A Tool for Supporting Graphical Interactive Assessment

Klaus BOEHM, Leonhard DIETZE
i3mainz, Institute for Spatial Information and Surveying Technology
FH Mainz, University of Applied Sciences
Holzstrasse 36, D-55116 Mainz, Germany
{boehm|dietze}@geoinform.fh-mainz.de

Abstract: Online-Assessment is widely accepted both for self-assessment and also as the basis for examinations. However its potential has not been fully utilized in current systems. Good instruments for innovative assessment are e.g. modern graphics based questions. Unfortunately their production is still very time and cost intensive. The GIA project develops a framework, which provides a solution that enhances the process of “authoring, providing and evaluation” of graphics based questionnaires. A new kind of question templates skeleton was developed in order to enable easy as well as time and cost reduced authoring of highly interactive graphical questionnaires. Moreover an innovative approach for fully automatic evaluation of the graphical questions was embedded in the new templates. Various graphical question templates were developed and tested in different kind of questionnaires – even in different domains, e.g. surveying as well as in medical education.

Introduction:

This paper describes the work of a currently ongoing research project “Graphical Interactive Assessment” (GIA) at the research institute “i3mainz”. The project that is funded by the German BMBF (Federal Ministry of Education and Research) concentrates on automatic generation and evaluation of online-tests with a focus on graphical interactive questions. The application environment, where the results of the research will be applied to is the area of higher education within the teaching of the university of applied sciences in Mainz and the medical continuous education (Boehm 2003). In this paper, first the motivation for investigating in this field will be outlined; second the general concept of GIA including data representation aspects will be sketched. Special attention within the paper is drawn on the new approach for the authoring and automatic evaluation of graphics based questions by the means of new innovative templates.

Motivation:

Online assessment is widely accepted both for self-assessment and also as the basis for examinations. A good example demonstrating the success of online quizzes is the GIScience program (Veenendaal 2001). A well accepted example of a tool for online tests is “Webassign” (Brunsmann 1999), it has been in use e.g. since 1998 by the “Fern-Uni Hagen”, the most popular Open University in Germany. Advantages arising from the use of such online-testing systems for authors are e.g. the reusability of test questions and better statistical feedback (Natal 1998). Most recently a very renowned university in Berlin introduced the first online test for the examinations in the medical education. Even though it’s launch in February 2004 (SPIEGEL 2003) was not successful due to server problems, it showed already the openness towards this new technology in the educational sector.

A lot of the current existing E-Learning and assessment systems focus on single or multiple-choice questions (MCQ) with several answers and radio-buttons to select the correct one. The forefathers of computer based questions were designed like this, because MCQ are fast and easy to build – and can afterwards be corrected automatically. These types of questions seemed (and still seem) to be the ones that some teachers and professors in the educational sector prefer. To use these kinds of questionnaires saves time and cost, especially with large groups of learners. Moreover, faculties like medicine traditionally have long MCQ-tests where the students e.g. have to find the correct definitions or names of muscles. All this is very easy to verify, because the
results are obvious. It should be noted that the pure application of MCQ is not undisputed – but the decision might be left to responsibility of the WBT author, or the responsible professor.

Reasons for new graphical based questionnaires (GBQ): Current pure text-based questionnaires are less intuitive for the user because they require mapping of the learned material and the subject of the question to predefined cases (a, b, c). This may subtract the concentration from the subject of the question. Moreover, it has been found out that learners are more eager to circumvent the typical MCQ instead of doing the tests in a proper way (learn first and then carry out the test). Another reason for this might be the fact that most of the current questionnaires are boring. Often they are enhanced by images that relate to the topic, but these questions are still simple MCQ with radio-buttons to select from. Alternatively to those very simple kinds of questionnaires, more innovative and well accepted solutions have been invented. Nowadays there are more types of questions that are asked already in computer-based tests. But mostly this has been done by the usage of Plug-Ins like Shockwave and Flash that allow questions with more graphical elements and interaction from the users. For example somebody is asked to drag an element to a specific position on an image. These questions show that there is a need for more interactivity and they are also a first step towards new concepts for GBQs.

Current solutions: A set of very powerful tools like e.g. the Macromedia Dreamweaver or the ToolBook Instructor allow to build complex questionnaires with innovative and content dependent graphical questions. The development of those questionnaires still requires profound skills in the usage of these tools. The effort can thus be calculated similar to full regular WBT learning courses. Additional development is needed for the presentation of exam statistics to the professor and to the learner. Thus innovative knowledge assessment by the usage of modern graphics based questions is still an expensive mean.

