Visualisation and Analysis of Students' Interaction Data in Exploratory Learning Environments

M.Mavrikis, Z.Zhu, S.Gutierrez-Santos, A.Poulovassilis
London Knowledge Lab, University of London

WebET 2015, May 2015

londonknowledgelab

Outline of the talk

1. Motivation
2. Multi-dimensional visualisation
3. More targeted visualisations
   - Frequency of indicator types
   - Transitions between indicator types
4. Conclusions and future work

1. Motivation

- Much research and development work focusses on open-ended interactive educational applications that encourage students' experimentation within a knowledge domain
- For students to benefit from interaction with such *Exploratory Learning Environments* (ELEs) there is a need for explicit pedagogical support to be provided to students
- Has led to research and development of intelligent techniques for providing adaptive support to students in order to foster their productive interaction with ELEs
- Data gathered from students' interactions with the ELE can help educationalists understand how students are interacting with the system, and technical experts to develop enhanced or new intelligent support features

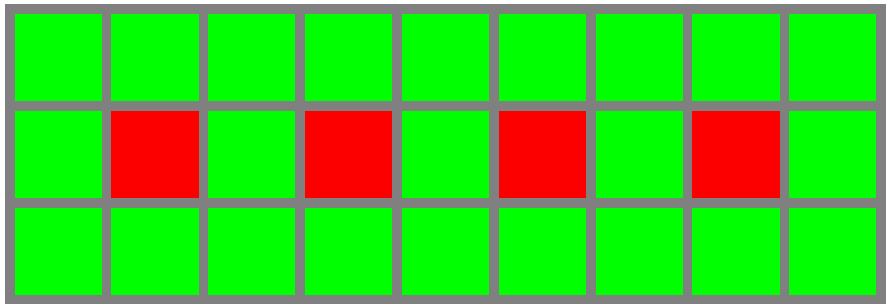
Motivation

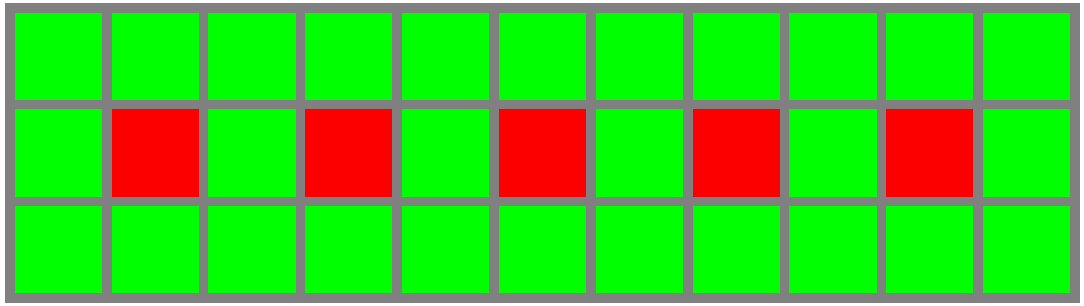
- Log files from ELEs can contain large quantities of data, making their interpretation for a difficult and costly task for researchers, teachers and systems designers
- Also, it is not possible to always know in advance what data are relevant for analytical purposes and therefore an exploratory analysis of the data gathered may be needed
- In this paper we describe transformation of students' interaction data from an operational online database into a data warehouse, thereby enabling data visualisation and exploration based on online analytical processing and other visual tools
- The aim is to increase the speed of data retrieval and analysis, and allow stakeholders to explore such data, enabling decision-making by pedagogical and technical researchers

2. Multi-dimensional visualisation

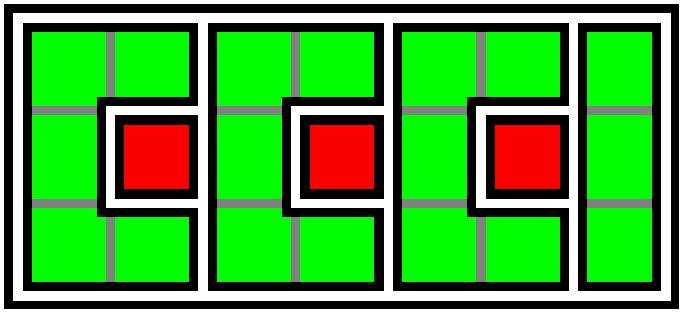
- Our case study is the *MiGen* system (www.migen.org), an intelligent ELE that fosters 11-14 year old students' learning of algebraic generalisation
- In MiGen, students under take construction tasks in a microworld called *eXpresser*
- These tasks ask students to create “generalised models” consisting of 2-dimensional tiled, coloured patterns constructed from one or more building blocks
- In parallel, students are asked to formulate algebraic rules specifying the number of tiles of each colour that are needed to fully colour their models

