonymy", and "Antonymy".
- **Conceptual Knowledge:**
  - Knowledge of classifications and categories: geological time periods. In Fig. 1, it would be stored through "Hypernymy", "Hyponymy", "Time", and "Rank".
  - Knowledge of principles and generalizations: In Fig. 1, it would be stored through "Hypernymy", "Hyponymy", "Comparison", and "Multimedia Attachment".
  - Knowledge of theories, models and structures: In Fig. 1, it would be stored through "Theory/Model", "Cause/Effect", and "Multimedia Attachment".
- **Procedural Knowledge:**
  - Knowledge of subject-specific skills and algorithms: In Fig. 1, it would be stored through "Formula".
  - Knowledge of subject-specific techniques and methods: In Fig. 1, it would be stored through "Procedure".
  - Knowledge of criteria for determining when to use appropriate procedures: In Fig. 1, it will be stored through "Condition".

2.3 Test Item Structure Ontology

The test item structure ontology includes an intelligent online test scoring mechanism [28], which includes various parameters for dealing with fill-in-the-blank tests. In Fig. 2, the item structure ontology includes four question types: true-false, multiple-choice, multiple-response, and fill-in-the-blank. The ontology also includes original and variable item types. The question steam of original items can be generated based on primitive online material knowledge, in which case the structure of the question steam does not require any special changes. The original item is primarily used to assess the "remember" level of the cognition process. The structure of the question steam of variable items differs from that for online material knowledge. Furthermore, the variable item is used to assess the "understand, apply, analyze, and evaluate" levels of the cognition process. The variable items are divided into structure variable items and operands variable items. The structure variable items are generated by changing the structure, words of material knowledge. Moreover, the operands variable items are generated by calculation or formula inference module.

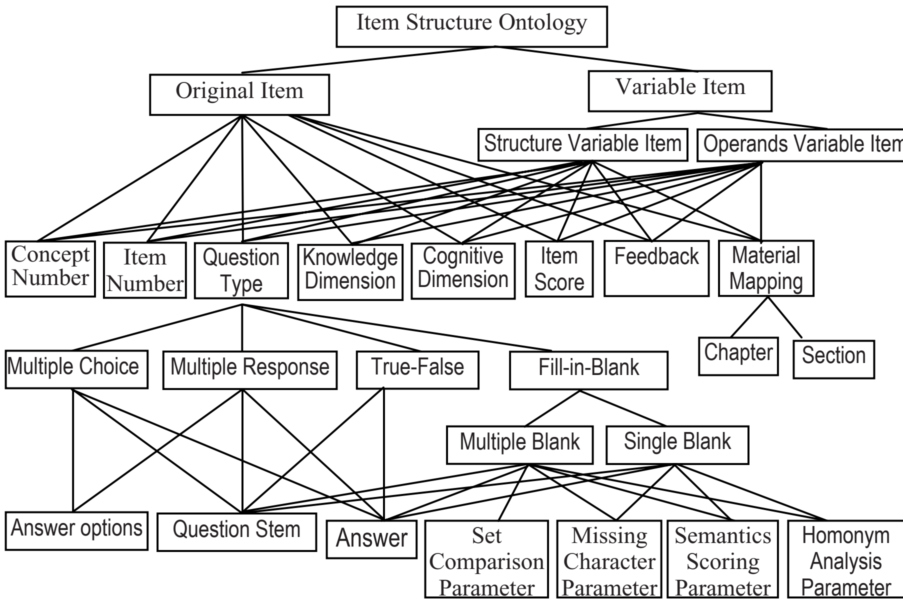

Fig. 2. Test Item Structure Ontology

2.4 CAGIS System Architecture

This study designed a computer-aided generation of items prototype system (CAGIS) in a three-tier Client/Server architecture. The back-end database server was Microsoft SQL Server 2000, which was used to implement trigger procedures and store the items, material, student data, scores, and so on. The web server was the Internet Information Server in Windows 2003. ASP language was adopted in the server-side. The architecture of the CAGIS E-learning system is shown in Fig. 3. The components are briefly described below.

