Diagram Exerciser: A tool for learning data modelling

Pete Thomas
Department of Computing
Open University
Milton Keynes, UK
+44 (0)1908652695
p.g.thomas@open.ac.uk

ABSTRACT
This paper describes a software tool that has been developed to help students learn to draw entity-relationship diagrams in the context of data modelling. The tool, called the Diagram Exerciser, presents a student with a set of questions each of which describe a relational model of a scenario and enables the student to draw ER diagrams that represent the relational models. The Diagram Exerciser marks each attempt at a diagram and provides several feedback perspectives on the student's answer. The Diagram Exerciser has been evaluated with students and the results of the evaluation are presented in this paper. The evaluation reveals that the tool was found to be very useful and has provided information on which to base a revised version of the Diagram Exerciser with improved usability.

Categories and Subject Descriptors
K.3.2 [**Computer and Information Systems Education**]: computer science education.

General Terms
Design, Experimentation.

Keywords
Diagrams, automatic marking, feedback, revision tool.

1. INTRODUCTION
An Entity-Relationship diagram (ERD) is a graphical notation often used for describing a high-level data model of the type of information that is to be stored in a database. An entity represents a discrete type of object and is named as in: a book, an employee, a sound, or a chemical. A relationship captures how two or more entities are related to one another.

In ERDs there are two types of relationships: associations and subtypes (often informally referred to as 'has-a' and 'is-a' relationships, respectively).

Figure 1 shows a typical ERD containing both types of relationship and illustrates the notation used in the database course for which the Diagram Exerciser tool is intended. Entities are shown as round rectangles, associations are indicated by a named line joining two entities and a subtype relationship is shown with the subtype entity drawn inside the supertype entity. There can be recursive associations in which an entity is associated with itself: SlotIn is an example because ExternalEvent 'is-an' Event.

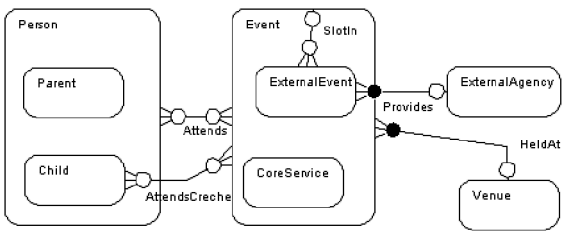

Figure 1. A typical ERD.

An association can have attributes. The degree of the association (one-to-one, one-to-many or many-to-many) is indicated by 'crowsfeet', and the participation conditions (mandatory or optional) are indicated by closed or open circles.

The Drawing Exerciser tool contains a number of data modelling questions, presented as relational models of scenarios, about which students attempt to draw representative ER-diagrams. The tool grades a student's diagram by comparing it with a specimen solution and provides feedback on the accuracy of the diagram. Furthermore, the tool allows the student to interrogate the specimen solution. The tool is intended primarily for formative assessment, to be used following instruction in data modelling techniques [14].

The tool described in this paper is an application of two strands of research (1) the machine interpretation of diagrams and (2) the automatic marking (grading) of assignment and examination questions [6, 7, 8, 10, 11].