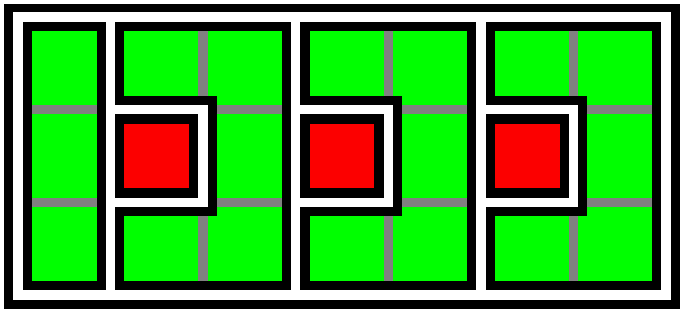
Generalised models

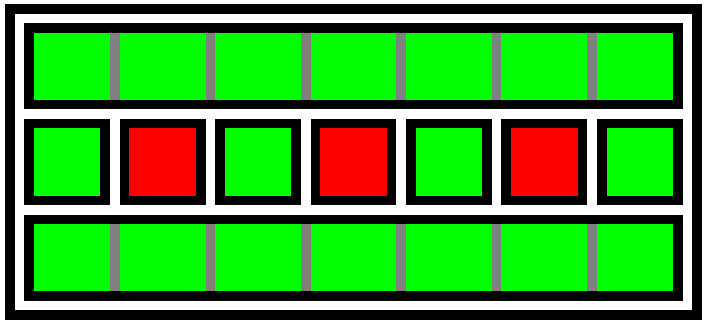


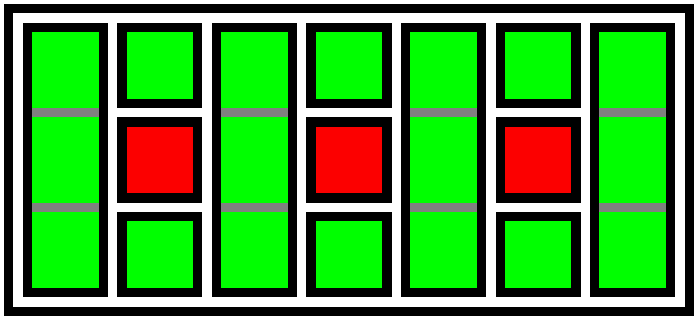


Possible construction approaches

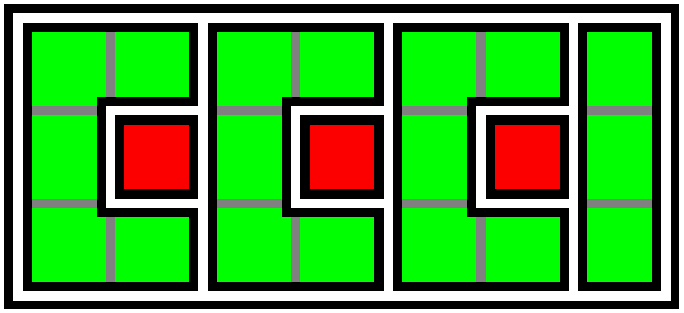




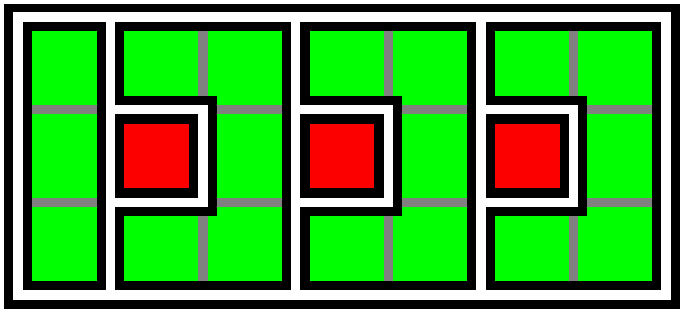




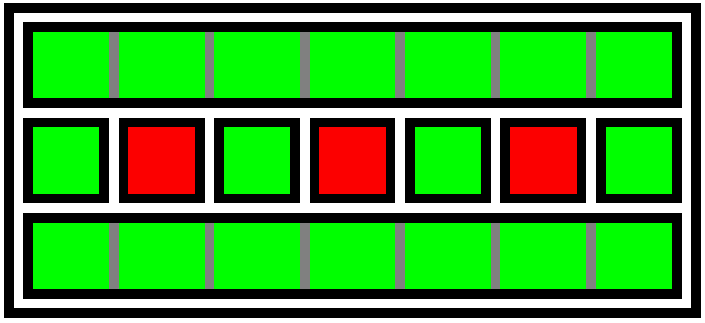
Possible rules



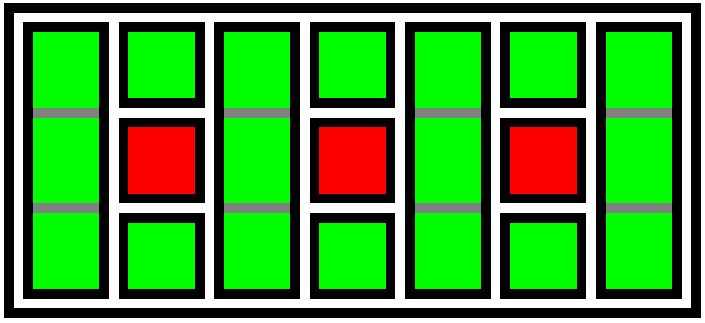
5 x reds + 3



3 + 5 x reds

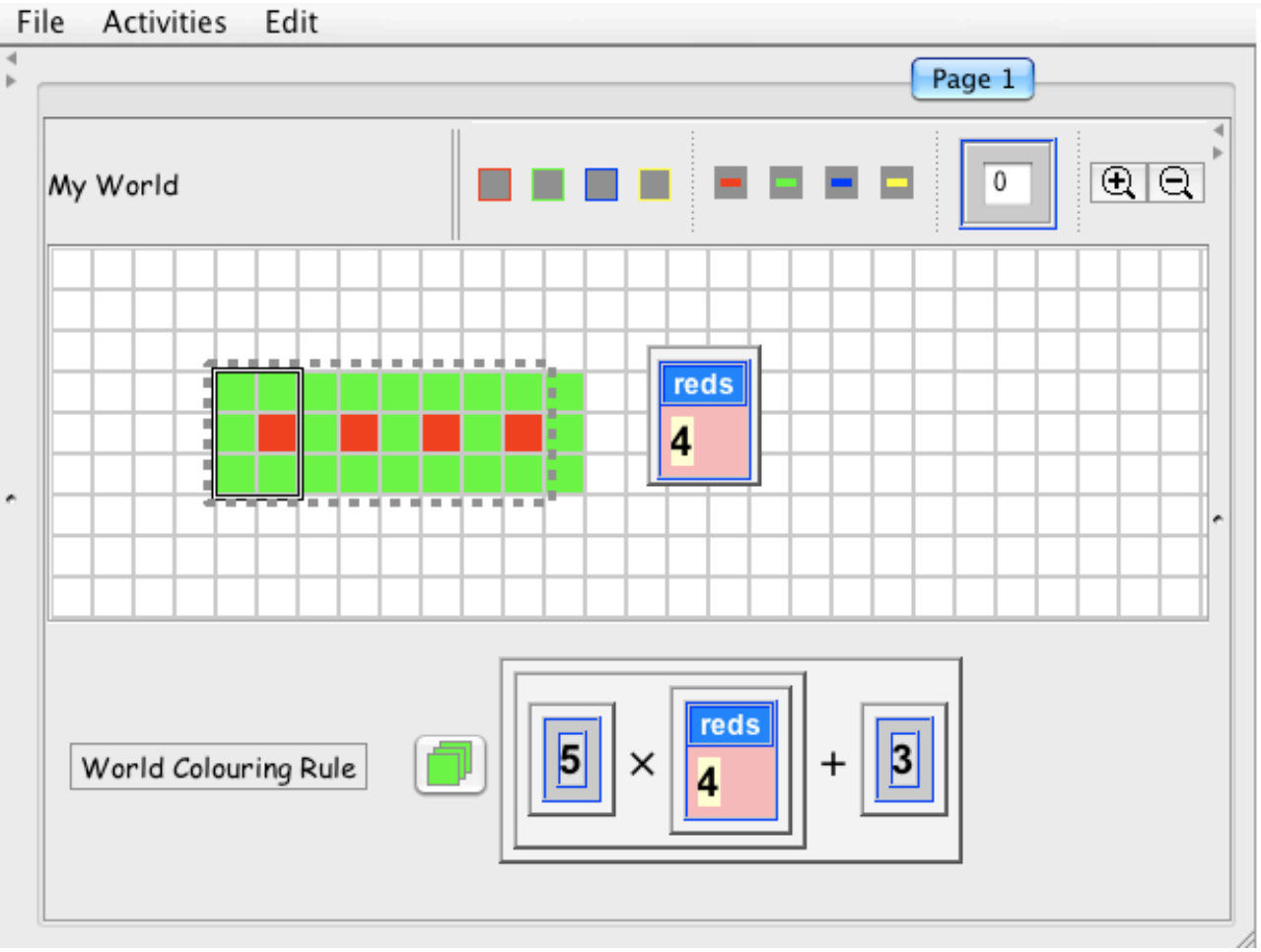


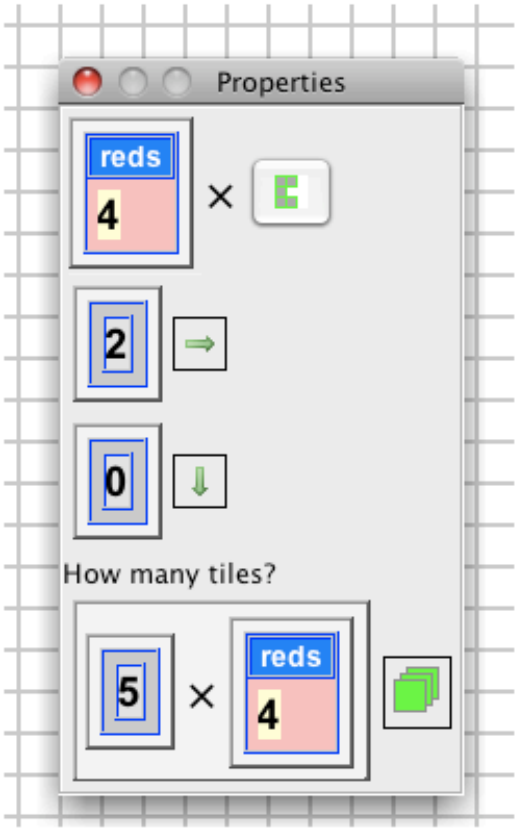
2 x (2 x reds + 1) + reds + 1