This structure includes two user interfaces, five subsystems and 18 relevant databases. They are briefly described below. The Word Segment Process Subsystem segments the Chinese words in the primitive knowledge article, and stores the segmented results in the Expertise WS Knowledge Base. The Computer-Aided Generation of Material & Presentation Subsystem retrieves the segmented material knowledge from Expertise WS Knowledge and uses it to generate an online material knowledge, and stores it in the Material Knowledge Base. It can also dynamically generate teaching material pages that students can learn online. The Computer-Aided Generation of Item Subsystem, the focus of this study, can analyze the content of the Material Knowledge Base, generates various item types by referring to Item Structure Ontology and rules of item generation, and stores these items and standard answers in the Item Bank. The Online Test & Intelligent Scoring Subsystem manages testing and scoring. The Assisting Learning Tool Subsystem provides tools to assist learner leaning.

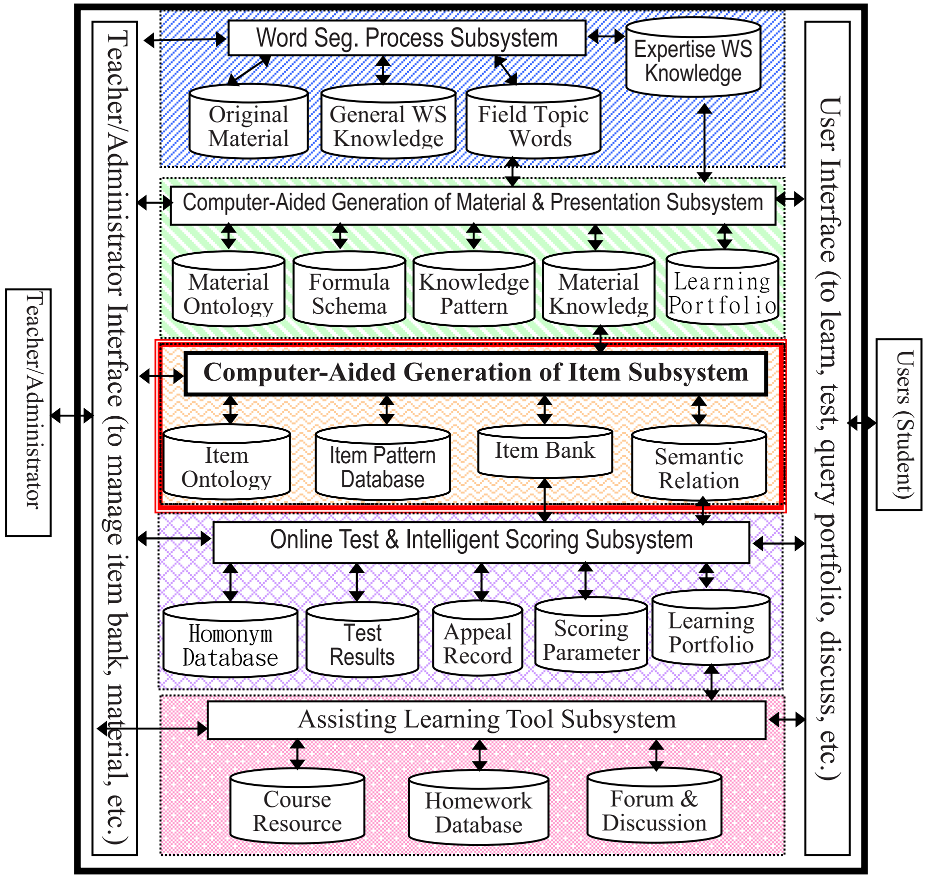

Fig. 3. CAGIS E-learning System Architecture

2.5 Computer-Aided Generation of Item Subsystem

Figure 4 shows he architecture of the Computer-Aided Item Generation Subsystem. From a 3*5 table of Bloom's taxonomy ("factual, conceptual, procedural" knowledge, and cognitive levels of "remember, understand, apply, analyze, evaluate"), teachers could assign numbers of four types of automatically generated test items: true-false, multiple-choice, multiple-response, and fill-in-the-blank. The components are presented below:

