

Teaching and Learning Applications Related to the Automated Interpretation of ERDs

Kevin Waugh, Pete Thomas and Neil Smith
*Department of Computing, Faculty of Mathematics and Computing
Open University, Milton Keynes, UK, MK7 6AL
{k.g.waugh, p.g.thomas, n.smith}@open.ac.uk*

Abstract

In this paper we report ongoing work in the development of diagram interpretation and assessment applied to entity-relationship diagrams (ERDs). We report on our development of an automated marking tool and the results of two large-scale experiments with this tool. We give an overview and the results of some initial, informal, evaluation of a student revision tool that we developed as a result of our experiences working on the automated marking system.

1. Introduction

There is growing interest in diagram interpretation and machine (assisted) grading of diagrams. Batmaz and Hinde [1] have a system that helps the human grader to assess conceptual database diagrams and Tsintifas [2] has produced a framework for assessing diagrams.

We have been developing an approach to the interpretation of diagrams that is based on the identification and recognition of minimal meaningful units (MMUs) within a diagram [3]. We consider a complex diagram to be a collection of meaningful units (MUs) formed from combinations of MMUs. This interpretation is comparable to the way in which transformational grammars for natural languages [4] represent complex sentences as transformed combinations of base sentence forms.

We have focussed our work on the development of an automated assessment tool for Entity Relationship Diagrams (ERDs) [5]. When working in automated assessment it is necessary to recognise both correctly drawn and incorrectly drawn diagrams. If the assessment process is to offer qualitative feedback than

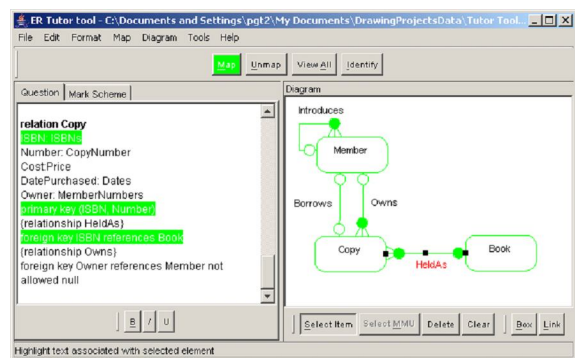
it is necessary to interpret the incorrectly drawn diagrams and comment on the possible mistakes that may have led to their production.

While developing our automated marking tool we have considered additional teaching and learning support tools based around our diagram interpretation approach. Our main result is the development of a learning support tool for student's independent use.

In the remainder of this paper we will give a brief outline of the automated marking tool and report on two experiments we undertook with the tool. In the final sections of the paper we will look at the tool we developed to support independent student practice and revision.

2. The diagram drawing tool

In order to capture ERDs in a machine readable form we developed a basic drawing tool component which has subsequently been embedded in the tools we produced, see Figure 1. Devising our own drawing tool enables us to control the format of the internal



files and to customize the notations to those the student is already familiar with from their study.

Figure 1. The drawing tool, here embedded in an ER Tutor tool, consists of the right hand pane and the buttons beneath it.

To avoid prompting students, the terminology chosen for features in the drawing tool avoids the use of the standard ERD terminology of entity type, relationship, degree and participation condition. However, the drawing tool does draw syntactically correct diagrams; relationships must begin and end on an entity type box, and the tool places the degree and participation condition features in the correct locations on the relationship lines. Text used for labelling entity types and relationships is unconstrained, so students can choose their own names for the features in the diagram they draw. The drawing tool does not attempt to interpret the diagram drawn and consequently the student can make errors of omission (for example, by failing to supply participation conditions) and semantic errors within the syntactically correct diagrams.

3. Experiments in the automated marking of ERDs

After an initial feasibility experiment reported in [5, 6] we undertook a large scale experiment using diagrams produced by students in an unseen, end of course examination. The model solution used in this experiment used only entity types and relationships. We followed this experiment with another experiment using examination question that included entity subtyping. These experiments are discussed in the following section.

Our marking approach [6], embedded in the marking tool, consists of a series of identification and matching stages between a student answer and one or more sample solutions.