The new GIA Approach

The research performed within the GIA-project addresses this situation and tries to overcome the current limitations by developing a solution that supports the whole assessment circle by the use of adequate tools. The Editor and the Provision module can be seen in (Fig. 1).

Figure 1: Support of the assessment circle by the use of the GIA Editor and the Provision module.

The GIA framework will provide the following functionality:

- A web-based Editor for developing graphical interactive questions. It allows an author fast and easy building of these types of questions through the use of an extendable catalog of question templates.
- A question-browser that enables the author to manage the questions within the database. It lists the existing questions sorted by topics and therefore allows an easy reuse of them.
- A new concept to store graphical elements from within these questions in a database including the procedure for the automatic verification of them.
- A question provision module that displays the questionnaires to the learner by combining the template and the question-text with the overall layout.
- An evaluation module on the server for an automatic verification of the completed questionnaire.

Data Representation in the GIA Framework

In order to develop a framework as described above, a well-defined data representation as the core of the whole development had to be worked out. The data representation is the base for all tools such as the editor,
the provision module as well as the evaluation module. In order to ensure interoperability with other learning management systems, the application of an established standard or specification such as the Question & Test Interoperability (QTI) have to be considered.

Those existing standards/specifications provide a good basis for traditional MCQ. However, these typical MCQ differ from the GBQs that we are targeting in the GIA work. Two major differences can be identified. First, the graphical information for those questions is far more complex than the graphics information currently considered in the standards because it consists of a combination of raster, vector and time related data. Second, the evaluation of GBQs is more complicated than for MCQ, because it might not be possible to reduce the correct answer to a ‘case’. Therefore the verification cannot be a simple comparison. Instead, the evaluation will require evaluation scripts or routines which are specifically applied for each different question template and thus has to be mapped to the standardized format in addition to the questions. See (Fig. 2).

Application of Standards

There are several existing specifications or standards concerning the E-Learning sector proposed by organizations like the IEEE, W3C, ADL, AICC, IMS. The GIA Framework selected the “Question and Test Interoperability”-specification (QTI) currently specified by the IMS (IMS 2002). Rather than other specifications QTI focuses on material for tests and therefore suits more the needs of the GIA project. Beside the technical advantages, QTI is well accepted and is currently used respectively can be imported/exported in products of companies like Questionmark, WebMCQ and Macromedia.

Inclusion of Advanced graphics: The QTI-specification currently does not include questions with vector-data and other more advanced graphics. But simple graphical questions using raster images are integrated. This approach was adapted to fit our needs, while at the same time stick to the given specifications.

That was one task within this project: to map the contents of the new questions to QTI. The proposed conceptual GIA solution for including advanced graphical elements is using the SVG format (W3C 2003). SVG has the advantage, that it is an XML-Format for 2D-Vector graphics that can be stored like the course material itself. The SVG-elements therefore are mapped to QTI and together with the textual material inserted in the provided XML-structure. This XML-file then contains all the information necessary for examination and evaluation. It can therefore be written into the question-database for later retrieval by the provision module.

More detailed conceptual considerations regarding the data representation of the GIA framework are described in [Boehm, Dietze 2003].

Authoring and Evaluation

Background and Concept

It was already described above that the process of building graphical interactive questionnaires is relatively time-consuming for authors. The initial objective of our investigations was to reduce this time exposure. It shouldn’t take the author or professor more than several minutes per question. As a possible approach the skeletal structures of numerous questions were analyzed and generalized. A second step was the extraction of question-specific parts which were classified in different templates, the so called “question-templates”. The
general components in turn were integrated into the system. The question-templates not only consist of the components for the authoring and presentation, but also of the routines for enabling an automatic evaluation. The specific information that was defined for each question-template was then segmented into its five sub-components. They are described below and visualized in the left image in (Fig. 3):

- **Authoring**
  This component contains the form to insert the parameters needed for the question during the authoring process.

- **Appearance**
  The overall layout of the question is described within this component. It provides the framework of the question and embeds the other components.

- **Evaluation**
  This component contains the routines that are needed to evaluate a graphical question.

- **Graphics**
  This component contains the vector graphics that built the graphical part of the question as well as the underlying image that is used within some of the templates.

- **Action**
  This component includes the routines that catch the events from the learner and handle the reactions within the question. It controls the behaviour of the presentation component.