- **Formula Schema Database:** Storing the knowledge rule of mathematical formulae, logic operations, or equations.
- **Knowledge Pattern Database:** Storing the regular rules of Chinese grammar structure, semantic relations between words, and notation of word segments corresponding to Chinese sentences in general textbooks.
- **Material Knowledge Database:** Storing the knowledge content of the material. The knowledge was stored based on Material Knowledge Ontology. Relevant

knowledge can be linked by semantic relations. It is a knowledge source for generating online material in the Computer-Aided Generation of Material Subsystem and generating items for the Computer-Aided Generation of Item Subsystem.

- **Module of Item Pattern:** It provides a function for managing and maintaining the rules (characteristics) of item patterns, semantic relations, and question types for item generation.
- **Item Pattern Database:** Storing the rules (characteristics) of item patterns, semantic relation, and question type.
- **Module of Item Ontology:** This module provides a function for managing the item structure ontology.
- **Item Ontology Database:** Storing the item structure ontology.
- **Computer-Aided Generation of Item Module:** It executes the tasks involved in item generation. The module takes the knowledge content newly entered from the Material Knowledge Base, seeks other correlated existing knowledge concepts and checks the rules governing the item pattern. If the check is passed, the computer automatically generates the item and stores it in the item bank.
- **Item Bank:** Storing the items generated by Computer-Aided Generation of the Item Module. Alternatively, items created manually by teachers can also be stored if necessary.
- **Semantic Relation Database:** Storing the semantic relationships among words, including semantic words, correlation types (Near Synonymy, Synonymy, antonymy, etc.), and correlation ratios.

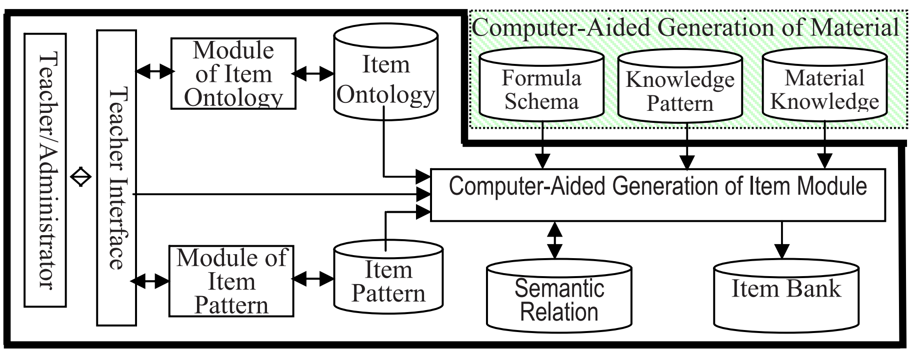

Fig. 4. Architecture of Computer-Aided Generation of Item Subsystem

2.6 Structure Rules of Knowledge Type and Item Generation Method

The Computer-Aided Generation of Item subsystem generates ten types of knowledge, Description, Property, Theory/Model, Cause/Effect, Sequence, Semantic Relation, Comparison, Formula, and Instance, and Others. The Formula Knowledge was created based on the formula schema set by teachers, the other nine knowledge types have their structure rules. These rules identify the knowledge type of original article contents, and store material knowledge that has been segmented to corresponding relation tables of the database. For illustration, some item generation methods are briefly described below.