For each sample solution and alternative marking scheme:

- Identify the MMUs present in the student answer and the sample solution.
- Align the MMUs present in the diagrams by finding the best overall match including similarity measures for imprecise text matching.
- Find best match between the structures of the MMUs not fully matched in the alignment stage.
- Compute similarity measures for each matching MMU and for the diagram as a whole.
- Apply a marking scheme, using the consistent matches and overall similarity measures to compute the mark to be awarded.

The advantage of computing a numeric similarity measure is that it enables the graceful handling of

imprecision in a student's diagram. Student's frequently deviate from the use of the names given in an exam question; for example by misspelling words, abbreviating words or using synonyms. Currently we use edit distance between the names as the basis of the similarity measure. This simple scheme has the advantage that trivial spelling errors can be taken into account ('Empoyee' is very similar to 'Employee'). However, the simple scheme really only works well when the student uses simple names that are close to those used in the specimen solution.

Dealing with more complex naming requires synonyms to be handled effectively. In addition, the names assigned to entity and relationship types can be 'composite' – formed from several words – to provide more meaning than would be conveyed by a single word (for example, 'HeldAs' in Figure 1.) Therefore, we have modelled entity names as noun phrases and relationship names as verb phrases. The subsequent processing attempts to extract the noun from the noun phrase and the verb from the verb phrase and bases the similarity of names primarily (but not exclusively) on the nouns and verbs. (Stop words are discarded.)

Further processing attempts to deal with abbreviations, which are common in data modelling especially when students are under time pressure as in an examination. Hyphenation that does not follow normal conventions is also popular. The final form of a name is a list of words each one of which is stemmed to deal with differences in parts of speech (so, 'receives', 'received' and 'receiving' would all be considered synonymous).

Since an aim of our research is to produce usable systems that can be incorporated into teaching and learning tools which execute on typical PCs with reasonable response, we have to balance the extent of the natural language processing with performance.

3.1. Experiment one

The examination question gave students a textual description of a scenario and required them to produce an ERD to capture the data requirements presented in the scenario. Students answered the questions under exam conditions using traditional pen and paper scripts. Human markers (a team of 7 were used) employed a marking scheme based on a sample solution ERD, shown in Figure 2.

When the scripts were released by the examinations office, the drawing tool was used to transcribe the student's diagrams into electronic form (2 scripts from the 593 available for the experiment could not be redrawn using the drawing tool, these were omitted from the experiment). Layout and

positioning of the diagrams were preserved during this transformation and a verification stage undertaken to ensure accuracy in the transcription process.

The original marks awarded by the exam markers were recorded with the transcribed diagram to form a corpus of exam answers and marks. A moderation stage checked the exam mark with those of a second marker to identify any errors in the original marking process. The corpus of moderated diagrams was then used in our marking experiments.

Our corpus consisted of a total of 591 diagrams which we split into a development set, containing 197 diagrams, and a test set, containing 394 diagrams. The

question carried seven marks and the human markers produced the distribution of marks shown in Figure 3.

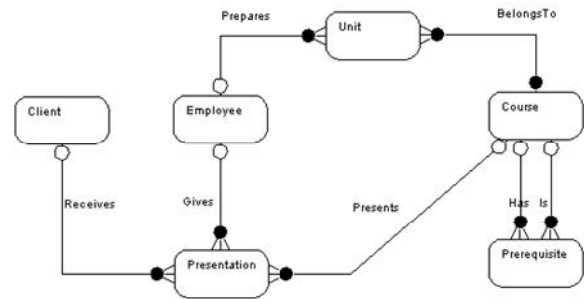


Figure 2. The ERD model solution used in our first experiment.

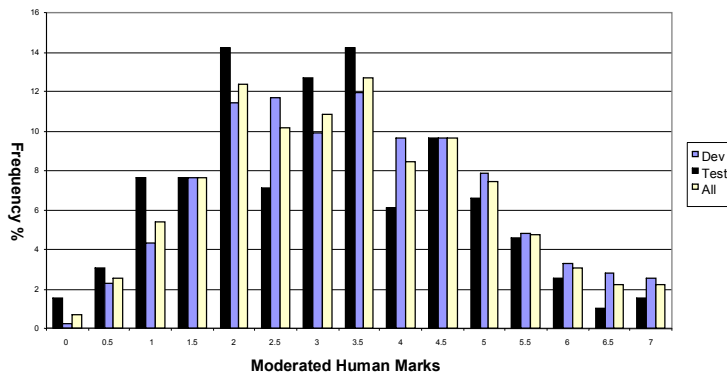


Figure 3. The distribution of marks awarded by human markers in the first experiment.

