Automation tools for assessment

John Kozma, PhD
ECPI University
7410 Northside Drive, Suite 100
North Charleston, SC 29420
jkozma@ecpi.edu
843-414-0350 x67943

ABSTRACT
Multiple choice tests can be used as both assessments and learning resources. When scored automatically they reduce instructor workloads while providing instant feedback for students. Various tools exist for assisting in the creation multiple choice questions. Most such tools, however, require significant manual effort, either in preparing text for questions or post-editing automatically generated text. For example, Moodle, an open-source Learning Management System, includes a graphical interface that allows copying and pasting of question text, but manual editing is required to ensure that the questions are logically and grammatically correct. It is proposed to incorporate minimally edited questions into quizzes administered to large numbers of students. Moodle includes tools that determine the discriminative efficiency of questions within the context of a single quiz. It is further proposed to refine these and other existing tools to facilitate automated analysis of response statistics from larger contexts of multiple courses to determine which questions should be revised or deleted.

Author Keywords
Moodle; Multiple-choice test generators; Natural language processing; Discriminative efficiency

Categories and Subject Descriptors
K.3.1 [Computers and Education]: Computer Uses in Education--- Computer-assisted instruction (CAI), Computer-managed instruction (CMI); J.1 [Administrative Data Processing]: Education

INTRODUCTION
Multiple choice reading quizzes are an effective means for improving students' mastery of course objectives [9]. Administering quizzes has been greatly facilitated by the use of online content management systems specifically designed for distance learning. Moodle [12], an open-source Learning Management System (LMS), has been adopted by our school, ECPI, for many courses, both online and classroom based. Moodle's quiz and question modules automatically grade true-false, multiple choice, and short answer quizzes, providing instant feedback to students. Quiz questions can be created using a graphical interface, and test banks can be imported in a variety of formats, including Blackboard, which is used by many textbook publishers.

While the use of an LMS can reduce grading workloads, this advantage depends on the availability of relevant quiz questions. Natural language processing (NLP) can be used to create multiple choice questions from machine readable text [5]. Output from even the most sophisticated NPL tools, however, may include ambiguous expressions, unresolved anaphora, grammar inconsistencies, or similar problems that must be corrected manually.

QUIZ STATISTICS AND LEARNING ANALYTICS
The widespread adoption of Massive Online Open Courses (MOOCs) and LMSs has led to the collection of large sets of data, but educators have only just begun to extract useful information from this data that will help improve the learning process [2, 3]. Learning analytics have been applied to the Kahn Academy platform to characterize individual problem solving habits [14]. Another suggested approach to learning analytics [4] uses Moodle with open source applications for data and process mining. Of particular interest for the present proposal is the data maintained by Moodle on quiz results. For example, questions answered incorrectly by a high percentage of students may be readily identified and revised accordingly.

Moodle users can create a quiz statistics report [15] with information about both a quiz as a whole and the individual questions in it. Along with simple items like the number of students attempting each question, the report gives information about discriminative efficiency based on formulas that correlate the score for each question with the score for the whole quiz. Unfortunately, the statistical significance of these discriminative efficiency measures is limited by the number of students taking the quiz. The same questions may be used in more than one quiz or even in different courses, but there is no ready way to combine information from multiple quiz statistics reports.

Since Moodle is a modular open source platform, it may be adapted to satisfy particular requirements. Examples of extensions for Moodle proposed by researchers include a virtual classroom [6], a Learning Analytics Enriched Rubric [7], self-guided lab activities [16], games [17, 18], and a course authoring interface [19]. Another technique used to adapt Moodle for specific needs is to link the gradebook to external applications and assignments [1]. The approach of

the current proposal is to develop a Moodle extension that will allow data and statistics from multiple quiz statistics reports to be combined. Information from the combined reports would then be used to identify test bank questions that should be modified or discarded.

Additional data could be utilized in the combined reports to correlate quiz performance with specific course curricula, instructions methods and other variables. The most recent version of Moodle allows the use of certainty based marking (CBM), which could be used to obtain even further information from quiz statistics reports. The use of automated techniques for grading free response questions [10, 11] may allow application of the proposed analysis to short answer questions as well.