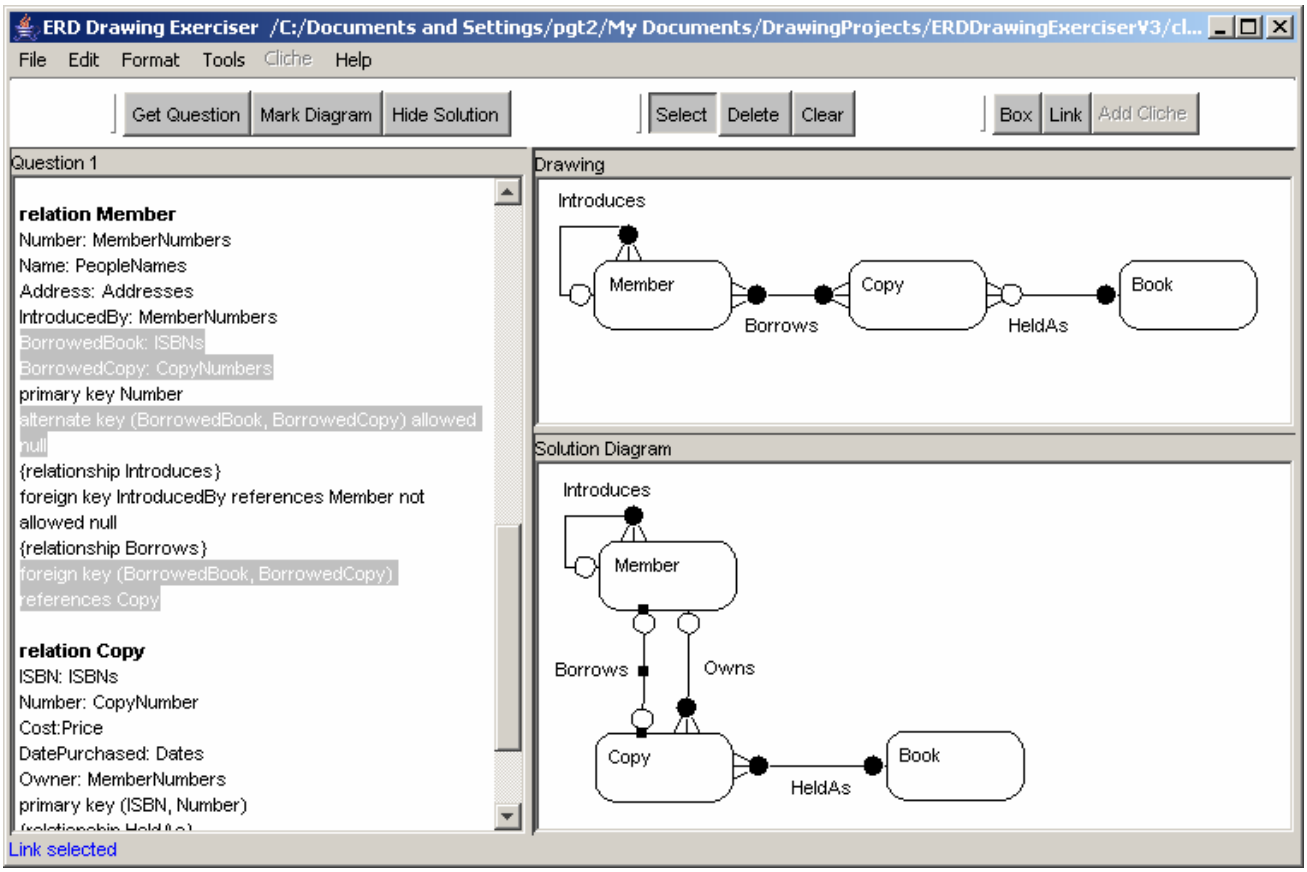

Figure 2. The Diagram exerciser Interface.

Our investigations of machine interpretation of diagrams have resulted in a framework [5] which has been exploited in a tool to automatically mark student answers to assignment and examination questions involving diagrams [10].

Trials of the automatic marker have provided confidence that our marking algorithm performs well enough for formative assessment [7].

There is growing interest in diagram interpretation and machine (assisted) grading of diagrams. Batmaz and Hinde [2] have a system that helps the human grader to assess conceptual database diagrams and Tsintisfas [13] has produced a framework for assessing diagrams.

This paper is structured as follows. Section 2 describes the capabilities of the Diagram Exerciser, section 3 reports on the evaluation of the tool and section 4 discusses future developments.

2. THE DIAGRAM EXERCISER TOOL

Figure 2 shows the user interface of the Diagram Exerciser tool. When opened, the tool presents the student with a question in the form of a scenario (in left-hand Question pane).

The student is expected to draw an appropriate ER-diagram that models the given scenario in the top-right Drawing pane. During the diagram construction phase the diagram in the Solution Diagram pane (bottom-right) is kept hidden. The solution diagram, a specimen answer to the question, is only revealed when the student presses the Show Solution button (which is relabelled as Hide Solution once the solution is revealed).

The Exerciser incorporates a drawing tool for drawing ER diagrams that has been deliberately designed not to be overly helpful because we are aiming to provide a tool that will enable students to express their own understanding of data modelling – which could be less than perfect – and then to provide feedback on this understanding. We do not want a tool that supports the drawing of correct diagrams. This fits in with our notion of imprecise diagrams used in our work on machine interpretation of diagrams [5].

The student is encouraged to attempt a question by first drawing a diagram and then submitting their diagram for marking (grading) by clicking on the Mark Diagram button. The immediate consequence of the marking is a simple dialog box, an example of which is shown in Figure 3. The student can ask for further details, and would be presented with a feedback window similar to that shown in Figure 4.

The feedback window shows the correspondences found by the tool between the relationships in the student's diagram (answer relationships) and those in the specimen solution (solution relationships).

There is a second level of feedback in which the student's answer diagram is shown overlaying the specimen solution diagram as illustrated in Figure 5.