The result of applying our automated marking tool to the sample set is shown in Table 1. Further analysis of the results showed that the distribution of marks was comparable to that of the human marker across the entire range of marks, and that the rank ordering of scripts was comparable to that produced by the human markers.

Table 1. First experiment: differences between the automated marker and the human markers. (Set size = 394 diagrams, 7 marks available)

Marking difference (magnitude)	0	0.5	1.0	1.5	2.0	2.5	3.0	>3.0
Number of scripts	271	89	32	1	0	1	0	0
% of scripts	68.8	22.3	8.1	0.26	0	0.26	0	0.0
Cummul. %	68.8	91.4	99.5	99.8	99.8	100	100	100

3.2. Experiment two

We performed a second experiment using a corpus drawn from another examination. This question also gave a short scenario and required the students to represent the data requirements for that scenario using an ERD. The solution required the inclusion of entity sub-typing as shown in Figure 4.

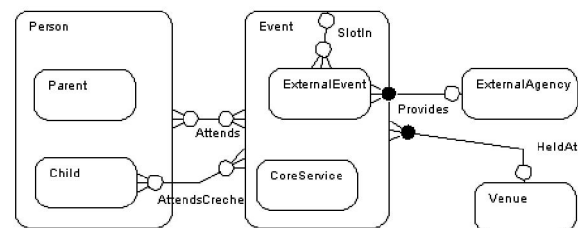


Figure 4. The ERD model solution used in our second experiment.

The full set of examination scripts has not yet been made available to us so our corpus for this experiment contains only 30 scripts, although marking samples from six markers were present in this sample.

With such a small sample of scripts we used the entire sample as our test set. When the full set of examination scripts is released we expect to be able to work with over 500 scripts for this examination question.

Table 2 shows the result of marking the second examination question; this question also had seven marks available. As with the first experiment the range of the marks awarded by the human markers and the rank ordering of scripts was comparable.

Table 2. Second experiment: differences between the auto marker and human markers. (Set size = 30 diagrams, 7 marks available)

Marking difference (magnitude)	0	0.5	1.0	1.5	2.0	>2.0
Number of scripts	11	18	0	1	0	0
% of scripts	36.7	60	0	3.3	0	0
Cummul. %	36.7	96.7	96.7	100	100	100

While the number of exact matches is not as good as in the first experiment (36.6% against 68.8%) the percentage of scripts differing by no more than 0.5 marks is comparable.

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

4. Student revision tool

To maximise the benefit to students and support their learning assessment mechanisms should offer both quantitative and qualitative feedback. Our experiments have shown that quantitative feedback (marking) is achievable within the approach we have followed. To further our work we therefore considered the development of a tool to give structured feedback to students in a formative setting. We have informally evaluated the tool with a small, self-selected, group of students.

The Drawing Exerciser tool (the implemented name for the Student Revision Tool) contains a number of data modelling questions, presented as relational models of scenarios, about which students attempt to draw representative ERDs. 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.

We intend this tool to be used by students who had previously been taught how to read and generate ERDs and how to translate ERDs into relational model representations. In our database course we have asked students to demonstrate their understanding of the relationship between ERDs and the relational representation by including coursework questions that ask them to reverse engineer an ERD from a given relational model (this is in addition to the more usual mapping from ERD to relational model questions). This is a different, and more constrained, form of question than that used in the two examination-based experiments reported earlier. This form of question limits the range of correct solutions the student should produce, both structurally and textually – effectively there is only a single correct ERD for each relational model supplied. We initially wanted to explore student experiences with the tool and were not confident that we could accurately process the more open, scenario-based, questions of the kind used in the experiments reported earlier. We expect to extend the tool to include both forward and reverse engineering style questions in the near future as we are now confident that the marking tool used in the examination marking can be usefully embedded in the revision tool.

4.1. Student revision tool, description

When opening the revision tool students are initially presented with the text of a question and a diagram drawing tool window in which they can draw their diagrams, see Figure 5 (when first opened the lower right hand solution pane is empty). At any time they can choose to have their attempt marked and, if they decide not to develop their solution further, they can examine the sample solution for comparison purposes. If they return to their drawing without opening the sample solution they can either redraw their solution, or ask for feedback on the mark currently awarded and use that feedback to attempt to improve their solution.

