

Re-usable OnLine Assessment Materials for Teaching Artificial Intelligence: An Overview

Dr Jim Smith
School of Computer Science
University of the West of England, Bristol
James.smith@uwe.ac.uk

Keywords: on-line assessment, accessibility, artificial intelligence,

Abstract

This report is concerned with the creation of a bank of different types of questions at various levels for enhancing the teaching and assessment of Artificial Intelligence (AI). This subject has ever-increasing importance in today's "knowledge economy", but is frequently perceived as a somewhat "dry" and "difficult" subject area therefore the development of alternative methods for delivering and assessing these materials is particularly timely. The materials produced are tagged with meta-data to facilitate the creation of multiple targeted assessments relating to specific topics, at various levels of difficulty.

This report also considers issues relating to the accessibility of electronically delivered assessment to various user groups, focussing in particular on issues arising for students suffering from dyslexia and/or who use screen readers.

Aims

The aims of this project were to:

- facilitate the introduction of AI materials to students.
- assist in the incorporation of AI material into other courses.
- promote widening participation by supporting learning of AI in a structured and self-paced manner beyond the traditional modes of delivery.
- support students through the transition to higher education via the provision of rapid and targeted feedback on their progress and providing guidance on remedial steps.
- examine issues related to the presentation and accessibility of e-learning materials and how they affect, for example people with dyslexia and partially-sighted users. This is particular important in this field as AI material tends to be jargon and acronym-heavy.
- address issues of retention and progression via automated means for providing detailed formative feedback available to students, enabling more targeted support based on students' needs, and identification of students experiencing difficulties engaging with the programme of study.

Outcomes

The first outcome of this project was a pool of over one hundred questions for assisting in the teaching and assessment of Artificial Intelligence. These were used and refined in practice as part of the delivery of a level one undergraduate module "Introduction to Artificial Intelligence" which is a core part of the awards "Computer Science " and "Games Technology" at UWE. This is a relatively large module

(cohort of more than one hundred) which makes extensive use of various e-learning tools provided by a Virtual Learning Environment (VLE) – in this case “Blackboard”[1] -- to deliver content and facilitate group working. The question bank was used to create formative assessment tests, which the students could use to check their understanding of each week’s materials. Blackboard’s “adaptive release” mechanisms were used encourage participation by making release of subsequent materials dependent on attempting each week’s tests. The ability to monitor test results and usage statistics for the course materials on a weekly basis was particularly valuable as it permitting rapid identification of those students who did not appear to be engaged with the course, which could then be followed up appropriately via UWE’s Graduate Development Programme and pastoral arrangements. While it is too early to assess the full impact of retention and examination results, feedback from the students has been overwhelmingly positive, and receptive to the use of a variety of learning and assessment methods. It is intended to use a similar bank of questions to conduct the final end-of-year exam, which will be added to the question pool.

A second, and equally important outcome resulted from the discussions with the cohort, and use of additional volunteers to assess the impact of switching to computer mediated assessment for people with dyslexia, and users of screen readers. Both groups expressed the opinion that on-line assessment offered benefits over traditional “paper” methods, particularly for learners with dyslexia. However it has to be emphasised that this benefit only accrues when the setters take particular care in the way that they express and set questions, over and above the established guidelines such as [2,3]. Notwithstanding the small sample sizes, the findings made it clear that certain types of question, such as “ordering” can cause particular difficulties for dyslexic students. More importantly, some types of question, such as “jumbled sentences” may be presented by VLEs in a way that is not properly accessible to users of screen readers. In order to make the learning materials accessible, and in particular if they are to be used for summative assessment, either these should be avoided or specific combinations of screen readers and VLEs should be tested for compatibility before the final assessment is prepared.

Deliverables

The principal outcome of this project is a pool of 107 questions which are available from the url www.cems.uwe.ac.uk/~jsmith/rlo . They are packaged as zip archives of xml files, which can be directly imported into Blackboard, or into other formats by using a suitable xml translator such as those available in Respondus (www.respondus.com). Once imported the question pool may be used to create a range of tests as required.