3 x (reds + 1) + 2 x reds

eXpresser microworld

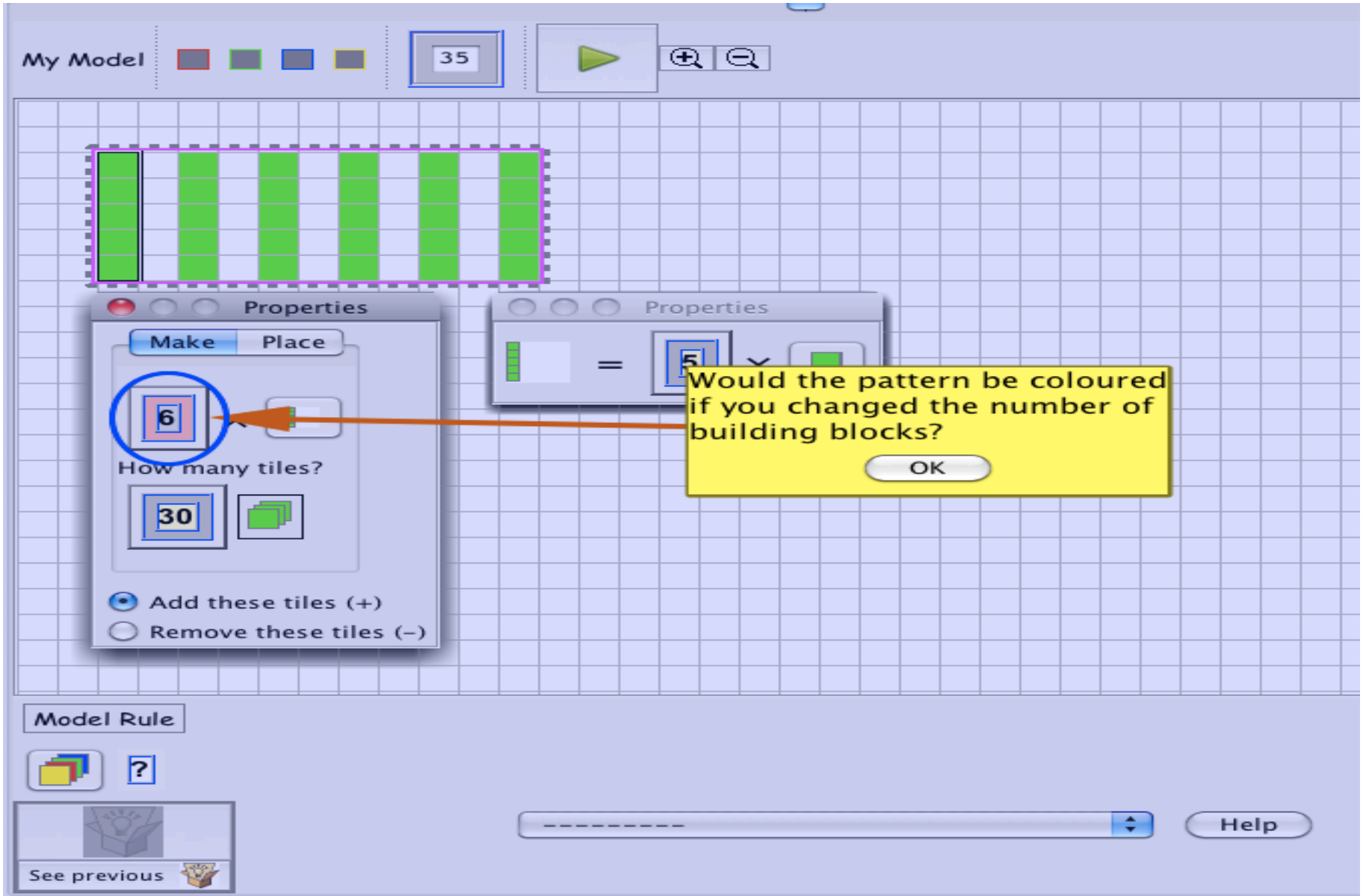




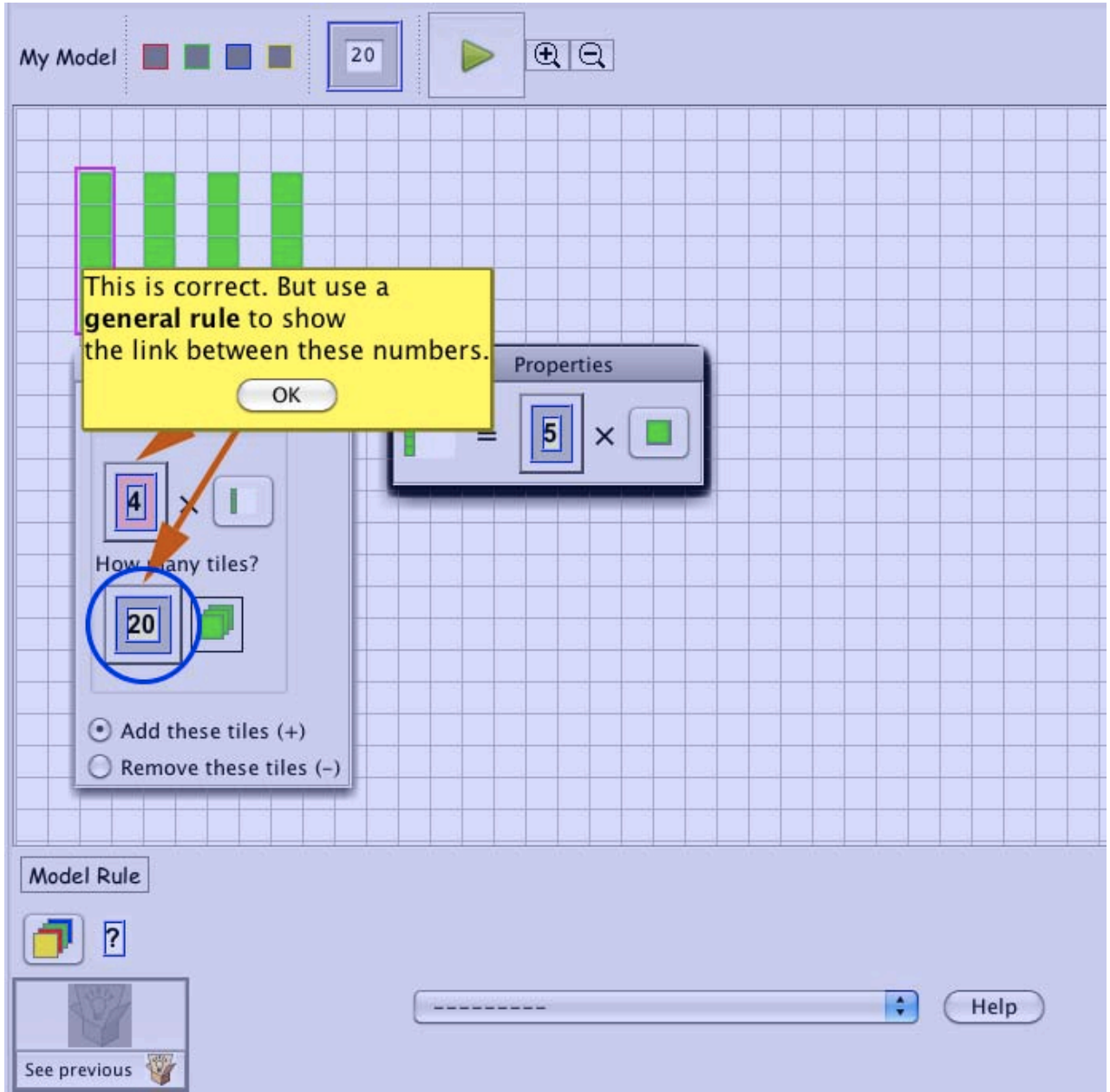
Intelligent Support

- An intelligent component called the *eGeneraliser* provides both unsolicited and on demand personalised feedback to students
- This support relies on the detection of *task-independent* (TI) and *task-dependent* (TD) indicators as students are interacting with the eXpresser
- Examples of TI indicators: `student has placed a tile on the canvas', `student has made a building block', `student has unlocked a number'