Feedback to the students is available in a number of forms. The basic mark awarded for their solution can be shown (Figure 6). This can be supplemented with a detailed breakdown showing a comparison between the student's solution and the model solution (Figure 7), or as an overlay of the student attempt onto the solution diagram (Figure 8).

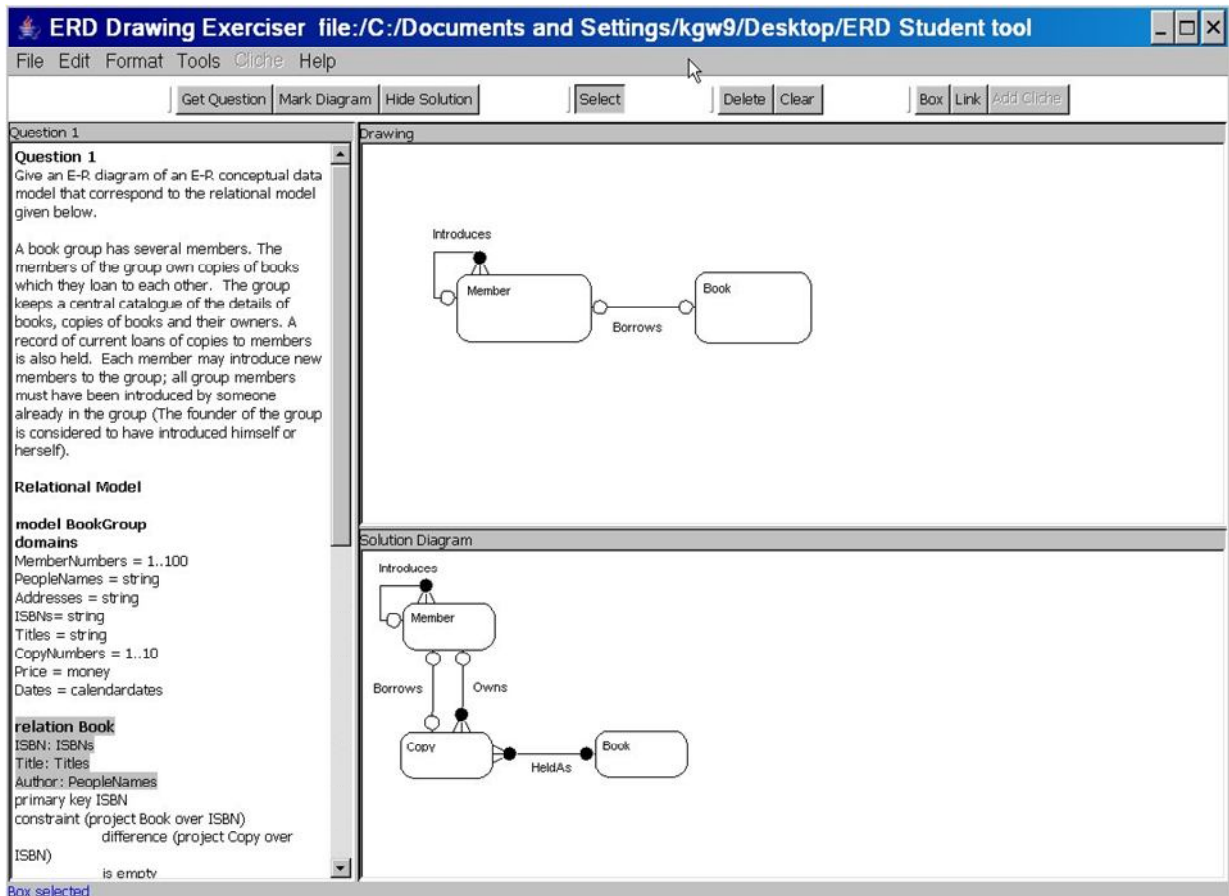


Figure 5. The revision tool showing various active windows; on the left is the question pane, on the top right is the student drawing pane, on the lower right is the sample solution diagram.

