PUTTING PEDAGOGY IN THE DRIVING SEAT WITH OPENCOMMENT: AN OPEN SOURCE FORMATIVE ASSESSMENT FEEDBACK AND GUIDANCE TOOL FOR HISTORY STUDENTS

Denise Whitelock and Stuart Watt

Paper published in the Proceedings of the 12th CAA International Computer Assisted Assessment Conference, 8/9 July 2008, edited by Farzana Khandia ISBN 09539572-7-6 pp.347-356

Putting Pedagogy in the Driving Seat with OpenComment : An open source formative assessment feedback and guidance tool for History students

Denise Whitelock,
Open University,
Walton Hall, Milton Keynes
D.M.Whitelock@open.ac.uk

Stuart Watt
The School of Computing
The Robert Gordon University
Aberdeen
s.n.k.watt@rgu.ac.uk

Introduction

The Open Comment project sits within an external demand for electronic assessment from policy makers together with the QCA and SQA (see Whitelock & Brasher 2006 and the final report to the JISC on a Roadmap for e-Assessment http://www.jisc.ac.uk/elp_assessment.html). Universities too, together with Further Education establishments, are embracing e-assessment (see Whitelock et al, 2006). There is a recognition that e-assessment accompanied by an appropriate feedback to the student is beneficial for learning (DiBattista et al, 2004; Pitcher el al., 2002; Whitelock & Raw, 2003). Distance Learning too is forging ahead with electronic delivery of courses together with addressing the complexities of e-assessment for large cohorts of students.

One of the more challenging aspects in the current e-assessment milieu is to provide a set of electronic interactive tasks that will allow students more free text entry and provide immediate feedback to them. In other words, being able to repeat in some small measure what occurs ordinarily in many student texts where self-assessed review questions are raised, the readers then reflect upon them and the answers can be found in the back of the book. The electronic approach would provide a set of interactive tasks. Students type in their answers, hints are given if the response is incorrect and the student can try again. If completely baffled, an answer can be provided. This pedagogical strategy would be the ideal type of electronic formative assessment. Disciplines such as Science and Mathematics have been able to use this

approach, for example, as in multimedia activities used in the Open University's "Discovering Science" first-level course. However, in disciplines that require more free text entry, such as is found in the Arts, this has yet to be achieved. The JISC funded e-Assessment Case Study project http://kn.open.ac.uk/document.cfm?docid=10817 which investigated 17 sites of excellence in the UK did not identify any work that was going on within the Arts arena. There was, therefore, a need to explore free text entry response systems with automatic marking.

Free text response processing is at the cutting edge of linguistics research and the team were under no illusions but that what was being attempted was very ambitious. Certainly a completely human-like response to free text is well beyond the state-of-the art, but experience has shown that sometimes it is possible to provide effective responses based on surface features of a free text response, as was achieved in OpenMentor (Whitelock et al, 2003). Carefully constructed language, conversational in form, can be even more important to guiding learning than the content being communicated (Holmberg, 1983). Instead of providing feedback on the answer, the project's approach was, to some extent like ELIZA (Weizenbaum, 1963), to couch just enough analysis of the text in reflective language to help the learner assess their own work.

The specific objective of the project was to construct some simple tools in the form of Moodle extensions that allow a Moodle author to ask free-text response questions that can provide a degree of interactive formative feedback to students. In parallel with this was the aim to begin to develop a methodology for constructing such questions and their feedback effectively, together with techniques for constructing decision rules for giving feedback. Open Comment is a formative feedback technology designed to be integrated in the Moodle virtual learning environment. Put simply, it provides a simple system allowing questions to be written in Moodle, and for students' free text responses to these questions to be analysed and used to provide individually customised formative feedback. Open Comment is related to traditional free text assessment technologies, such as the ETS e-rater system and Landauer et al.'s (1998) IEA, although it has a very different emphasis. In particular, it makes no attempt to provide grading information; instead, it provides reflective feedback, designed to guide the students in their learning.

Although Open Comment was designed principally for Moodle, it is an open and flexible framework, and there should be no significant difficulties adapting it to embed its functionality into any other formative assessment system.

It was a deliberate and eary decision to separate the feedback engine from the VLE as a web service. This is in keeping with JISC's emphasis on service-oriented architectures. However, generating feedback is computationally intensive, and Moodle is implemented in a language that is not suited to computationally intensive processing. Using this approach allows the load to be balanced, with the VLE running on one set of servers, and feedback generation on separate systems if required. A second benefit of this is that

only the presentation aspects of the system need to be adapted to additional VLEs.

Open Comment has been developed as an open source system, and consists of the following components:

- A Java-based feedback system
- A web service shell
- A Moodle-based question type
- A graphical interface for testing
- A forms-based editing tool

Unlike most prior work, Open Comment does not commit to any particular technologies. Although latent semantic analysis (Landauer et al., 1998) has been used successfully under some circumstances, it is not the only game in town, and it does require significant effort developing a training set. However, in many cases, keyword or phrase matching can be just as helpful. Open Comment allows many different classification engines to be used to recognise evidence of understanding and use of knowledge, and their results integrated into feedback to the learner.

