USING COMPUTER ASSISTED ASSESSMENT TO FACILITATE RESOURCE SELECTION

Mhairi McAlpine

Using Computer Assisted Assessment to Facilitate Resource Selection

Mhairi McAlpine
Robert Clark Centre
University of Glasgow
G12 8QQ

Abstract

The number of computer mediated learning resources is growing at a phenomenal rate, and the advent of the web has made access easier than ever. This has however led to difficulties identifying resources suitable for the particular teaching purpose needed. One way round this is through directories, however the information that they provide is often limited to a few parameters. Where human evaluation is included, the review is often too short to give meaningful advice on its quality and most appropriate usage, or too long to enable quick identification.

Modern computer assisted assessment packages are capable of storing and analysing vast amounts of information on student learning. With appropriate analysis this data can be used to pinpoint the strengths and weaknesses of individual students and match these to learning resources that meet their needs.

This paper outlines a proposal to use the analysis of Rasch residuals to compile profiles of individual students. Rasch analysis allows computation of a candidate's average ability and a question's average difficulty. Where an individual candidate's ability on one question is below their average, they can be considered to be weaker on that question than on others. Students can then be directed to resources that address that weakness. This overcomes the problems of manually identifying student needs.

Computer assisted assessment generates a great deal of data which can be utilised far more fully than is currently the case. This paper urges the CAA community to look beyond the creation of items and towards their future analysis.

Keywords

Computer Assisted Assessment, Item Analysis, Rasch, Classification, Learning Resources, Profiling

Introduction

Computer assisted assessment is finally coming of age. Although it has a long history, prior to the development of software that enabled easy creation and banking of items, it was of limited use to the average university teacher, who had either to develop significant programming skills, or to rely on pre-prepared items. Software such as Question Mark made item creation and selection possible for an average computer literate academic. The advent of web based software has taken this further and enabled distributed test creation, test taking and facilitated easy item sharing. Most recently, the IMS QTI specifications have allowed items to be classified under standardised parameters further facilitating item sharing by providing metadata guidelines which can be used for selection. These three features, easy creation, distribution and selection of items, go a long way to enabling large scale item banks to be constructed and indeed there are a number of projects ongoing within UK HE which are doing exactly that.

At the same time the web has facilitated the creation and distribution of a great many computer mediated learning resources. Some of these resources could not be provided except in a computer-mediated environment (e.g. simulations) however a great deal are not novel (e.g. online coursenotes) although they may well be in a different format (e.g. textbooks vs. hypertextbooks). It is the explosion of access to material that has made these resources so critical. Finding appropriate, high quality resources has now become a significant challenge. Directories of learning resources, even those with ratings and/or reviews are of limited use, and can be extremely frustrating. Classification parameters are often limited and broad, while the reviews are often too short to provide meaningful advice, or too long to enable quick identification. Again IMS is providing metadata guidelines, and while these are to be welcomed, it is yet to be seen whether these will be taken up in resource classification as comprehensively as is required and if they will have a significant impact on resource identification.

This is essentially a theoretical paper. No system which does this is currently in available, although work is ongoing at the Robert Clark Centre at the University of Glasgow to develop a suitable item bank which can facilitate such a technique

Profiling students through item analysis

Large scale CAA increasingly involves students taking a variety of items extracted from a bank. These items are generally used several times over and can be given in different orders to different students, or indeed students can be given entirely different, yet equivalent tests ensuring the security of items. Items can also be selected real-time on the basis of previous responses to create adaptive tests. As students do not all take the same items, classical test theory Is not appropriate, however Latent Trait Theory approaches can be used to analyse students' responses. One such method of that of Rasch analysis (Wright and Stone, 1979)

It is beyond the scope of this paper to fully describe the technique, but essentially Rasch analysis involves the log linear regression of student responses to items, based on the assumption that a student with twice as much ability in an area has twice as much chance of answering a question correctly. It yields difficulty estimates for each item and ability estimates for each student, irrespective of the students who actually attempted the question, making it a suitable analysis technique for use in item banking where it is widely used. It is worth noting that in Rasch, ability and difficulty are measured on the same continuum and are essentially the same thing.

A Rasch analysis of a person taking a ten item dichotomous test, would provide a measure of the candidate's overall ability, together with a probability estimate for each question (i.e. how likely the candidate is to answer the question correctly) based on their ability estimate and the difficulty of the item.

Rasch is most generally used to ensure the quality of items entered into an item bank and provide statistics on them for future selection. It can however be used to identify questions which are biased, i.e. that some candidates have a greater chance of correctly answering the question, based on a factor other than their ability (e.g. gender, linguistic background etc). This is done by comparing the expected and actual achievement (residuals) of identifiable candidate subgroups, on the items; where these are significantly different the question can be said to be biased (McAlpine, 1999). As Rasch is a "symmetrical" system (i.e. that ability and difficulty are measured on the same scale) the notion of biased people, as well as biased items is also feasible. A mathematics student could, for example, be said to be biased toward algebra questions by over-achieving in this area as compared with his overall ability. Where questions comprise definable subgroups, such as algebra, geometry, number analysis etc. this can be a powerful tool for student profiling.