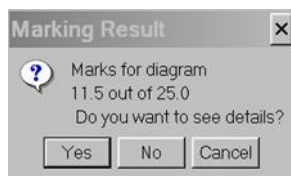


Figure 6. Feedback showing marking result.

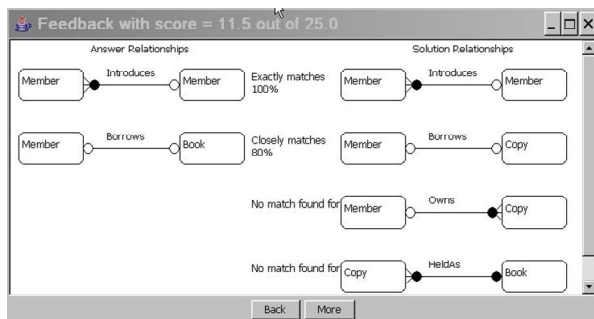


Figure 7. Feedback showing comparison between student's solution and sample solution.

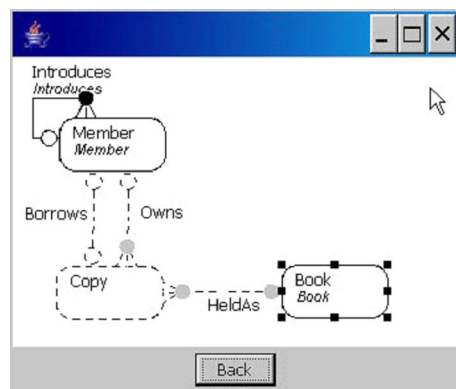


Figure 8. Feedback showing the overlay of the student solution on to the sample solution, the italicized text is from the student solution.

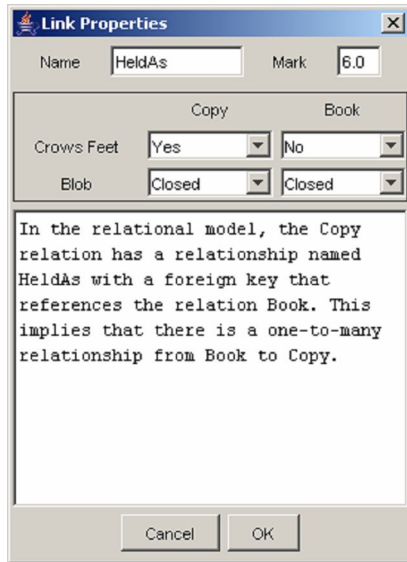


Figure 9. Feedback showing a textual description of why a feature appears in the sample solution.

If the student has chosen to reveal the sample solution then it becomes possible to interrogate the sample in two different ways. The student can highlight a feature in the sample diagram and this will result in the corresponding text in the question pane being highlighted. Students can see where in the relational model the evidence for the inclusion of each feature in the diagram can be located. In Figure 5 this is shown for the Book entity type in the sample solution diagram which has resulted in the corresponding text in the question being highlighted. In addition to this highlighting the student can get a textual description of the evidence for a selected feature. Figure 9 shows the feedback for the HeldAs relationship also shown in highlighted form in Figure 1.

4.2. Student revision tool, informal user evaluation

The Diagram Exerciser was distributed to around 80 volunteer students who were studying a third level 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. At this time we have made no attempt to correlate student use of the tool against performance in summative exercises. Our initial evaluation aim is to assess the usability of the tool and the students own perceived usefulness of the tool and the feedback it

gave. Future versions of the tool will be subject to a more rigorous evaluation of their benefits.

4.2.1 Usefulness. The tool was supplied to students with 10 questions of an equivalent standard to those normally found in coursework on our 3rd level database course. One of the questions was taken from the coursework set for the 2003 presentation of the course and all questions were set by the same author.

On average, the students attempted over 4.4 questions each which is 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. One student indicated that the feedback from the tool was not sufficiently detailed to support their revision requirements.

Typical responses were:

“I particularly liked the feature in ‘Mark diagram’ that gave a % of match to the solution – it gave me hope!”

“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.

4.2.2 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 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. Three students reported that they did not find the drawing tool easy to use with two stating that they only attempted a few of the questions because the drawing tool was too difficult, or took too long, to use.

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.

The responses also indicated that the tool contained a number of faults (software errors) as well as some annoying 'behaviours' (design and usability flaws). These responses included requests for 15 additional features or drawing capabilities. This information was used to improve the tool.