The pool has been designed to provide a wide range of different question types as described in Table 1. The questions cover a range of topics in Artificial Intelligence, and each question is tagged with meta data concerning its primary topic (introduction, philosophy, search, knowledge representation, problem representation, artificial neural networks, evolutionary computation) and level of difficulty (easy, moderate, difficult). The questions have been extensively tested and refined via their use for weekly formative assessment in a large level 1 undergraduate module. They also contain extensive feedback to aid their use in formative as well as summative assessment. Feedback is provided for correct and incorrect answers, and also where applicable at the level of individual responses.

Question Type	Description
multiple choice	User selects exactly one of a series of options.
multiple answer	User selects exactly one of a series of options.
numerical calculation	User provides numerical response in dialogue box
missing blanks	User provides series of text responses in dialogue boxes to complete sentences.
jumbled sentences	User makes series of selections from pull-down boxes to complete sentences.
matching	User makes series of choices form pull-down boxes to match two sets of items
ordering	User makes series of choices form pull-down boxes to rank a set of items

Table1: Question Types Used

The second deliverable is a report on some case studies of how accessible the materials were [4]. It contains detailed comments on certain types of questions, and an assessment of how one particular combination of screen reader and Virtual Learning Environment (VLE) interacted. In this test certain kinds of question were found to be inaccessible to the volunteer for technical reasons. While this is of course a limited sample, it raises the issue of whether Higher Education establishments should be expected to either (i) test every combination of screen reader with their particular VL, (ii) force students to use a screen reader which is known to work with their VLE, or (iii) avoid the use of certain question types.

One obvious benefit to the academic of employing computer-marked assessment, is a reduction in the marking task, particularly when fast feedback is required. It must be emphasised that in this author's opinion, these materials provide a way not of reducing the amount of time spent on a module, but of using it more profitably. To give an example, in the next academic year it is intended to extend the range of questions, and provide extra voluntary tests for students who wish to either extend their knowledge of a topic and test it at a deeper level, or to spend more time consolidating their understanding of the topic as taught. Also under consideration are ways of automating the provision of extra materials that students can self-select depending on how they have fared with different topics.

Background

As noted above, Artificial Intelligence is a subject that is frequently regarded a somewhat "dry" but which is of increasing importance on a variety of courses from Computer Science to more application specific areas such as Games Technology. The module chosen as a pilot study attracts a wide range of students from different backgrounds and therefore provides a good test-bed for evaluating the use of e-learning tools.

In addition to the general aims and objectives outlined above, the specific benefits that we hoped to achieve, and measure in our evaluation are:

- improved retention by providing self-assessment tools to be used after each lecture, therefore easing the transition into higher education for the students, and providing a means for the tutors to monitor those students who are not engaging with the course,
- more transparent assessment, and
- more rapid feedback to students than a humans-markers can achieve.

Putting it into Practice

There are a variety of ways in which these materials could be incorporated into the delivery of Artificial Intelligence. The tests can be used for both formative (e.g. weekly quizzes) and summative assessment (e.g. end of module exam). The use of meta-data permits the tutor to select questions by type, category and difficulty level as appropriate to their course.

The route followed at UWE was one of quite intensive computerisation of the delivery of content and assessment, coupled with carefully designed large group tutorials. Electronic discussion groups and group work areas were used to provide the possibilities for smaller scale collaborative work and individual communication, and appear to work reasonably well. Students responded well to the clarity and openness of the assessment, and the immediate feedback, although “missing blanks” questions were unpopular because it is hard for the tutors to predict the full range of possible responses. On their part, the staff involved have reported that it is more enjoyable and productive to spend time throughout the year devising good electronic questions, and providing detailed feedback, than spending the same amount of time marking a large number of similar exam scripts.

