Indicator Visualisation for Adaptive Exploratory Learning Environments

Sergio Gutierrez-Santos, Manolis Mavrikis, Alex Poulovassilis, Zheng Zhu
London Knowledge Lab, University of London, UK
{sergut,m.mavrikis,ap,zheng}@dcs.bbk.ac.uk

ABSTRACT

This paper presents our approach to identifying areas of improvement in the intelligent components of adaptive Exploratory Learning Environments. Students' interaction data from an online operational database are first transformed into a data warehouse in order to allow visualisation and exploration using online analytical processing (OLAP) tools. Using a microworld for secondary school algebra as a case study, we also present some more targeted visualisations of the students' interaction data. We demonstrate the possibilities that these visualisations provide for exploratory data analysis, enabling confirmation or contradiction of expectations that pedagogical experts may have about the system and ultimately providing both empirical evidence and insights for its further development.

Keywords

exploratory learning environments, indicators, visualisation

1. INTRODUCTION

In recent years there has been much research and development work focusing on open-ended interactive educational applications that encourage students' experimentation within a domain. These applications range from simple games to complex simulators and microworlds [1]. Although not new, they are becoming more common due to the new forms of interaction afforded by tablets and increasing ease of creation through related authoring tools. In parallel, the appreciation that in order for students to benefit from interaction with such Exploratory Learning Environments (ELEs) there is a need for explicit pedagogical support [2] has led to the development of adaptive support components [1].

The design and improvement of such adaptive exploratory environments is not a trivial task. Following a principled, evidence-based approach needs to rely on data gathered from students' interactions, which can help educationalists to understand how students are interacting with the system and technical experts to prioritise the development of enhanced or new support features. However, log files from ELEs contain large quantities of data that render their interpretation for researchers, teachers and systems designers quite a difficult and expensive task (cf. [3]). In addition, one does not always know in advance what data are required for analytical purposes and therefore an exploratory analysis may be needed. Lastly, logging of students' interaction data typically takes place in a manner that is optimal for recording and supporting students' interaction but not necessarily for subsequent analysis and decision-making.

In this paper, our case study is the MiGen system, which provides an intelligent environment to support 11-14 year old students' learning of algebra concepts In MiGen, students undertake tasks in a microworld called eXpresser. These tasks ask students to create models consisting of 2-dimensional tiled and coloured patterns — firstly specific instances of such models and then generalised versions in which one or more of the numbers in their construction are replaced by so-called "unlocked" numbers (i.e. variables). In parallel, students are asked to create rules specifying the number of tiles of each colour that are needed to fully colour their models (for more details see www.migen.org).

As students are interacting with the system, MiGen's intelligent support component [1] applies rule-based and case-based reasoning techniques to infer the occurrence of a wide range of significant task-independent and task-dependent indicators from the students' actions. These inferrences are used to provide both unsolicited and on-demand feedback to students. In addition, the indicators are stored in the operational online MiGen database, leading to large volumes of such data.

The question we address in this paper is: how might this data be visualised and explored in order to determine the effectiveness of the intelligent support provided by the system and to improve it? We have investigated several possible visualisations, including (i) multi-dimensional data visualisation and exploration using online analytical processing (OLAP) tools, and (ii) more targeted visualisations of the frequency of occurrence of different types of indicators and the transitions between them.

2. INDICATOR VISUALISATION

We first transformed data from the online MiGen database into a data warehouse that categorizes indicator occurrences according to several dimensions (e.g. when in occurred, the student and task it relates to, what kind of indicator it is). Multi-dimensional visualisation of this warehouse data using standard OLAP tools allowed the MiGen team and other experts to see what kinds of positive, neutral and negative behaviours are occurring as students are undertaking a task.

2.1 Frequency of indicator type occurrences

The visualisation in Figure 1 illustrates the conditional relative frequencies of different types of indicators (indicators of Status -1, 0, 1, 2) in three successive classroom sessions (Sessions 1, 2, 3). The widths of the bars correspond to the relative frequencies of indicator occurrences between the sessions. We can see that the number of indicator occurrences grows with each successive session and that the frequency of occurrence of negative indicators is decreasing with each successive session. This may be because students are becoming more familiar with using the system — a hypothesis that could warrant further investigation.