![Diagram showing the components within the GIA system](image)

**Figure 3:** The components within the GIA system

**Implementation**

The approach taken in the GIA-project to implement this modular concept was to build files that each inherit the tasks of one of these components. One question template thus consists of five files that are based on the basic components that were described above (see right image in (Fig. 3)). The template-files are implemented in PHP and build the presentation window using HTML for the frame and the 2D-Vector standard SVG (W3C 2003) for the graphical elements. The interaction is handled by JavaScript, which allows to build and interactively adjust the SVG-image by modifying the DOM (Document Object Model).

These template-files are separated from the other application logic in order to enable an easy creation of new questions independently of the core system. When a new question-template is defined, the files have to be linked into the system and written to the database. But no changes have to be made within the existing system files. This ensures the extensibility of the platform for new question templates.

The five files define the question template skeleton and they have to be adapted for each question-template to be developed. They incorporate the following tasks:

- **Authoring.php:** This file contains the html-form in which the textual information for each question including the parameters for the graphics-component have to be inserted. It is only needed during the authoring process.
• **Appearance.php**: It builds the html-frame with the question-details, a background-image where necessary, and embeds the “Graphics” as well as the “Action”-scripts. This file is used for the authoring and presentation process.

• **Evaluation.php**: This file includes the evaluation routines. These algorithms are only needed during the evaluation process.

• **Graphics.php**: The content of this file is the SVG information that is needed to build the SVG-DOM for each question. The file is used during authoring and presentation process and it is integrated in the “Appearance”.

• **Action.php**: It contains the routines necessary for the graphical interaction within the questions. As mentioned above, this interaction is realized by Java Script routines which modify the SVG-DOM. The file is used during authoring and presentation process and it is integrated in the “Appearance”.

### Development of New Question Templates

The concept above provides the skeleton for the development of question templates. The question-template-developer is requested to elaborate contents for all five files. It should provide the developer with sufficient freedom to build adequate presentation and authoring environments for a large variety of innovative and highly interactive questions. Since the evaluation of the questions is performed by algorithms, described by the template developer, there is virtually no limitation on this aspect in terms of complexity and extensiveness. When a new question is built based on one of the templates, the actual descriptive data of the concrete question is stored in the database of the core system.

The relations of the skeleton, a question template and a concrete question might be described by the terminology of the object-oriented paradigm.

<table>
<thead>
<tr>
<th><strong>OO-Terminology</strong></th>
<th><strong>GIA-Term</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract Class</td>
<td>Question Template Skeleton</td>
</tr>
<tr>
<td>Class derived from the Abstract Class</td>
<td>Question Template</td>
</tr>
<tr>
<td>Instance of a Class</td>
<td>Question built out of a Template</td>
</tr>
</tbody>
</table>

### Authoring in GIA

The authoring of questions and tests is performed by using the GIA core platform which supports the author by a set of possible templates, lists of own questions and questionnaires as well as functions for assigning the tests to the learner. For describing a new question, the author has to select a question-type out of the automatically generated list of question-templates. After that he has to insert the descriptive data such as the title, the elaboration of the question text and further needed parameters and metadata into the authoring-component. Now the appearance-, the graphics and the action-components are activated in order to allow the author to built the graphical part of the question. (Fig. 4) shows the authoring-component exemplarily for one question-type.

![Figure 4: The GIA Editor](image)
After having built several new questions, the author may want to create a new test (questionnaire) and include the new as well as eventually older questions. In this step the author also has to define the total amount of points that can be assigned in each question. The total of these points is then set as 100 percent of the points in the whole test.

Evaluation within GIA

After a learner performed a test, the automatic evaluation can take place. At that point in time, the evaluation algorithms described in the “Evaluation”-files are applied. The first step of the evaluation is to get the description of the correct answers defined by the author and the answers given by the learners out of the database. Then they are evaluated using the respective Evaluation-files where several calculations according the specific algorithms are done in order to find out if the question has been answered correctly. The results of this evaluation process are stored in the user’s account and may be checked by him. In contrast to multiple choice test, some graphical questions may leave space for interpretation of the learners answers. This fact, of course, increases with the flexibility within the questions. Therefore the GIA evaluation management allows the author to decide if there should a post manual evaluation be done after the automatic evaluation process. During this optional post manual evaluation the author views the graphical result of the learner together with the own suggested answer. The author may then alter the evaluation result suggested by the automatic process.