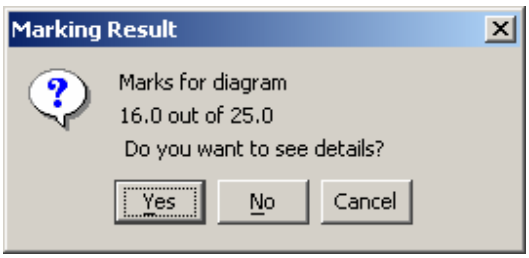

Figure 3. The Marking Result.

The aim is to identify those parts of the student answer which have been matched with the specimen solution, indicated by solid lines and circles. Those parts of the specimen solution which do not appear in the student's answer are indicated by dashed lines and grey circles. Any extraneous parts of the student answer are listed separately (none in Figure 5).

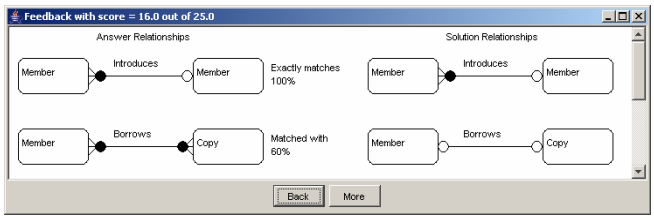

Figure 4. The Feedback – First Level.

Students can revise their diagrams and resubmit them for marking as many times as they like. Alternatively, the student can reveal the specimen solution and interrogate it. This is achieved by selecting an element of the solution diagram after which the tool highlights those parts of the scenario which give rise to the selected diagram element (an example can be seen in Figure 2 where the Borrows relationship has been highlighted as have the relevant parts of the question text). Right-clicking on a diagram element reveals a properties window which gives more details of the diagram element and its derivation (see Figure 6).

3. EVALUATION OF THE EXERCISER

3.1 The evaluators

The Diagram Exerciser was distributed to volunteer students who were studying a database course. Twenty-eight students responded to a questionnaire designed to elicit opinions of the usefulness and usability of the tool. The evaluation took place immediately prior to a summative assessment in which students were asked to draw an ERD.

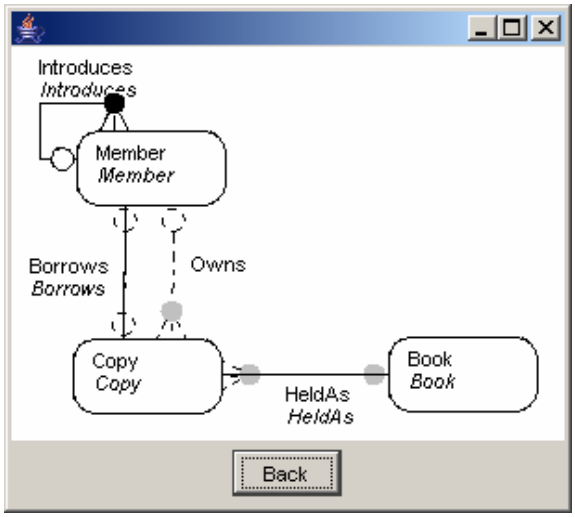

Figure 5. The Feedback – Second Level.

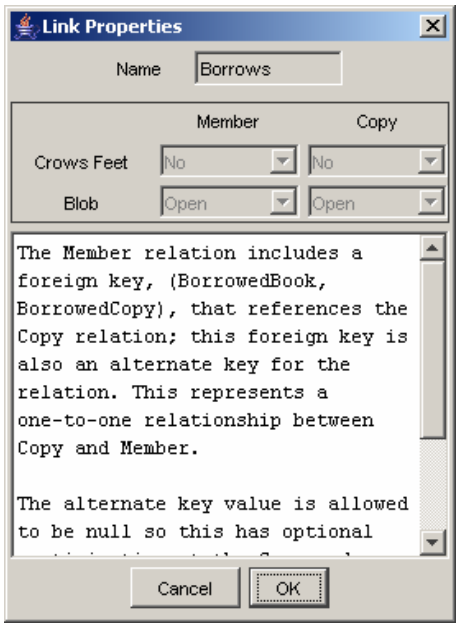

Figure 6. A Properties Window.

3.2 Usefulness

The tool was supplied with 10 reasonably demanding questions for students to attempt (the questions were of an equivalent standard to those normally found in assignment questions).

On average, the students attempted over 4.4 questions each which is certainly enough to gain a good impression of the usefulness and usability of the tool. The main reasons given by students for not attempting more questions were either that they had completed enough questions to satisfy themselves that they understood ER-diagramming adequately or that they had insufficient time to do more because of the need to complete an

imminent assignment (although several students said that they would be returning to the tool during their revision).

