article:2091

A Knowledge Modeling Approach to Evaluating Student Essays in Engineering Courses

One of the challenges of being an instructor in higher education is in deciding the best method for evaluating a student's level of understanding of course material. The overall question addressed in our research is whether a standard process can be developed to evaluate student responses to essay-type questions and whether that process can be automated. In order to answer this question, two preliminary issues must be addressed. Can models be developed from student responses and then compared against an expert model to effectively determine a meaningful score for that test question? If so, can these models be used as an evaluation of the student's comprehension? A promising starting place for the automation process is the use of a knowledge representation to formally specify a model for a good essay response, and then use that expert model as the standard by which the student responses are judged. In the past, latent semantic analysis has been used to automatically evaluate student essays. However, latent semantic analysis has been characterized as a "cheaper but less accurate" solution than knowledge representation methods such as those described in this paper.

Conceptual graphs, developed by Sowa [1], are a knowledge representation formalism. They provide a graphical representation of first order logic and are used for many computer-based applications such as natural language understanding, information retrieval, and expert systems. Conceptual graphs were chosen as the knowledge representation to be utilized in this research because they:

- are a mature technology, with a large body of international researchers making progress through a long-standing conference series (e.g., [2] [3] [4]),
- have a formal basis in first-order logic, making it possible to perform automated reasoning and inference [5],
- can be exchanged with other systems through use of an international standard for logic [6], and
- have a graphical display form that helps human understanding, and thereby assists in the validation of the model.

Examples of conceptual graphs can be found at http://www.jfsowa.com/cg/. The basic methodology used in this research consisted of the following steps:

1. Create a conceptual graph representing an expert model of the desired response.
2. Obtain essay question responses from students in an engineering course.
3. Create conceptual graphs representing each student's response.
4. Establish scores for conceptual graphs generated from student responses and evaluate the reliability of the scores.
5. Compare students' graphs and the expert model.

6. Evaluate comparisons among researchers and with respect to an independent expert.

We decided that the most direct approach to modeling the student responses was to begin with the expert model and mark those concepts, processes, tools, methods, and relations that were mentioned in the student response, eliminating those elements that were not mentioned. After all concepts and conceptual relations that had not been matched to the student's response had been removed, the remaining concepts and conceptual relations constituted the conceptual graph of the student's response to the question. Note that any additional information the student provided was thereby ignored, because it was outside the template provided by the expert model.

Two different methods were used to obtain scores from the conceptual graphs of students' responses. First, the Summation Method consisted of counting the number of concepts and relations that were present in graphs of student responses. Using this method, the assigned grades would simply be a reflection of the total number of concepts and relations identified, regardless of which particular concepts were used in the response. The second method employed was the Importance Method. The instructor of the course classified concepts as high, medium, or low in terms of importance to the desired response. Using this method, the assigned grades would be a reflection of the importance of the particular concepts and relations identified in the response.

For both the Summation and Importance methods of scoring conceptual graphs derived from student's essay responses, the reliability between coders of the same responses were shown to be within acceptable levels. To ascertain the validity of the conceptual graph scores, an independent expert (outside the research team) assigned a numerical grade to each student essay response. For the Summation method, there were no significant correlations between students' conceptual graph scores and the grades given by the independent expert. These results suggest that a simple count of matching concepts and/or relations does not capture the complexity of the process of grading an essay response. To assess the validity of the Importance method in predicting student grades, the instructor of the course classified each concept appearing in the expert model according to its level of importance as high, medium, or low. The number of concepts in each of these categories was submitted to a regression analysis as predictors, with the instructor's rankings of the student essay responses as the variable to be predicted. Both the number of highly important aspects ($b = 0.16$, $p = 0.056$) and the number of low importance aspects ($b = 0.25$, $p = 0.006$) were significant predictors of the instructor's rankings. In addition, the regression indicated that 36.3% of the variance in rankings could be attributed to these components of the conceptual graphs. This evidence indicates that most students included the most important concepts, but those student responses that were more complete (i.e., also including concepts of lesser importance) received higher grades.

Potential applications of using conceptual graphs for assessing student essay responses include grading consistency, potential automation of grading, and identifying teaching deficiencies. First, when instructors grade essay responses or evaluate the quality of term papers, they should have a rubric that they use in their assessments [7]. A *rubric* is a scoring system that identifies various levels of performance in a process or product [8]. If instructors design a rubric for assessing students' work, it is likely that they will be more consistent in their evaluations across students. Conceptual graphs could be used to model the rubric in such a way as to increase the consistency of grading across students' responses. Thus, conceptual graphs could provide a common source or reference for instructors to derive judgments when grading a large set of exam papers. Second, teachers sometimes grade a question by simply looking for certain key information and ignoring any extraneous information. This is the approach an automated tool employing knowledge modeling would use. However, some teachers grade a question by taking off points for incorrect additional information or extraneous additional information. An automated tool might be able to handle this grading approach by similarly subtracting points for extraneous concepts and relations. Finally, conceptual graphs can also be used to identify teaching strengths and weaknesses. Instructors can use an expert conceptual graph to map information that students should know. For any given topic, aspects of the students' and instructor's conceptual graphs that do not match could then be targeted by the instructor for additional attention or a different type of presentation in future classes. In another educational application, there is substantial evidence in the research literature of education, psychology, and organizational training that people learn more easily when the instructional style matches their own learning style (see [9], for a review). Conceptual graphs could be used to determine the extent to which different delivery methods or instructional styles are appropriately aligned to students' or trainees' learning styles.

References

[1] Sowa, J.F., *Conceptual Structures: Information Processing in Mind and Machine*, Reading, Mass.: Addison-Wesley, 1984.

[2] Tepfenhart, W., and W. Cyre, "Conceptual Structures: Standards and Practices," *7th Intl. Conf. Conceptual Structures*, Blacksburg, VA, USA: Springer-Verlag, 1999.

[3] Chein, M., and M.-L. Mugnier, "Conceptual Structures: Theory, Tools and Applications," *Lecture Notes In Artificial Intelligence*, Montpelier, France: Springer-Verlag, 1998.

[4] ICCS, "International Conference on Conceptual Structures," 2006, http://www.iccs.info.

[5] Sowa, J.F., *Knowledge Representation: Logical, Philosophical, and Computational Foundations*: Brooks/Cole, 2000.

[6] ISO, "Information technology — Common Logic (CL) – A framework for a family of logic-based languages - FDIS 24707": ISO/IEC JTC1/SC 32/WG 2, International Organization of Standards, 2006.

[7] Routman, R., *Invitations: Changing as teachers and learners K-12*, Portsmouth, NH: Heinemann, 1991.

[8] Weiner, R., and J. Cohen, *Literacy portfolios: Using assessment to guide instruction*, Columbus, OH: Merrill, 1997.

[9] Cassidy, S.," Learning styles: an overview of theories, models, and measures", *Educational Psychology* Vol. 24, No. 4, 2004, pp. 419-444.

Author 1: Sandra L. Carpenter; carpens@uah.edu Author 2: Harry S. Delugach; delugach@cs.uah.edu Author 3: Letha H. Etzkorn; letzkorne@cs.uah.edu Author 4: Phillip A. Farrington; paf@ise.uah.edu Author 5: Julie Fortune; julie.fortune@qsm-inc.com

Article Link:www.asee.org

: Back to 2008 Fall Issue, Vol. 4, No. 1

: Back to List of Issues

: Back to Table of Contents