- TD indicators are detected by the analysis and reasoning sub-components of the eGeneraliser based on students' actions and on knowledge specific to the current task, e.g. `student has made a plausible building block for this task', `student has unlocked too many numbers for this task', `student has achieved task goal n'.
- All the occurrences of TI/TD indicators are logged in the operational online MiGen database

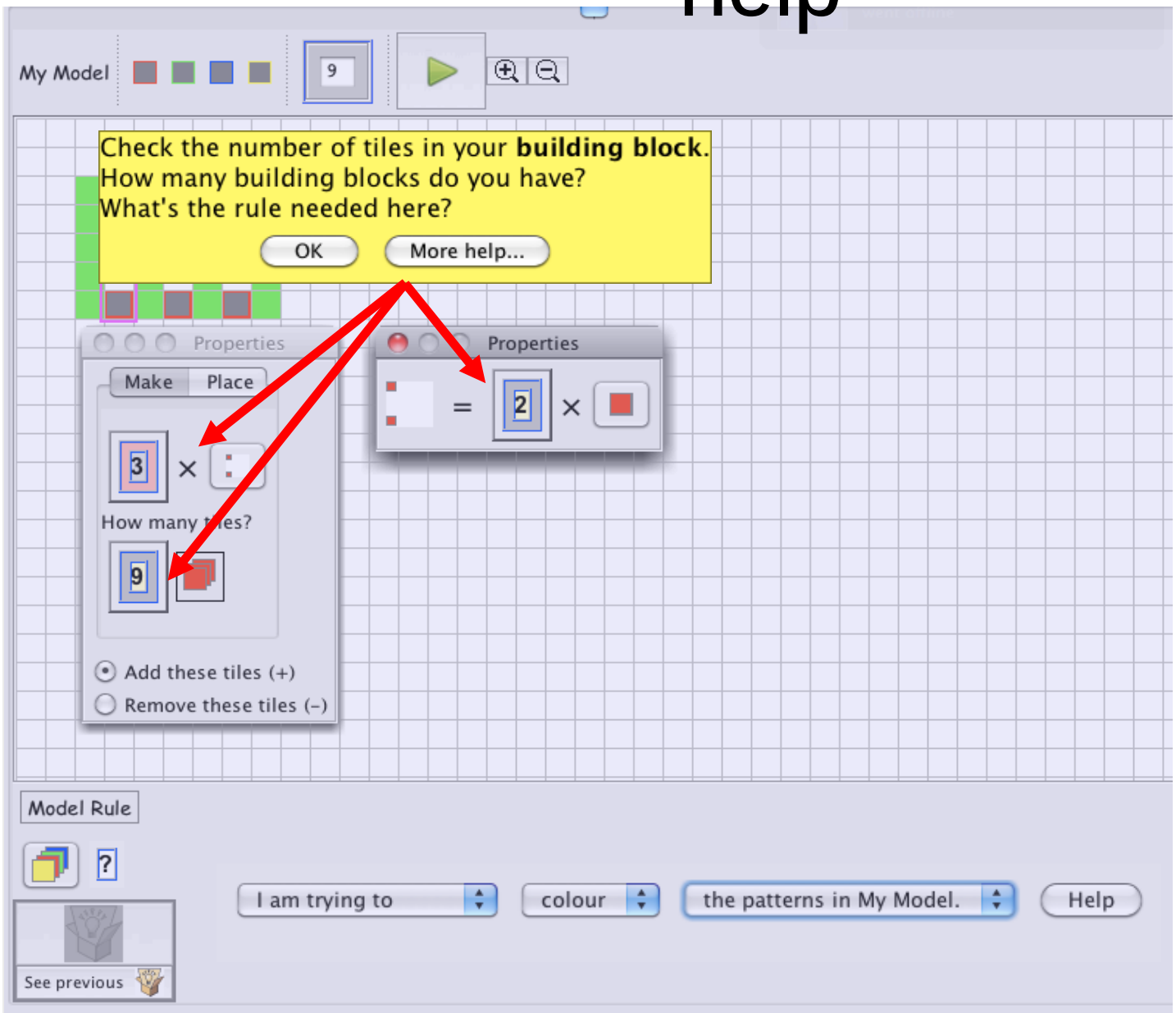
student feedback: a nudge...