We were satisfied that the responses to the questionnaire were based on a sufficiently good understanding of the tool and its capabilities.

All the students who returned the questionnaire said that the tool was helpful in their studies and 6 volunteered that the tool was *very* helpful. All students said that they would recommend the tool to other students with 6 volunteering that they would *definitely* recommend the tool. Most students provided additional positive comments on the usefulness of the tool.

Typical responses were:

"I found it a great learning experience."

"I love it – being able to seeing where parts of the diagram come from is magic."

"I particularly liked the feature in 'Mark diagram' that gave a % of match to the solution – it gave me hope!"

"I enjoyed using the software and thought that it was very helpful in enhancing my understanding of E-R diagrams."

"I liked the explanation … describing the origin of the selected element."

"It was the final piece of the jigsaw for me. It was a good opportunity to test understanding of the theory of the course."

While the questionnaire had been kept quite short to encourage feedback, the general remarks covered all of the tool's main functionality with at least one appreciative comment.

3.3 Usability

The average time to become familiar with the tool was just over 13 minutes with most students quoting 10 minutes. This was a good result because students were provided only with a brief description of the tool's features (little more than appears in this paper) and had to rely on the tool's help facility for some details. 17 students used the help facility but 9 said that they found the tool sufficiently intuitive that they did not access the help.

Generally, the information needed from the help facility was confined to only a few drawing features: how to select the degree and participation conditions of a relationship, how to re-route a link for aesthetic reasons, and how to draw recursive relationships. This information was invaluable in revising the interface.

Table 1 summarises the responses about ease of use.

Table 1. Ease of use statistics

Easy to Use	Not Easy to Use	No response
21	3	4

The responses also indicated that the tool contained a number of faults (software errors) as well as some annoying behaviours (design faults).

In addition, the instructions for the tool (accessed via the tool's help menu) contained errors and were incomplete in places. It is remarkable that these students did not give up using the tool. However, no student reported that the issues with usability detracted from the usefulness of the tool.

3.4 Further evaluation

Following the evaluation discussed above, the tool was revised, primarily to remove bugs and improve usability (the user interface shown in Figure 1 is actually the revised interface). The students who had responded to the first questionnaire were given the opportunity to test the revised version at the time they were revising for their final exam. Only 11 students responded to the second questionnaire, but they were uniformly in agreement that the tool was easier to use. Typical comments were:

"Pleasure to use, thank you."

"I think this is a very useful tool and I used it extensively for revision…"

"… [I] would suggest it is offered to all students."

"[I] will definitely be using it next year for my project"

"Overall, a very useful and helpful tool. I could save hours if I could produce an assignment ER diagram using this tool."

"I found it easy and intuitive. It helped me with my revision. … I felt much more confident during the exam than I think I would have been."

The responders also provided suggestions for further functionality and usability enhancements.

4. CONCLUSION

We conclude that the students who responded to the questionnaire found the tool to be very useful despite the usability deficiencies in the first version. There was little criticism of the main functionality of the tool and indeed many students said that wished such a tool had been available as part of the course materials. The overwhelming positive reaction to the tool together with the fact that all respondents said that they would recommend the tool to other students gives us confidence that there is a need for such a tool and that the current implementation has appropriate functionality and adequate usability.

5. RELATED AND FUTURE WORK

The Exerciser works with three distinct items of data: a question (also known as a scenario), an ER diagram, and a set of associations (relationships) between the question and diagram on which feedback is based. The construction of a question and its associations with a diagram is performed by a separate tool called the Tutor Tool. The Tutor Tool enables the teacher to describe the diagram in terms of the scenario and build associations which form the basis of the teaching in the Exerciser. The Tutor Tool is being used to create a new set of questions for the new database course.

The automatic marking tool and the feedback mechanism developed for the Diagram Exerciser are being exploited in a tool to help human markers. We have detected that in some circumstances the way in which a student's diagram is presented can affect the marker's perception of the diagram. Therefore, we

are experimenting with re-drawing a student's diagram in a form which more closely resembles the specimen solution but without changing any of the original relationships.

It should then be quicker for the human markers to recognise both correct and incorrect portions of a diagram, and hence be able to mark diagrams more quickly and more accurately.

The tools we have discussed here are applications of our work on diagram understanding which we are continually improving and extending. This should enable us to provide better automatic marking tools and improved feedback mechanisms for the Exerciser, an area of wider interest [2].