CURRENT PROGRESS AND FUTURE WORK
A prototype tool [13], based on a javascript tokenizer created by Ariel Flesler [8], has been implemented for generating multiple choice questions in Moodle GIFT format from English text passages. Javascript was chosen for the implementation because it can be added to Moodle course pages by instructors without invoking administrative privileges. The prototype performs only rudimentary lexical analysis, splitting sentences at prepositions. The resulting output can be imported into Moodle, but would require significant editing to make the questions usable. Future efforts will focus on development of a more functional question generating tool using syntactic and semantic analysis and a Moodle extension for combining quiz statistical analysis reports.

CONCLUSION
When relevant multiple choice question test banks are available, they may be incorporated into a variety of learning resources. The task of creating multiple choice questions can be greatly expedited by the use of automated NLP tools combined with learning analytics. These tools can be implemented as extensions to Moodle, a widely adopted open source LMS.

REFERENCES
1. Baumstark, L. and Rudolph, E. Automated Online Grading for Virtual Machine-based Systems Administration Courses. *SIGCSE 2013*, ACM Press (2013), 477-481.

2. Clow, D. MOOCs and the funnel of participation. *LAK 2013*, ACM Press, 186-189.

3. Conde, M. and Á. Hernández-García, Á. A promised land for educational decision-making?: present and future of learning analytics. *TEEM 2013*, ACM Press (2013), 239-243.

4. Costa, C. Analysis of e-learning processes. *OSDOC 2011*, ACM Press (2011), 37-40.

5. Curto, S. Automatic generation of multiple-choice tests. Master's thesis, University of Lisbon, 2010.

6. Di Cerbo, F., Dodero, G., and Succi, G. Extending Moodle For Collaborative Learning. *ITiCSE 2008*, ACM Press (2008), 324.

7. Dimopoulos, I., Petropoulou, O., and Retalis, S., Assessing Students' Performance Using the Learning Analytics Enriched Rubrics. *LAK 2013*, ACM Press (2013), 195-199.

8. Flesler, A. String Tokenizer for Javascript, URL: http://flesler.blogspot.com/2008/03/string-tokenizer-for-javascript.html

9. Johnson, B. and Kiviniemi, M. The Effect of Online Chapter Quizzes on Exam Performance in an Undergraduate Social Psychology Course. *Teach Psychol*. 2009 January 1; 36(1): 33–37

10. Klein, R, Kyrilov, C., and Tokman, M. Automated Assessment of Short Free-Text Responses in Computer Science using Latent Semantic Analysis. *ITiCSE 2011*, ACM Press (2011), 158-162.

11. Kleiner, C. Tebbe, C., and Heine, F. Automated Grading and Tutoring of SQL Statements to Improve Student Learning. *Koli Calling 2013*, ACM Press (2013), 161-168.

12. Moodle, URL: http://moodle.org/

13. Moodle GIFT format question generator, URL: http://jkozma.com/MoodleMCqGenerator.html

14. Muñoz-Merino, P., Valiente, R., and Kloos, C. Inferring higher level learning information from low level data for the Khan Academy platform. *LAK 2013*, ACM Press (2013), 112-116.

15. Quiz statistics report, URL: http://docs.moodle.org/26/en/Quiz_statistics_report

16. Radenski, A. Digital CS1 Study Pack Based on Moodle and Python. *ITiCSE 2008*, ACM Press (2008), 325.

17. Sampaio, B., Morgado, C., and Barbosa, F. Building Collaborative Quizzes. *Koli Calling 2013*, ACM Press (2013), 153-159.

18. Tiusanen, M. and Lahtinen, R. Maze-Moodle Module for Games of Exercises. *Koli Calling 2011*, ACM Press (2011), 76-82.

19. Watson, C., Li, F., and Lau, R. A Pedagogical Interface for Authoring Adaptive e-Learning Courses. *MTDL 2010*, ACM Press (2010), 13-18.