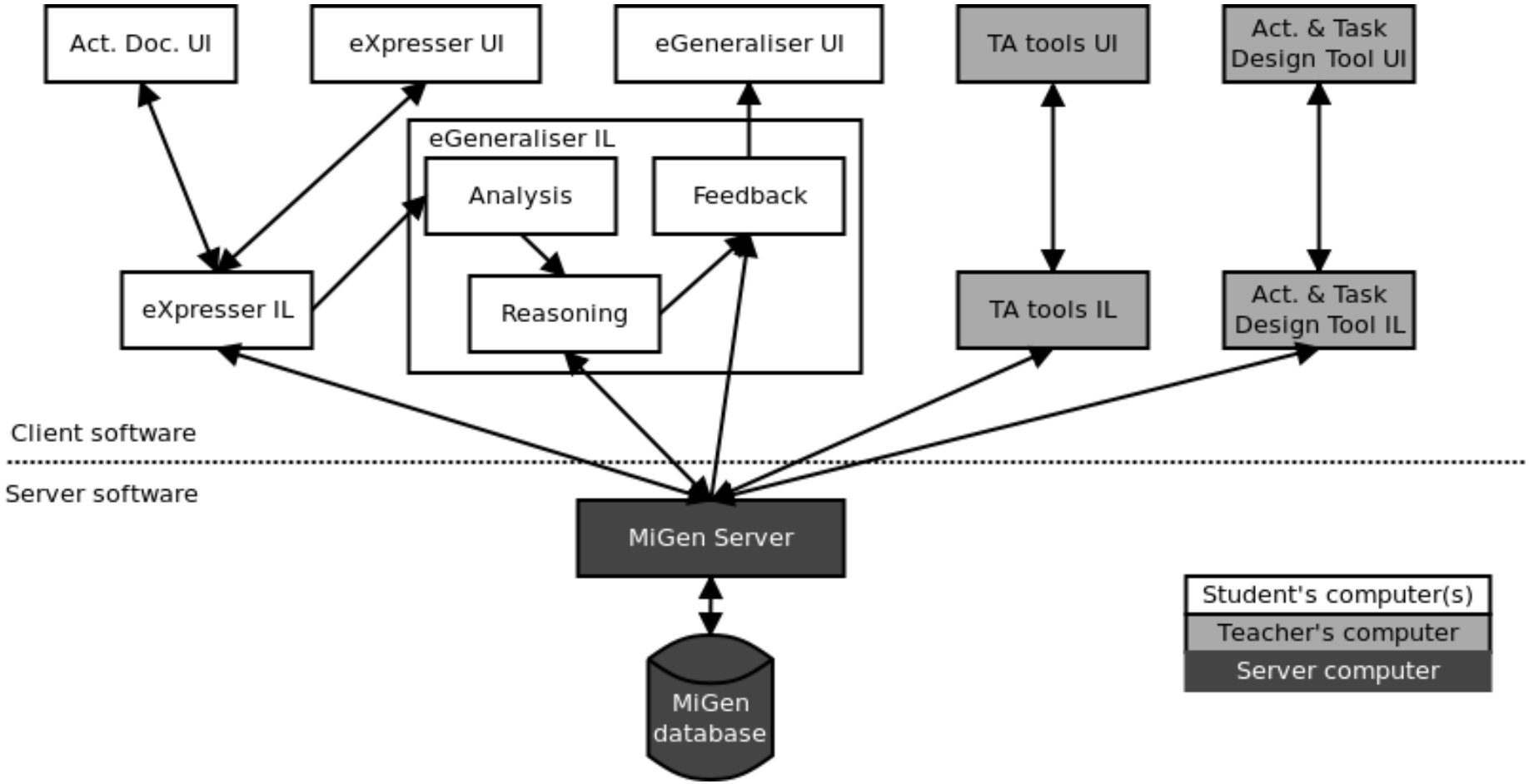
an unsolicited prompt...



feedback after a student's request for help



MiGen Architecture



The data being gathered

- Event-based data:
  - Log of students' interactions with the ELE
  - Occurrence of key indicators as students interact with the ELE
  - Provision of feedback by the ELE to students
- Students' constructions:
  - models
  - expressions
  - history of development of models and expressions

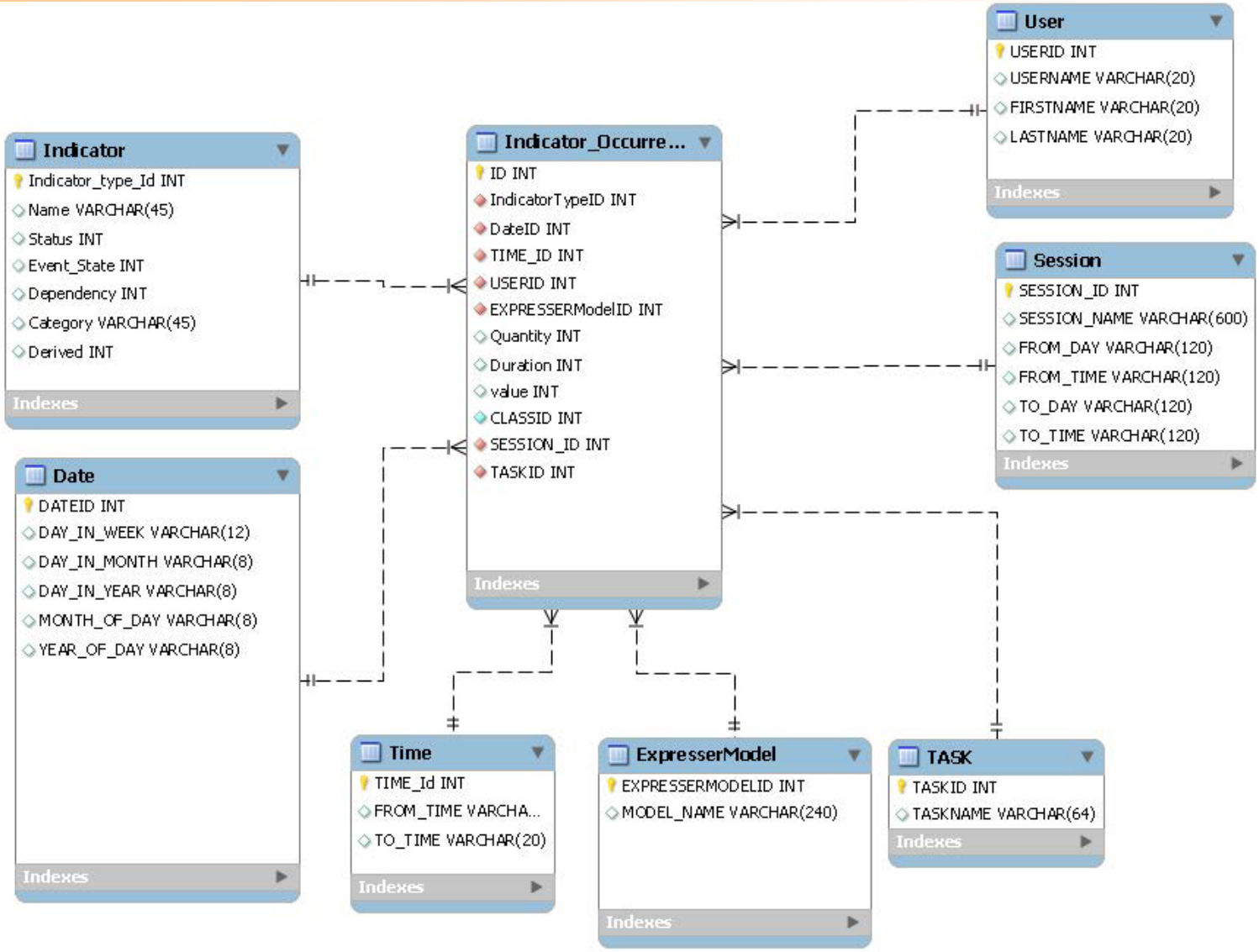
The data (cont'd)