Through the analysis of residuals, person bias estimates can be obtained. These can indicate areas where candidates have particular strengths or weaknesses, and lead them to appropriate further study

Classification of learning resources

Typically current databases of learning resources rely on individual judgements. Resources are classified by either the developer, the person who maintains the database, or an independent evaluator. Classification is performed on a variety of different variables including, variously, subject area (e.g. Maths, Physics, English), quality of resource (usually a ranking on a linear scale), appropriate level of resource (which may be either age or educational level related), type of resource (e.g. hypertext book, simulation), and context that the resource is set in. Although the IMS Content specifications may standardise what the classification schema are, it does not address the fundamental arbitrariness of having a resource classified by an individual, or even a group of individuals

Some of the most effective directories on the web make use of automated rating systems. From commercial providers such as *Amazon.com*, to computer resource sites, such as *Cnet.com*, and online auction houses such as *Ebay.com*, they make use of the information that they hold about individuals to rate their resources. This can be done in a number of ways. On Amazon you are given the opportunity to see which other books have been bought by people who have purchased the same book as you. On Ebay, each trader is given a rating depending on other users estimations and on Cnet users are encouraged to rate the resource for quality, as well as submit comments.

This aggregation of individual details makes for a more comprehensive view of each resource. Rather than resources being ranked or classified by an individual, with individual biases, aggregated data allows for a more balanced overview. Although all of these techniques could be adapted for use by learning resource directories and indeed could increase their effectiveness, even greater effectiveness could be achieved by integrating the vast amounts of data which are generated by CAA, and routinely stored in both commercial and proprietary packages.

Facilitating Resource Selection

The results of student profiling obtained through a Rasch analysis can be used to identify the strengths and weaknesses of candidates. This data can be combined with data on the use of learning resources. Several systems, such as *Pathware* are able to track students' paths through material. Where regular assessment is built in to the paths, data can be obtained about which material is proving to be most effective in raising student achievement. Taken together with the profiles of candidates' strengths and weaknesses, this may prove an effective tool for identifying which resources are most suitable for each student, giving them an individual program of study, tailored to their needs.

The flow-chart shown in Appendix A shows the interaction of the items with the resources. It can be seen that based on the students' responses to items, a resource is selected, once that resource has been undertaken, a further assessment, using different items (although from the same areas), is given. Through this process a student profile is obtained and additional information is added to that already held on the resources and items available. The system is dynamic, both in that the data is constantly being updated to reflect curricular changes and changes in students abilities as well as being extendable. Further items can be added to the item bank, and resources added to the resource classification on the basis of a little data – confident that their inclusion will eventually calibrate or classify them correctly.

The classification of the learning resources using this method does not have to be done by an individual, but is automatically obtained using the feedback data from the assessments. Resources that prove effective in combating students' weaknesses in one area can be highlighted for selection whenever a student displaying those

weaknesses is identified. Resources which prove suitable only for students within a particular ability band (perhaps because they are designed to be particularly accessible, or, alternatively, require a broader range of knowledge that might only be available to the higher ability student) can also be identified, and highlighted for selection only when a student meeting the overall ability criteria, as well as the specific profile, is found.

Conclusion

It is time for computer assisted assessment to look beyond its own narrow confines and look at the wider context of how it can be best used to encourage student learning. The data generated by CAA is vast and extensive, but rarely used at all, far less to its fullest potential.

Classification of electronic learning resources, in a manner which facilitates easy identification of appropriate materials is a significant issue. It is usually currently undertaken by individuals using arbitrary schema. Although the use of IMS content specifications may standardise the parameters on which learning resources are classified, this does not address the fundamental arbitrariness of classification. Automated classification schemes using aggregated data have been successful in classifying online resources in a number of contexts and can be applied to this area. One of the most powerful ways of ensuring that the resources are linked to appropriate student learning is to ensure that the resources used are raising achievement in the area in which they are designed to cover. This can be done by using a Rasch analysis of computer assisted assessment data to profile student needs and link them with resources which address those needs. Software is available both to gather and analyse data gathered both from CAA and from student paths through materials. Linking these together would make for a powerful tool to promote the best use of appropriate computer based learning resources and overcome the classification problem which causes difficulties in selecting materials.

References

McAlpine (1999) *Gender and Performance on Geography*, British Education Research Association Annual Conference, Queens University, Belfast

Wright and Stone (1979) *Best Test Design: Rasch Measurement* Chicago: MESA Press

Appendix A: Flowchart of Data collection and use