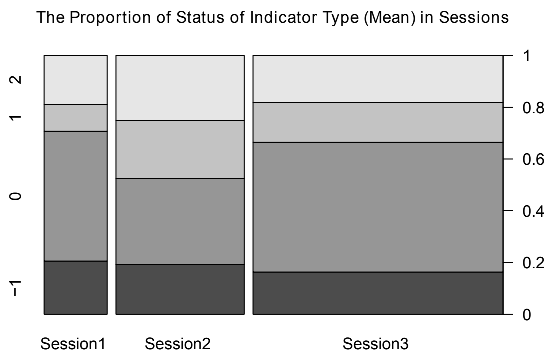

Figure 1: Per session proportion of negative (-1), neutral (0), positive (1) or feedback (2) indicators.

2.2 Transition of indicator type occurrences

Sequences of indicator types may be presenting patterns that can provide insight. Standard sequence analysis, however, provides patterns that are difficult to inspect. In order to facilitate the involvement of domain experts, we therefore investigated transition matrices, which are used to describe the transitions of a Markov chain.

Given a finite space of indicator types, $P_{ij} = P(j|i)$ is the probability of moving from indicator i to indicator j in one time step. Transition matrices can be normalised to quantify the transition probability from indicator i to any other indicator. We can also normalise the matrix to measure the incoming transition probability to indicator j from other indicators. In addition, we add artificial points to the system to capture the start and end of the interactions. Accordingly, for each model, s indicates the first indicator before the student begins construction of the model and e the last indicator at the end of the model's construction.

Transition matrices can be visualised using graphs such as those in Figure 2. Indicators shown with a circle round them indicate that there are transitions in the data where this indicator occurs in succession. The thickness of each line or circle indicates the value of the transition probability: the thicker the line, the higher the probability. The red (light grey) lines are associated with a probability less than 0.2 and the black lines a probability greater than or equal to 0.2.

Figure 2 shows an example transition matrix from Session 1 that leads to interesting insights. For example, consider the transition from indicator 3002, corresponding to numerical answer being provided by the student, to indicator 6001, corresponding to an intervention being generated by the system. In the visualisation of Session 2 (not shown here) there

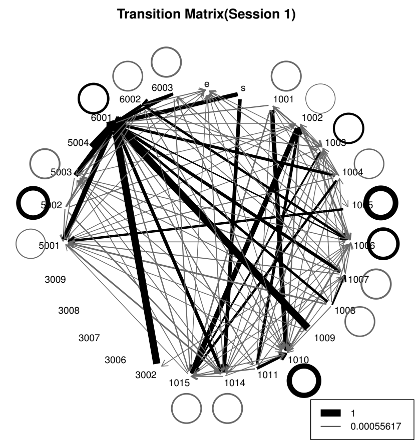

Figure 2: Incoming Transition Matrix (Session 1)

is no occurrence of this transition. This indicates that feedback received from the system in Session 1 was carried over to students' interactions in Session 2. Such an observation helps us raise a hypothesis for more detailed analysis or subsequent experimentation (e.g. "are students internalising the system's feedback and thus avoiding the same error in subsequent sessions or is this simply an artifact of their increasing familiarity with the system?").

3. CONCLUSIONS

We have developed several visualisations of learners' interaction data from an exploratory environment. We have discussed some insights derived from these and how they can inform decisions with respect to further research and design of the intelligent support provided by the system. Currently, our visualisations require the support of a technical expert in order to create them, using either standard OLAP tools or ad-hoc visualisations (mostly generated using R scripts). We plan to improve both their interactivity and their ease of use, in order to allow stakeholders with less technical expertise to be able to create such visualisations for themselves, to explore the data from their perspective, and to derive hypotheses worth further investigation.

Acknowledgements. The MiGen project (migen.org) and the data collection was funded from the ESRC-EPSRC TEL Programme. The work here has been co-funded by the FP7 EU projects iTalk2Learn (#318051 italk2learn.eu) and M C Squared (#610467 mc2-project.eu)

4. REFERENCES

[1] S. Gutierrez-Santos, M. Mavrikis, and G. D. Magoulas. A Separation of Concerns for Engineering Intelligent Support for Exploratory Learning Environments. *Journal of Research and Practice in Information Technology*, 44, 2012.

[2] R. E. Mayer. Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, 59(1):14–19, 2004.

[3] C. Romero, P. Gonzalez, S. Ventura, M. Del-Jesus, and F. Herrera. Evolutionary Algorithms for subgroup