- Task information
  - description
  - task learning goals
  - possible solution approaches
- Students' learner models:
  - task short-term model
  - task long-term model
  - domain model

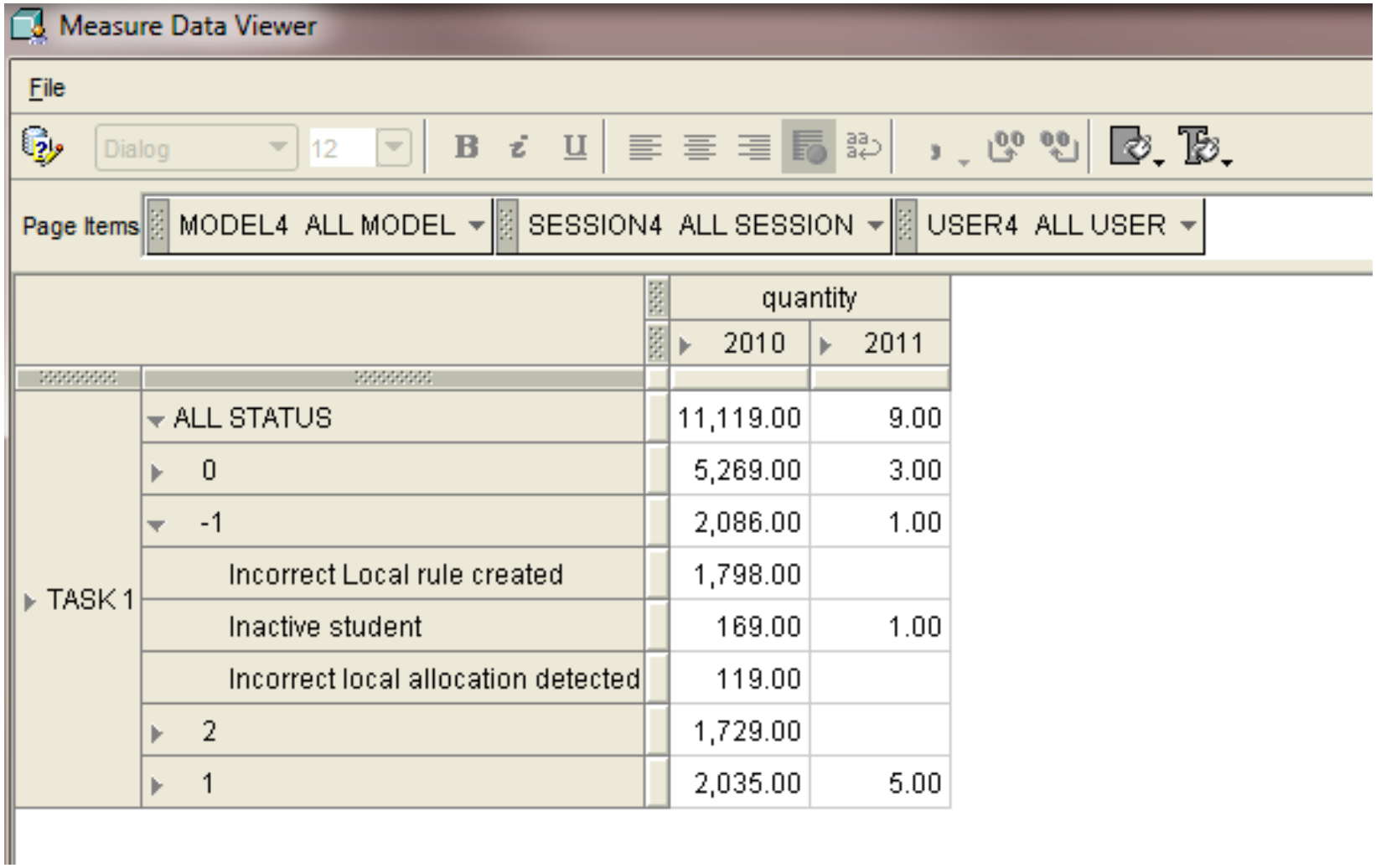
Data Warehouse and OLAP

- We extracted and transformed data from the operational MiGen database into a *data warehouse* – this Extraction-Transformation-Load (ETL) process has the aim of transforming day-to-day operational data from its original format into a form that is more efficiently and easily analysed by domain experts
- In the case of MiGen's interaction data, we identified
  - two numeric *measures* - the quantity and the duration of the indicator occurrences; and
  - seven *dimensions* relating to each occurrence of an indicator – the Date and Time it occurred, the Expresser Model and Task it relates to, the User who generated it, the Session in which it was generated, and its Indicator type

DW Schema (Oracle)



OLAP Views (Oracle Analytic Workspace Manager)



		quantity	
		2010	2011
TASK 1	ALL STATUS	11,119.00	9.00
	0	5,269.00	3.00
	-1	2,086.00	1.00
	Incorrect Local rule created	1,798.00	
	Inactive student	169.00	1.00
	Incorrect local allocation detected	119.00	
	2	1,729.00	
	1	2,035.00	5.00