Pedagogical principles driving the feedback engine

This paper wishes to report on the feedback engine and the pedagogical principles which drove its development since the pedagogical rationale for this development was to engage students in a series of electronic formative assessment tasks that would provide more free text entry with automatic feedback. This would promote a more challenging experience for the students than just checking their learning for revision purposes and promote a more personalised learning environment for self-reflection.

The guidance text arose from our analysis of what feedback actually was, and how learners used it. It built on our earlier work on Open Mentor (Whitelock & Watt, 2007).Throughout the development work, we worked closely with expert tutors in several Arts disciplines, using a range of techniques to elicit the processes they used to provide appropriate feedback. These ranged from role play (becoming a student) through to analysing collections of real answers and constructing sample solutions.

A preliminary analysis of 68 History assignments together with 100 plus assignments from different disciplines revealed a common pattern of tutor responses. These were clustered around the main categories of praise, advice on structure and presentation, particular misunderstandings, and developing and understanding particular issues.

The underlying model of feedback centred around:

- Identification of salient variables

- A description of these variables
- Identification of trends and relationships between these variables

The result of these analyses were formalised as an operational model for formative feedback generation, as set out in the table 1 below.

Table 1. Operational feedback model for Open Comment

Stages of Analysis by computer of students' free text entry for Open Comment	**Advice with respect to Content**	**Socio-Emotional Support**	**Stylised Example**
STAGE 1a: DETECT ERRORS E.g. Incorrect dates, facts. (Incorrect inferences and causality is dealt with below)	Instead of concentrating on X, think about Y in order to answer this question	Recognise effort (Dweck) and encourage to have another go	You have done well to start answering this question but perhaps you misunderstood it. Instead of thinking about X which did not…….. Consider Y
STAGE 1b: **IF NO INCORRECT STATEMENTS GO TO 2**			A good start……
STAGE 2a: **REVEAL FIRST OMISSION**	Consider the role of Z in your answer	Praise what is correct and point out what is missing	Good but now consider the role X plays in your answer
STAGE 2b: **REVEAL SECOND OMISSION**	Consider the role of P in your answer	Praise what is correct and point out what is missing	Yes but also consider P. Would it have produced the same result if P is neglected?
STAGE 3: **REQUEST CLARIFICATION OF KEY POINT 1**	Explain X more fully What do you mean by X	Confirm and concur about what is correct encourage to take the analysis further	
STAGE 4: **REQUEST FURTHER ANALYSIS OF KEY POINT 1** (Stages 3 and 4 repeated with all the key points)	Analyse X more fully	Confirm and concur about what is correct encourage to take the analysis further	Very interesting point – X is very complex perhaps it would have been effective to look at things slightly differently and consider how X contributes to Y
STAGE 5: **REQUEST THE INFERENCE FROM THE ANALYSIS OF KEY POINT 1 IF IT IS MISSING**	Request the conclusion that can be drawn from X.	Praise effort and reiterate progress is being made	This is a sound description but it would be good if you explain what X is contributing to this situation.

STAGE 6: REQUEST THE INFERENCE FROM THE ANALYSIS OF KEY POINT 1 IF IT IS NOT COMPLETE	What is X causing in this situation?	Reaffirm progress but encourage student to take the analysis process one step further	Yes what you have written is correct but can you elaborate and explain what X means?
STAGE 7: CHECK THE CAUSALITY	What is X, Y and Z causing in this situation?	Praise persistence and effort and ask the user to think about the reasoning behind a particular response.	You are certainly improving your answer to this question. Well done. In order to improve your answer further could you say something about the role X played in Y I'm thinking particularly of the following example where X was seen with respect to Z.
STAGE 7: REQUEST ALL THE CAUSAL FACTORS ARE WEIGHTED	Do X, Y and Z contribute in the same way to producing situation C, i.e. do the variables have equal weighting	Praise persistence and effort and ask the user to think about the importance and relative weightings of the causal factors	You have made a good stab at this question. From your answer I think you are allowing a considerable role to X. Does this mean you accept that X alone causes Y

As this table shows, this model operates by and large through a sequential set of rules identifying sources of evidence within the student's response, and escalating in level of analysis, in some sense following Anderson, Krathwold, and Bloom's (2000) revised taxonomy of educational objectives. Importantly, also, there is a strong causal element to many of the rules[1]. These rules are implemented in a bespoke feedback engine within Open Comment – by and large, all the other components are only there to make it accessible in a usable form, through a VLE or through an interactive interface. Although we have set out the main principles behind the feedback system, it is worth being more specific about the details.

Much of this model is implemented in JavaScript rules[2], which make the bridging inferences between the levels. Simple errors of omission or commission can be immediately added to the response; otherwise, the analysis passes on to more detailed feedback on later stages. Each question is analysed using a script in a configuration file, allowing many questions to be configured and handled from the same main feedback engine. Each question will typically provide its own configuration file, although this is not always necessary, as in some cases several questions may be closely related, and share aspects of inference about appropriate feedback.