A significant part of the marking process is the matching of entity names between a student diagram and the specimen solution in which the recognition of synonyms is crucial. In some circumstances the name of an entity can be viewed as synonymous with another entity name when it is suspected that they are hyponymous. However, when explicit subtype relations exist in either the student answer or the specimen solution, the names need to be distinguished and not be treated as synonyms. The current synonym handling algorithm works well in the absence of subtype relationships, and an enhanced algorithm is being implemented to deal with explicit subtype relationships.

The Diagram Exerciser and the associated Tutor Tool are scheduled to be tested with the first cohort of the new database course. This will include both students and their tutors. The feedback from this testing will be incorporated into a final version of the tool which will be subjected to our rigorous quality assurance procedures prior to deployment on the new course in 2008.

6. ACKNOWLEDGMENTS

Part of this work was undertaken while one of the authors held a UK HEFCE funded Centre of Excellence for Teaching and Learning Teaching Fellowship of the Centre for Open Learning in Mathematics, Science, Computing and Technology at the Open University, UK.

7. REFERENCES

[1] Anderson, M., McCartney, R. (2003) *Diagram processing: Computing with Diagrams*. Artificial Intelligence **145** (1-2) 181-226.

[2] Batmaz, F. and Hinde, C.J. (2006) A Diagram Drawing Tool for Semi-automatic assessment of Conceptual Database Diagrams. In *Proceedings of the 10th Annual International Conference in Computer Assisted Assessment*. Loughborough University, Loughborough, UK, July 2006, 68-81.

[3] Manning, C.D. and Schutze, H. (2002) *Foundations of Statistical Natural Language Processing*. The MIT Press, Cambridge, Massachusetts. ISBN 0-262-13360-1.

[4] Marriott, K., Meyer, B., Anderson, M., Cheng, P. and Haarslev, V. (2000) Non-standard logics for diagram interpretation. In *Proceedings of Diagrams 2000, International Conference*, Edinburgh, ISBN 3-540-67915-4-42-57.

[5] Smith, N, Thomas, P.G. and Waugh, K. (2004) Interpreting Imprecise Diagrams. In *Proceedings of the Third International Conference in the Theory and Application of Diagrams*. March 22-24, Cambridge, UK. Springer Lecture Notes in Computer Science, eds: Alan Blackwell, Kim Marriott, Atsushi Shimojima, 2980**,** 239-241. ISBN 3-540-21268-X.

[6] Thomas, P.G. (2003) Evaluation of Electronic Marking of Examinations, In *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education* (ITiCSE 2003), Thesaloniki, Greece, 50-54.

[7] Thomas, P.G. (2004) Drawing Diagrams in an Online Exam, In Proceedings of the 8th Annual InternationalConference in Computer Assisted Assessment. Loughborough University, Loughborough, UK, 403-413.

[8] Thomas, P.G., Waugh, K., Smith, N. (2005) Experiments in the Automatic marking of E-R Diagrams. In *Proceedings of the 10th Annual Conference on Innovation and Technology in Computer Science Education* (ITiCSE 2005), Monte de Caparica, Portugal, 158-162.

[9] Thomas, P.G. (2003) Evaluation of Electronic Marking of Examinations, In *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education* (ITiCSE 2003), Thesaloniki, Greece, 50-54.

[10] Thomas, P.G., Waugh, K., Smith, N. (2005) Experiments in the Automatic marking of E-R Diagrams. In *Proceedings of the 10th Annual Conference on Innovation and Technology in Computer Science Education* (ITiCSE 2005), Monte de Caparica, Portugal, 158-162.

[11] Thomas, P.G., Waugh, K., Smith, N. (2006) Using Patterns in the Automatic Marking of ER-Diagrams. In *Proceedings of the 11th Annual Conference on Innovation and Technology in Computer Science Education* (ITiCSE 2006), Bologna, Italy, 403-413.

[12] Thomas, P.G., Waugh, K. and Smith, N. (2006) An approach to the Automatic Grading of Imprecise Diagrams. *Computing Department Research Report* 2006/16, Open University.

[13] Tsintsifas A. (2002) *A Framework for the Computer Based Assessment of Diagram-Based Coursework*, Ph.D. Thesis, Computer Science Department, University of Nottingham, UK.

[14] Waugh, K.G., Thomas, P.G., Smith, N. (2004) Toward the Automated Assessment of Entity-Relationship Diagrams. In Proceedings of the 2nd LTSN-ICS Teaching, Learning and Assessment in Databases Workshop, Edinburgh. http://www.ics.ltsn.ac.uk/pub/databases04/index.html (accessed 15/01/06).