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.

4.2.3 Further evaluation. Following the evaluation discussed above, the tool was revised, primarily to remove bugs and improve usability (the user interface shown in Figure 5 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:

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

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

"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.

5. Future work

We will be continuing to develop our work on diagram understanding and its application to automated marking. However, as we currently do not use a framework for computerised submission of

examination scripts it is unlikely that the automated marking tool will be used to directly process examination scripts. We expect that examinations will continue to be the source of our experimental data.

The revision tool is well received by students and the nature of the questions can be controlled sufficiently that a high degree of accuracy in marking and directed feedback can be achieved. In 2007/8 we intend to offer the revision tool, with tutor support, to the students enrolled on our 3rd level database course.

At the Open University we do use electronic coursework submission with on-line marking and feedback by tutors. We are also increasing the provision of on-line supported tutor-student interaction. We are therefore developing tools which could be used in such a framework. Student's formative and summative coursework (non-examination) are already submitted in an electronic form. The drawing tool offers a mechanism to easily capture student attempts to draw ERDs and so can augment any electronic submission of coursework. This environment offers an opportunity to further develop tools covering student and tutor support for ERD assessment. As suggested earlier, the nature of our general coursework questions are too open for fully automated marking to be considered robust. We are continuing to work towards an automated marking system that will enable more open questions and a wider range of student solutions to be marked [7]. However, we have also identified scope for tools to assist human marking of summative coursework, and to support tutors producing formative student assessment.

5.1 Supporting human markers

The automatic marking tool and the feedback mechanism developed for the Diagram Exerciser are being exploited in a tool that we hope will support the marking of electronically submitted coursework. We have identified that in some circumstances the way in which a student's diagram is drawn 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 (Figure 10 and Figure 11).

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 possibly more accurately.

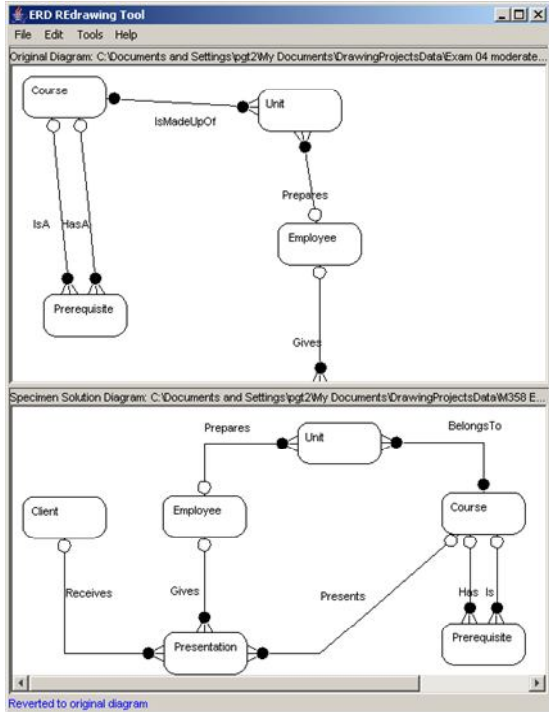


Figure 10. ER Redrawing tool showing sample solution and student's original submission.

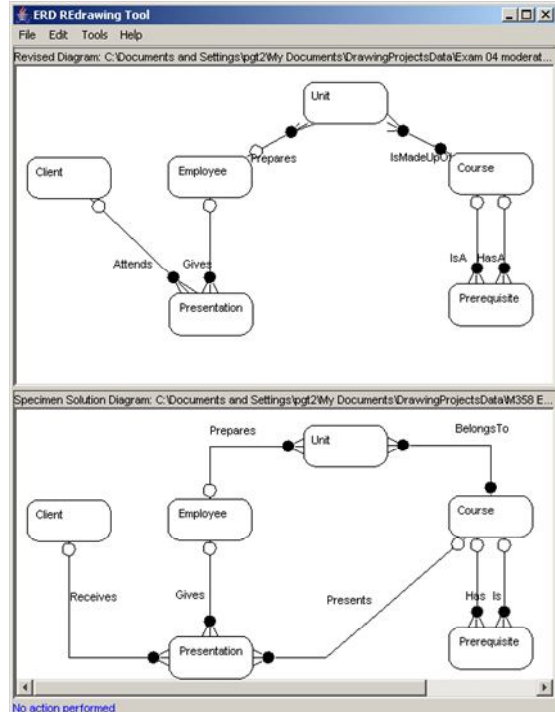


Figure 11. ER Redrawing tool showing redrawn student submission.