Example of the Authoring and Evaluation Process

As a proof of the concept an exemplary workflow for the process of authoring and evaluation for the “Area” question-template is described below. The “Area”-question requests the student to define a polygon by setting a number of points on a given image. The task might be to draw a traverse on an given street-map or a border around a specific area within an aerial picture.

- **Authoring:**
  During the authoring process the authoring-template prompts the author to insert the needed parameters: question-text, description and the metadata. He further has to upload the file of the background-image. The size of this image is scaled to the window-size of the template. Then the author has to insert the correct answer into this image, in this case a tolerance buffer for a polygon. This buffer is stored in the database as the “correct answers” to the question.

- **Answering the question by the learner**
  The polyline drawn on this image by the learner is stored as the “given answer” in the database.

- **Evaluation**
  The evaluation-template receives the correct and the given answers within two arrays. They are compared as shown in (Fig. 5) and can then be viewed together as drawn in (Fig. 6).

```
foreach segment gp in the given polygon
    foreach segment b in the buffer
        checkIntersection(gp, b)
```

Figure 5: Evaluation within the Area-template: correct buffer (black), given polyline (blue) and points of intersection (red).

The evaluation algorithm seeks for intersections between the buffer and the polygon given by the learner. Each segment of the given polyline is now compared with the buffer-lines for intersections.
Furthermore the length of the given answer is compared with the buffer. The sum of these criterions is used for calculating the numeric result of the answer. Thresholds per criterion define the requested minimum of correctness. The result of the automatic evaluation is shown to the author during the optional post manual evaluation step (see (Fig. 6)).

Further verification criteria and procedures still have to be developed, this will be done during further development of the system in the year 2004.

This workflow may differ for question-templates during authoring and evaluation. But all the specific calculations and the visualizations are only done within the components described above.

Currently several other question-templates have been developed. Most of them have been developed out of recurrent questions from examinations of the studies of surveying and geomatics.

![Figure 6: Screenshot of the „Area“- evaluation result](image)

**Current Results**

The GIA core system is running as a stable prototype. This prototype will be applied in the targeted application environment (higher education and medical domain) within the first quarter 2004. Currently nine different graphical question-templates have been developed. This includes one for the planning of an aerial flight, two for filtering a small raster image, one for evaluating the continuation of a graph, one for selecting a line within a text or source-code, one for evaluating the result of source-code. There are also several ones which allow to draw on a background-image: one which allows the author to mark a position, one for selecting one or several points and the above described “Area”-template.

The first user tests were carried out in November 2003 and January 2004. Altogether 52 students from different semesters (with different levels of experience) of studying geomatics at the university of applied sciences conducted the test. Although the number is not completely representative it provided important results and feedback, both to the concept and the system functionality. The students were very open-minded towards this new way of accomplishing tests. More than 80% of them liked these types of questions better than multiple-choice questions. A similar percentage of the students can also imagine to have such tests as assessment within the scope of lectures.

**Future Work**

The technological objective of the research project was the development of a running and proven core system and a framework for the development of new graphical question templates. It can be anticipated that this objective will be fulfilled and the provided software will be in a pre product phase at the end of the project in summer 2004. The cooperating company health&media GmbH ([www.health-media.de](http://www.health-media.de)) is planning to commercialize the results and to implement it in the medical domain in Germany.
The user tests that took place until now focused on the survey of the concept of the system and the usability for end users. Further user tests of the system will take place in spring 2004. In this last step of the development phase also the process of authoring should be evaluated by several professors from the study course of geomatics and surveying in Mainz.

With respect to the potential of automatic generated and evaluated graphical questions, the result of the project can be seen as the starting point for further investigations in various directions. Most important is the development and test of additional question templates. Here the development of appropriated evaluation criteria and routines might be most crucial. In addition to the use 2D graphics for questions, the inclusion of time dependent data, audio and video seem logical for the future development. Moreover integration of three dimensional scenes for questions should be considered.

The application of the current prototype within the academic test field as well as its evaluation within the business environment of the cooperating company has not requested the full implementation of the QTI export explained above. As long as the graphical questions are used within the GIA management system this functionality is not required. However it is foreseen to implement the concepts developed in the next version of the GIA.

References:


Spiegel online magazine (2003), article: Erstmals Medizinprüfung via Internet. Source: Internet (Nov 24th, 2003) (http://www.spiegel.de/unispiegel/stadium/0,1518,274123,00.html)