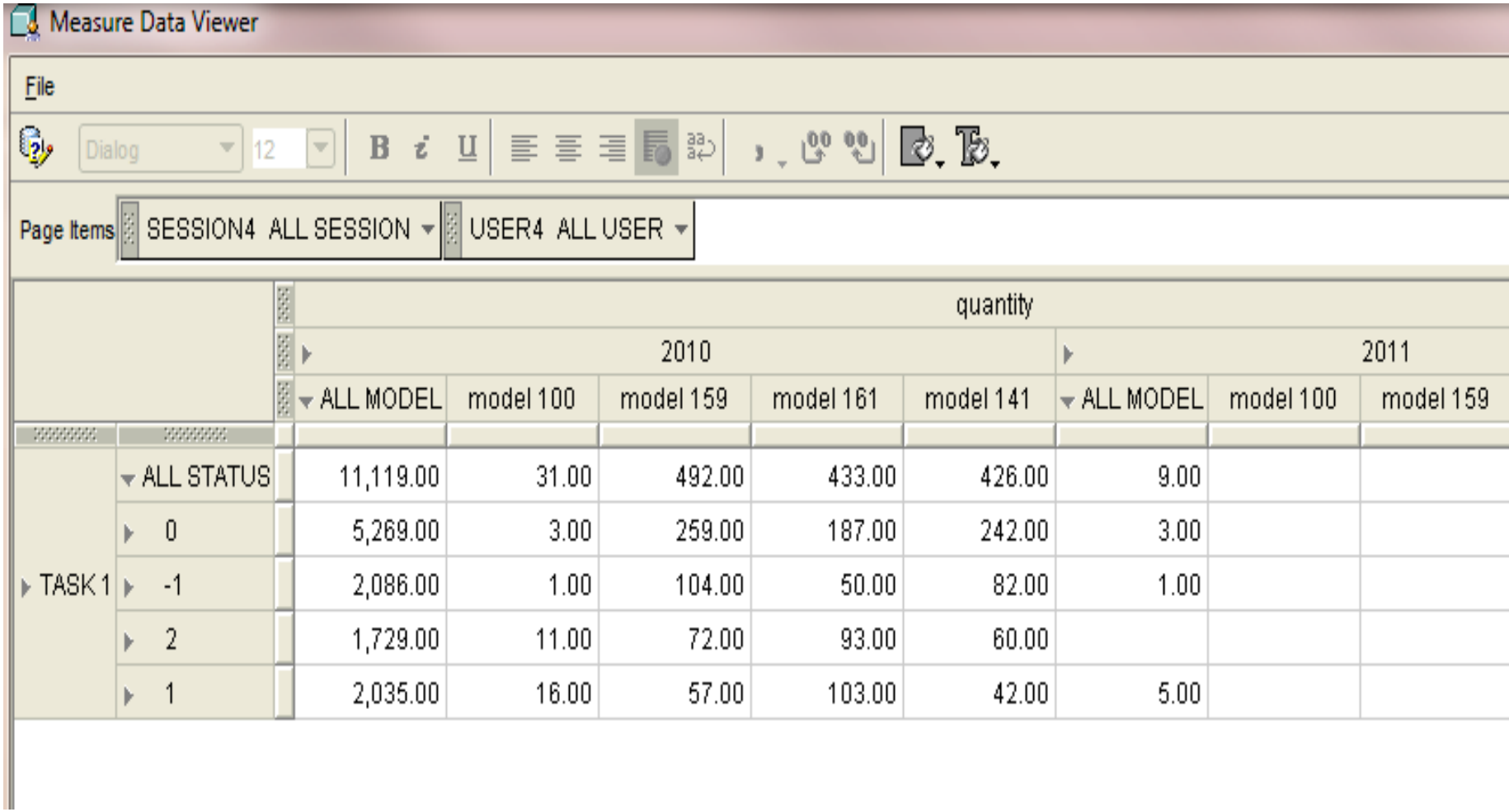
OLAP Views

- The previous figure shows a breakdown of the number of indicators occurrences by Task ID, Status and Year.
- There has also been some filtering of the data, so that data relating only to years 2010 and 2011 and Task 1 has been selected to be viewed. There has also been a ‘drilling down’ into the data relating to Status -1, to break it down into the specific indicator types that have this Status value (namely, Incorrect Local rule created, Inactive student, Incorrect local allocation detected).
- This kind of view allows us to see what kinds of positive, neutral and negative behaviours are commonly occurring as students are undertaking a task, in order to gauge students’ levels of engagement and productive interaction with a task.

OLAP Views

- Closing up again the Status -1 row and breaking down the data into the Models to which the indicators relate, we obtain the view shown in the next figure.
- This allows us to see the relative numbers of positive, neutral and negative indicators occurring during the construction of each model, allowing us to gauge, for example, the relative degrees of difficulty of different construction exercises.

OLAP Views



		quantity							
		2010					2011		
		ALL MODEL	model 100	model 159	model 161	model 141	ALL MODEL	model 100	model 159
	ALL STATUS	11,119.00	31.00	492.00	433.00	426.00	9.00		
	0	5,269.00	3.00	259.00	187.00	242.00	3.00		
TASK 1	-1	2,086.00	1.00	104.00	50.00	82.00	1.00		
	2	1,729.00	11.00	72.00	93.00	60.00			
	1	2,035.00	16.00	57.00	103.00	42.00	5.00		

3. More targeted visualisations

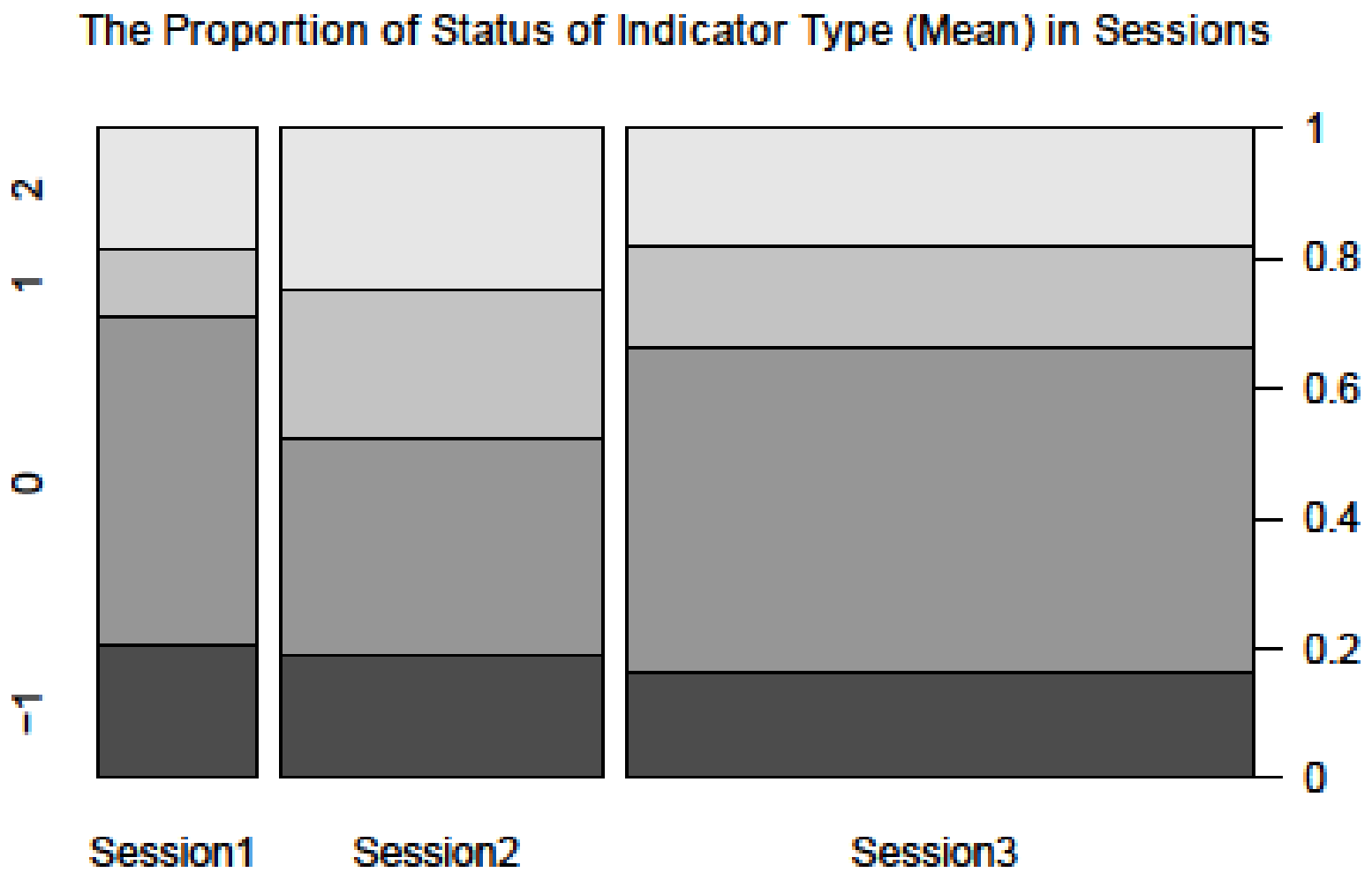
- The previous multidimensional visualisations allowed MiGen team members to interact with each other and with other pedagogical experts, the aim of gauging students' levels of engagement and productive interaction with the tasks set