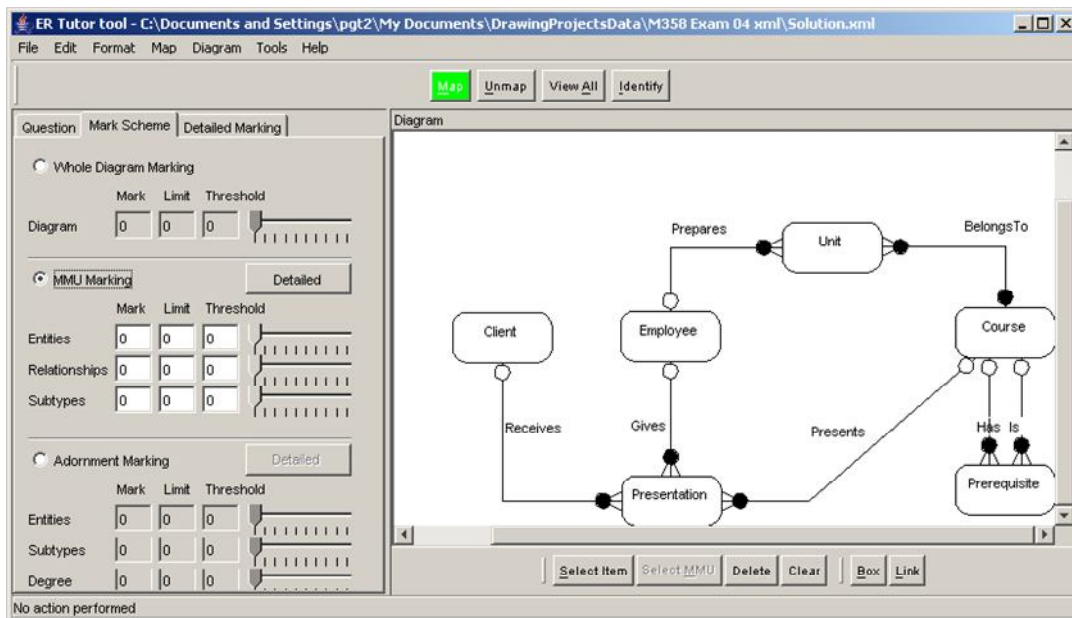


Figure 12. Tutor tool showing marking scheme setting pane.

5.2 Supporting tutors developing formative assessment materials

The student revision tool works with four distinct items of data: a question (also known as a scenario), an ER diagram, a set of mappings between the question and diagram on which feedback is based, and a marking scheme. To support a tutor in developing a question and its mappings with a diagram we have produced a Tutor Tool. The Tutor Tool enables the teacher to describe the diagram in terms of the scenario and to identify mappings which form the basis of the teaching in the Diagram Exerciser. The Tutor Tool is being used to create a new set of questions for the new database course (Figures 1 and 5 show the question pane of the tutor tool on the left of the screenshot, and Figure 12 shows the marking scheme setting features.)

The Student Revision Tool 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 working independently of the course team. The feedback from this testing will be incorporated into a final version of the tool which will be subjected to our rigorous quality assurance processes.

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 Revision tool, an area of wider interest [8].

6. Acknowledgements

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] 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.
- [2] Tsintsifas A., (2002), A Framework for the Computer Based Assessment of Diagram-Based Coursework' *Ph.D. Thesis*, Computer Science Department, University of Nottingham, UK.
- [3] Smith, N, Thomas, P.G. and Waugh, K. (2004) Interpreting Imprecise Diagrams. *Diagrams 2004 - Third International Conference on the Theory and Application of Diagrams*, University of Cambridge, UK.
- [4] Harris, Z.S. (1962) *String Analysis of Sentence Structure*, Mouton, The Hague.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] Burstein, J., Leacock, C., Swartz, R. (2001). Automated Evaluation of Essays and Short Answers. *Fifth International Computer Assisted Assessment Conference*, Loughborough University, UK.