- **Original Items:** The question steam structure refers to the same structure as the material knowledge base. For true-false questions, the answers are all true, which can be used to assess the ability of the "remember" process. The original items can generate items of other question types, e.g., fill-in-the-blank items, which can be used to "recall" ability.
- **Opposite Items:** If certain words in the question steam have the antonym sets in the Semantic Relation Database, the computer replaces them to produce the opposite items, which can assess the ability of confirmation in "remember" process level.
- **Grammar Inverting Items:** The material knowledge includes positive and negative concept sentences. If the computer exchanges and inverts the knowledge grammar structure of sentences, the sentences become the grammar inverting items. The grammar inverting items can be used to assess the ability of "understand" process.
- **Combined Same Subclass Knowledge of Single Concept Items:** These items were generated by the computer and combined with a lot of the same subclass (or sub-subclass) knowledge content from the single topic concept of materials. These items could be used to assess the confirmation ability in "understand" and "analysis" process levels. For example, since the concept "Expert System" has the following some characteristics: "Inference ability", "Explanation ability", etc. in the sub-subclass knowledge "General Characteristics", an item about "Expert System" concept can combine numerous "General Characteristics" knowledge.
- **Combined Same Subclass Knowledge of Multiple Concept Items:** These items were generated by the computer and used to combine a lot of the same subclass knowledge content from the multiple meaning-related topic knowledge contents of materials. For example, the concepts "Decision Support System" and "Expert System" could be compared with the "General Characteristics".
- **Combined Different Subclass Knowledge of Single Concept Items:** These items were generated by the computer and used to combine a lot of the different subclass knowledge contents from a single topic concept. For example, since the concept "Expert System" involves some knowledge in "General Characteristics", "Definition", "Condition", and "Meronymy", an item about "Expert System" concept could combine a lot of different subclass knowledge.
- **Combined Original Items of Same Concept:** These items were generated by the computer and combined a lot of original items of true-false of same topic knowledge from existing item bank. These original items could be combined to generate multiple-choice or multiple-response items.

3 Evaluation of System Effectiveness

This study compares computer-aided generation and manual item generation by teachers. The CAGIS used the same materials as the teachers used in a pilot study for item generation. Counting the different forms of the question stems and contents, CAGIS generated 18621 items, as shown in Table 3. However, certain items involve the same item concepts and meanings, because they were generated by procedure of combination and permutation in CAGIS. As a result, the CAGIS generated 1567 item groups with different assessment meanings (as listed in Table 4), which originated from 279 knowledge concepts of course materials. Each item thus can be replaced with an average of 11.466 (18621/1567) different forms of items. This study thus could solve the problems of shortages problem and excessive exposures of test items. In the pilot study, 15 teachers create 386 items in total. This CAGIS is more efficient than teachers on the quantity of items.

Furthermore, this study compares the effectiveness as follows. (1) The items produced by CAGIS include the assessment information of the knowledge and cognitive process dimensions. Such information can be used to provide learning suggestions for learners, and can also be used for teaching. (2) Teachers have difficulty creating the item of higher cognitive process level. In CAGIS, the items cover three types of knowledge and five dimensions of cognitive skills. (3) Regarding the degree of objectivity in selecting and generating items, teachers usually have personal subjectivity. However GAGIS follows the standard generation rules to select and produce items. (4) Regarding the effort spent on production and the quantity of items produced, 15 teachers produced 440 items manually and the average consuming-time of the teachers was 4.3 hours; CAGIS spent just 5 minutes producing the 1567 item group, and 18621 items. (6) Finally, because not all teachers underwent instructional strategy training, some items violated educational principles. However, these rules of preparing items are built into the Module of Item Pattern of CAGIS.

Table 3. Question Type of Items Generated by CAGIS

Question Type	True-False	Multiple Choice	Multiple Response	Fill-in-Blank	Total
Different Question stem and Answer Options	6.19% (1153)	35.51% (6612)	57.24% (10659)	1.06% (197)	100% (18621)
Different Assessment Meaning (Item Group)	32.04% (502)	20.49% (321)	37.97% (595)	9.51% (149)	100% (1567)

Table 4. Distribution of Items in Bloom's Taxonomy by CAGIS

Knowledge Dimensions	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Total
Factual	555 (35.42%)	0 (0%)		245(15.63%)	0 (0%)	809(51.05%)
Conceptual	137 (8.74%)	28 (1.79%)		108(6.89%)	0 (0%)	273(17.42%)
Procedural	17 (1.08%)	0 (0%)	2 (0.13%)	457(29.16%)	18 (1.15%)	494(31.53%)
Total	709(45.25%)	28 (1.79%)	2 (0.13%)	810(51.69%)	18 (1.15%)	1567(100%)

4 Conclusions and Future Research

Instructional designers and teachers have adopted Bloom’s taxonomy involved in all levels of education. This study applied ontology, Chinese semantic database, artificial intelligence, and Bloom's taxonomy, to propose a CAGIS E-learning system architecture to assist teachers in creating test items.