Weekly in-class assessment provides benefits in terms of monitoring engagement, but does not fit that well with an agenda of widening access and remote delivery. Another benefit that accrues from the use of VLEs to deliver self-assessment materials, is the ability to gain these benefits while providing of materials that students can work through in their own time. Of course students may not be aware when they have misunderstood a concept or topic, and so not see the value of taking a test. Our experience is that to encourage uptake, either the tests should count in some way towards the final mark, or techniques such “Adaptive Release” should be used, in which case the materials made available on satisfactory completion of each test must provide sufficient incentive to take the test as many times as necessary.

A vital factor in the provision of e-learning materials is making every attempt to ensure that fair access is available to all students. To this end, a number of groups have considered the issues and produced guidelines on web standards e.g. [2]. However the specific issues relating to making assessment tools accessible has been less well studied, although some good pointers may be found in [3]. Many of the issues relate to the correct use of punctuation, avoiding images, and providing clear and explicit instructions (for every question if a mixture of question types is employed). However some specific issues arise from the type of questions, and the way that VLEs such as Blackboard display them. The supplementary report [4] contains two case studies, which suggest that certain problem types such as “*order the following list according to ...*” may inherently cause particular problems for students with dyslexia. Additionally, the screen layout and format of certain question types as displayed by VLEs may cause problems with other types. As noted above, this is an issue that should probably be dealt with on an institutional level, bearing in mind the legal requirements against discrimination.

Issues and Debates

How can this be managed? It all seems like a lot of work...I don't like product X...

There are two main methods by which on-line assessment can be provided, namely using customised software, or within the context of a VLE provided at the institutional level. Despite many academic's misgivings about how well certain VLEs function, it is this author's opinion that the benefits of using the full range of monitoring and deployment tools available within a VLE – an environment with

which most students are having other module content delivered, far outweigh any shortcomings. Developing software, and materials in a customised format may be personally satisfying for the tutor. However it is surely desirable to use software with which the students are familiar, allowing them to focus on the content of the assessment rather than the method of delivery. Moreover, either the accessibility of any customised software has to be properly assessed, or the tutor faces the risk of inadvertently discriminating against some student groups.

What if I don't have Blackboard

The question bank developed is supplied in a format which can be immediately imported into Blackboard, and since it is in xml format, can be readily imported into other software via free or low cost tools such as (but not only) the Respondus software.

How can you adequately test understanding and knowledge at level X?

It is possibly the case that simple multiple choice questions may appear to limit the depth of understanding that can be assessed. However, most VLEs now support a wide range of problem types, such as multiple answer, numeric calculation, jumbled sentences, etc, and with due thought these can be used to probe understanding quite deeply. Using these tools undoubtedly requires a rethink of the assessment process, moving to a model whereby the majority of time can be spent refining and developing questions, which are then automatically marked, rather than on marking a number of exam scripts. The fact that this can be done over a longer time period than is traditionally allowed just for the marking process means that any limitations arise more from the imagination of the academic, than from the medium via which assessment is conducted.

Resources

A well publicised and significant resource is MIT's opencourseware in Artificial Intelligence (<http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-034Spring-2005/CourseHome/index.htm>), but this takes a very "traditional" view of AI in its scope, and the assessment tools are not particularly suitable for level 1 students

For more subject specific texts in a range of media, a variety of resources exist, of which most notable is the material on Evolutionary computation in the EvoNet flying Circus (<http://evonet.lri.fr/CIRCUS2/node.php?lang=fr&node=1>).

References

- [1] Blackboard version 7.1. <http://Blackboard.com>
- [2] W3C, Web Content Accessibility Guidelines 1.0 available at <http://www.w3.org/TR/WAI-WEBCONTENT/>
- [3] Online Assessment Tools: Do They Work for Students with Disabilities? Fiedor, L, and Primlani, S. Presented at EDUCAUSE Annual Conferences (10/23/2007). Available at http://delta.ncsu.edu/projects/is/accessibility/assessment/presentation/html/web_data/slides_and_notes.htm.
- [4] Smith, J.E. "Analysis of accessibility issues for creating and deploying on-line assessment tools". Technical report available from www.cems.uwe.ac.uk/~jsmith/rlo.