- Beyond these standard visualisations provided by OLAP tools, we also developed several other ad-hoc visualisations to help us to further explore students' interactions

Frequency of indicator types

- For analysing frequency of occurrence of different indicator types we found visualisations such as the one shown in the next figure very useful.
- The heights of the bars correspond to the conditional relative frequencies of indicator types with Status -1, 0, 1, 2 in each of the classroom sessions being analysed (the analysis here relates to three successive classroom sessions: Session 1, Session 2, Session 3). The widths of the bars correspond to the relative frequencies of indicator occurrences between the sessions.
- In this example, we can see that the number of indicator occurrences grows with each successive session and that the frequency of occurrence of negative indicator types is decreasing with each successive session. This may be because students are becoming more familiar with using the system or more proficient with tasks set – hypotheses that would warrant further investigation.

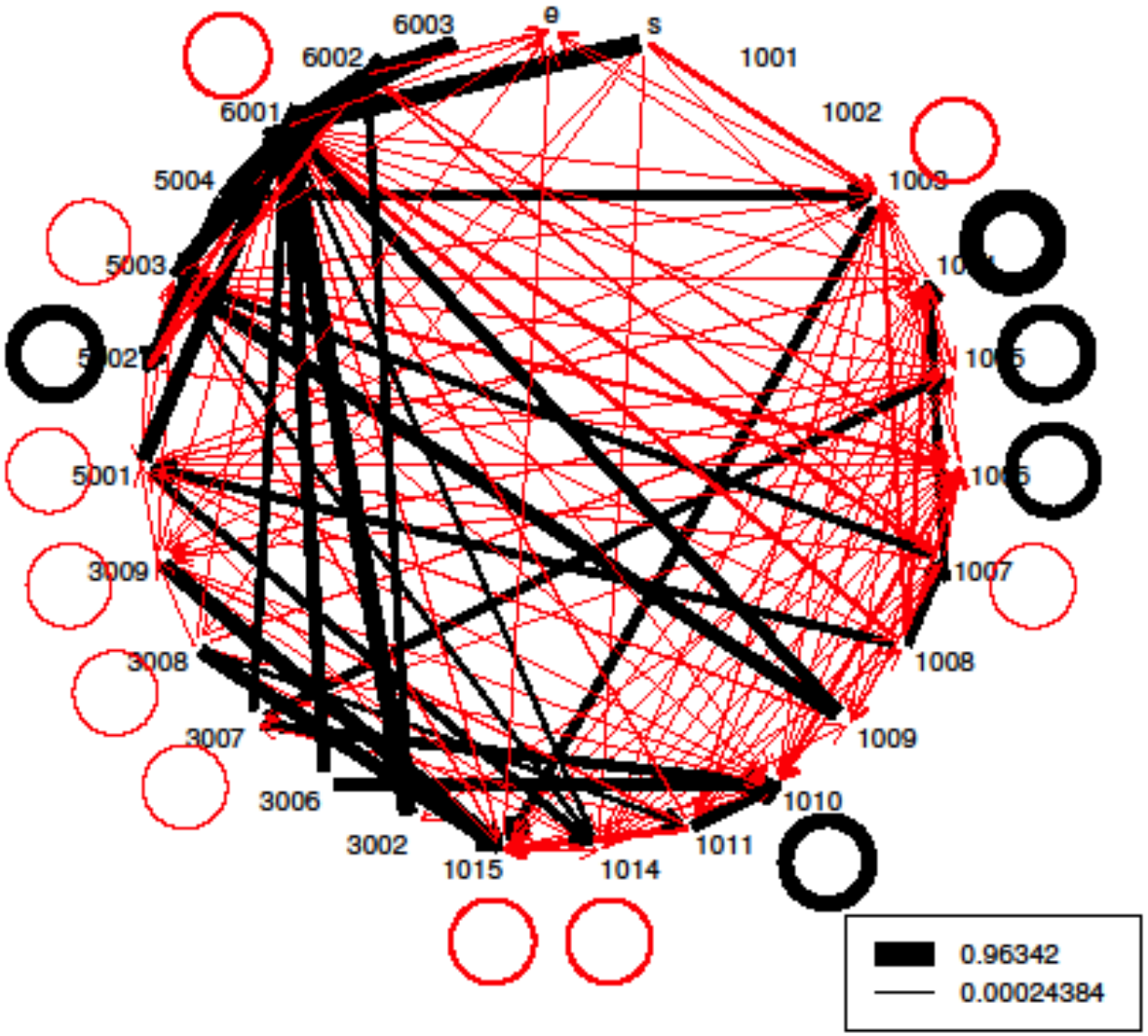
Frequency of indicator types



Transitions between indicator types

- Transition matrices can be visualised using graphs such as those shown in the next figure, which shows the normalised incoming transitions for a one-hour classroom session involving 22 students using MiGen
- We observe a black arrow 3007 -> 1005, indicating transitions from events of type 3007 (detection by the system that the student has made an implausible building block for this task) to events of type 1005 (modification of a rule by the student).
- Such an observation raises a hypothesis for more detailed analysis or further student observation, namely: does the construction of an incorrect building block lead students to self-correct their rules?
- Developing a better understanding of such complex interaction can lead to improvement of the system. For this particular example, we designed a new prompt that suggests to students to first consider the building block against the given task before proceeding unnecessarily in correcting their rules.

Transitions between indicator types



4. Conclusions and Future Work

- We have described the transformation of students' interaction data arising from ELEs from operational databases into a data warehouse, thereby enabling data visualisation and exploration based on online analytical processing and other visual tools
- We have presented several visualisations of the interaction data, have discussed insights derived from these, and how they can inform decisions with respect to further research and development
- Although developed in the context of the MiGen ELE, our approaches can be applied to any ELE in which key indicators relating to students' interactions are detected or inferred by the system
- In ongoing work we are investigating *graph-based* modelling of students' interaction data arising from ELEs, and development of meaningful graph-based queries, analyses and visualisations
- Our overarching aim is to investigate student-system interactions, leading to improvement of the intelligent support provided by the ELEs under investigation