Based on the results of this study, we recommend the following: (1) applying machine learning techniques and revising the item pattern rules to generate items for supporting higher level cognitive processes, (2) exploring the item difficulty and item discrimination indexes, (3) executing empirical research to explore the learning effects of CAGIS.

References

1. Chou, W.J.: Implementation of Computer-Assisted Testing on WWW. In: Proceedings of the Seventh International Conference on Computer Assisted Instruction, pp. 543–550 (1998)
2. Chang, L.C.: Educational Testing and Measurement, Wu-Nan, Taipei (1997)
3. Kreber, C.: Learning Experientially through Case Studies: A Conceptual Analysis. Teaching in Higher Education 6(2), 217–228 (2001)
4. Bloom, B.S., Englehart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R.: A Taxonomy of Educational Objectives: Handbook 1, The Cognitive Domain. David Mckay, N.Y (1956)
5. Anderson, W., Krathwohl, D.R.: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Blooms' Educational Objectives. Longman (2001)
6. Lin, Y.D., Sun, W.C., Chou, C., Wei, H.Y.: DIYexamer: A Web-based Multi-Server Testing System with Dynamic Test Item Acquisition and Discriminability Assessment. In: Proceedings of the ICCE 2001, vol. 2, pp. 1512–1520 (2001)
7. Ho, R.G., Su, J.C., Kuo, T.H.: The Architecture of Distance Adaptive Testing System. Information and Education 42, 29–35 (1996)
8. Devedzic, V.B.: Key Issues in Next-Generation Web-based Education. IEEE Transactions On Systems, Man, And Cybernetics-PART C. Applications And Reviews 33(3), 339–349 (2003)
9. Hwang, G.J.: A conceptual map model for developing intelligent tutoring systems. Computers & Education 40, 217–235 (2003)
10. Moundridou, M., Virvou, M.: Analysis and Design of a Web-based Authoring Tool Generating Intelligent Tutoring Systems. Computer & Education 40, 157–181 (2003)
11. Sun, K.T.: An Effective Item Selection Method by Using AI Approaches. The Meeting of the Advanced in Intelligent Computing and Multimedia System, Baden-Baden, Germany (1999)
12. Hwang, G.J., Tseng, J., Chu, C., Shiau, J.W.: Analysis and Improvement of Test Items for a Network-based Intelligent Testing System. Chinese Journal of Science Education 10(4), 423–439 (2002)
13. Mitkov, R., Ha, L.: Computer-Aided Generation of Multiple-Choice Tests. In: Proceedings of the HLT-NAACL 2003 Workshop on Building Educational Applications Using Natural Language Processing, Edmonton, Canada, pp. 17–22 (2003)
14. Sumita, E., Sugaya, F., Yamamoto, S.: Measuring Non-native Speakers' Proficiency of English by Using a Test with Automatically-Generated Fill-in-the-Blank Questions. In: Proceedings of the Second Workshop on Building Educational Applications Using NLP, Ann Arbor, Michigan, pp. 61–68 (2005)
15. Liu, C.L., Wang, C.H., Gao, Z.M., Huang, S.M.: Applications of Lexical Information for Algorithmically Composing Multiple-Choice Cloze Items. In: Proceedings of the Second Workshop on Building Educational Applications Using NLP, Ann Arbor, Michigan, pp. 1–8 (2005)
16. Flavell, J.H.: Speculation about the Nature and Development of Meta-Cognition. In: Weiner, F.E., Kluwe, R.H. (eds.) Metacognition, Motivation, and Understanding. Lawrence Erlbaum, Hillsdale (1987)
17. Yang, H.L., Ying, M.H.: An On-line Test System Framework with Intelligent Fuzzy Scoring Mechanism. Journal of Information Management 13(1), 41–74 (2006)