[1] This seems to be particularly important to the domain chosen (History). It is very likely that this will not be the same in other domains, although causal reasoning is expected to be important in a fair number of both related and unrelated fields.

[2] After careful investigation of the option of developing a domain-specific language for feedback, we felt that JavaScript smoothed the learning curve for developers. However, Open Comment uses an entirely different object model compared to web JavaScript, and it is this object model that enables access to evidence from a range of advanced text classification technologies.

So far, only a few questions have explored the higher stages of the feedback model, looking at causality. In our initial work on more detailed questions (and in higher level courses) this was more prominent than in the later, smaller, questions. It remains an important topic for further work.

One important result has been an increased understanding of the differences between even closely related disciplines. In both History and Philosophy, as with many humanities and social sciences, there is a greater emphasis on developing each students' ability to reason, and to use arguments and evidence in ways that are in keeping with a discipline-specific methodological ethos. Questions could rarely be taken at face value – especially in the more advanced levels. We found that our feedback systems focused far more on evidence than on getting the answer right; effective development of formative feedback technologies in these disciplines is totally dependent on effective involvement of tutors with both pedagogical and domain expertise.

Discussion

The first demonstration system was received favourably by Arts Faculty staff who have now become more aware of both the potential and limitations of automated systems based on free text responses. Lessons have been learned about the type of feedback that instructors think would be most useful. In particular, we have found that it appears to be worth distinguishing two main classes of feedback. These being:

- Specific to the question
- Generic for Arts-style questions

A certain degree of feedback to students on free text answers can be usefully generated, but cannot with the current state of the art, replace detailed feedback from a qualified academic. The benefit to the students is that helpful feedback can be given almost instantaneously. This should encourage more rapid progress and build student confidence. The benefit to the course tutor is that more off-the-point responses should be identified by the system so that the tutor's attention can be focused on more substantial issues that are pertinent to the students.

Acknowledgements

We appreciate the support of Clive Baldwin, Colin Beagrie, Malcolm Clark, Ian Craw, Frank Herrmann, Jan Holt, Chris McKillop, Mark Pittaway, Carolyn Price and Jeanette Rey for their assistance with this phase of the project.

References

Anderson, L.W., & Krathwohl, D.R. (eds.) (2000) A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. Allyn & Bacon.

DiBattista, D., Mitterer, J.O. & Leanne, G. (2004) Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing. Teaching in Higher Education, Vol. 9, No. 1, pp.17-28, ISSN 1356-2517

Holmberg, B. (1983) Guided didactic conversation in distance education. In D. Sewart, D. Keegan, and B. Holmberg (Eds.), Distance education: International perspectives (pp. 114-122). London: Croom Helm.

Landauer, T.K., Foltz, P.W. & Laham, D. (1998) An Introduction to Latent Semantic Analysis. Discourse Processes, 25, pp. 259-284

Masterton, S.J. & Watt, S.N.K. (2000) Oracles, bards and village gossips, or, social roles and meta knowledge management. Information Systems Frontiers, 2 (3/4)

Pitcher, N., Goldfinch, J. & Beevers, C. (2002) Aspects of computer bases assessment in mathematics. Active Learning in Higher Education, Vol. 3, No. 2 pp.19-25

Watt, S.N.K. (2006) Text categorization and genre in information retrieval. In A. Goker & J. Davies (Eds.) Information Retrieval, John Wiley & Sons

Weizenbaum, J. (1966) ELIZA: A Computer Program for the Study of Natural Language Communication between man and machine. Communications of the ACM, 9(1), 36-45.

Whitelock, D., Watts, S., Raw, Y. and Moreale, E. (2003) 'Analysing Tutor Feedback to Students: First steps towards constructing an Electronic Monitoring System'. Association for Learning Technology Journal (ALT-J). Vol. 11, No. 3, pp. 31-42 ISSN 0968-7769.

Whitelock, D. & Raw, Y. (2003) Taking an electronic mathematics examination from home: what the students think. In C.P. Constantinou and Z.C. Zacharia (eds.) Computer Based Learning in Science, New Technologies and their Applications in Education, Vol. 1, Nicosia: Department of Educational Sciences, University of Cyprus, Cyprus, pp. 701-713, ISBN 9963-8525-1-3

Whitelock, D. & Brasher, A. (2006) Developing a Roadmap for e-Assessment: Which Way Now? 10th International Computer Assisted Assessment Conference, Loughborough University, 4/5 July 2006, pp. 487-501. ISBN: 0-9539572-5-X

Whitelock, D., Mackenzie, D., Whitehouse, C., Ruedel, C. & Rae, S. (2006) Identifying Innovative and Effective Practice in e-Assessment: Findings from

Seventeen UK Case Studies. 10th International Computer Assisted Assessment Conference, Loughborough University, 4/5 July 2006, pp. 505-511. ISBN: 0-9539572-5-X

Whitelock, D. and Watt, S. (2007) 'e-Assessment: How can we support tutors with their marking of electronically submitted assignments?' Ad-Lib Journal for Continuing Liberal Adult Education, Issue 32, March 2007. ISSN 1361-6323