Identifying strategies in user's exploratory learning behaviour for mathematical generalisation

Mihaela COCEA [a,1], and George D. MAGOULAS [a]

[a] *London Knowledge Lab, Birkbeck College, 23-29 Emerald Street, WC1N 3QS, London, UK*

Abstract. The nature of the activities that take place in Exploratory Learning Environments allow generating a variety of learner trajectories and makes difficult to develop a model of all possible behaviours (correct or incorrect). To alleviate this situation, we propose an approach for knowledge representation and identification of strategies followed by the learners during exploratory learning. Our approach combines heterogeneous sources of information and defines appropriate similarity measures for strategy identification. Some scenarios from the domain of mathematical generalisation are presented.

Keywords. Exploratory learning, strategy identification, similarity metrics

Introduction

Exploratory Learning Environments (ELEs) [3] are characterised by freedom given to the learner and are usually suitable for domains where multiple solutions could be obtained, and where the exploration is important for understanding the characteristics of the domain. Some ELEs are simulation-based, the users being able to change the different parameters of the models they explore; others allow the learner to construct their own models. Among the later is a system called *eXpresser* [2] [5], which is developed for teaching mathematical generalisation in classrooms.

In ELEs, as there are rarely unique solutions or approaches to a task, from pedagogical point of view, it is useful to know what strategies learners adopt when solving a task. For example, this could be useful for generating personalised feedback in terms of learner's approach to that particular task rather than guiding the learner to a predefined solution that may not have anything in common with the learner's current thinking.

In this paper we present the knowledge representation and identification mechanisms used to recognise learners' strategies in solving a task, and illustrate their operation through scenarios. Our approach employs knowledge representation and inferencing mechanisms where partial and complete solution are represented as sequences of cases, associated with temporal and dependency relations; the identification mechanisms are based on similarity measures.

[1]Corresponding Author: Mihaela Cocea, London Knowledge Lab, Birkbeck College, 23-29 Emerald Street, WC1N 3QS, London, UK; E-mail: mihaela@dcs.bbk.ac.uk.

[2]Developed in the context of MiGen Project, funded by ESRC, UK, under TLRP e-Learning Phase-II (RES-139-25-0381); `http://www.migen.org`.

1. Strategies Representation and Identification

In our approach, strategies in building constructions that take the form of different shapes, i.e. rectangles, C-shapes, etc., are represented as a series of cases [4] with certain relations between them. A *case* is defined as $C_i = \{F_i, RA_i, RC_i\}$, where C_i represents the case and F_i is a set of attributes. RA_i is a set of relations between attributes and RC_i is a set of relations between C_i and other cases, respectively.

The set of *attributes* includes two types: numeric and variables. The *set of relations between attributes* comprises *value* and *dependency* relations; the *set of relations between cases* refers to temporal relations. A *strategy* is defined as $S_u = \{N_u(C), N_u(RA), N_u(RC)\}$, $u = \overline{1, r}$, where $N_u(C_i)$ is a set of cases, $N_u(RA_i)$ is a set of relations between attributes and $N_u(RC_i)$ is a set of relations between cases.

Strategy identification is based on scoring elements of the strategy followed by the learner according to the similarity of their attributes and their relations to strategies previously stored. Thus, to identify components of a strategy, four similarity measures are defined: (a) Numeric attributes - Euclidean distance: $D_{IR} = \sqrt{\sum_{j=v+1}^{w} \times (\alpha_{I_j} - \alpha_{R_j})^2}$ (I stands for the pattern the learner is constructing and R stands for patterns compared or recalled from the ones stored); (b) Variables: $V_{IR} = \sum_{j=1}^{v} g(\alpha_{I_j}, \alpha_{R_j})/v$, where g is defined as: $g(\alpha_{I_j}, \alpha_{R_j}) = 1$ if $\alpha_{I_j} = \alpha_{R_j}$ and $g(\alpha_{I_j}, \alpha_{R_j}) = 0$ if $\alpha_{I_j} \neq \alpha_{R_j}$. (c) Relations between attributes - Jaccard's coefficient: $A_{IR} = \frac{|RA_I \cap RA_R|}{|RA_I \cup RA_R|}$. A_{IR} is the number of relations between attributes that patterns I and R have in common divided by the total number of relations between attributes of the two cases; (d) Relations between cases - Jaccard's coefficient: $B_{IR} = \frac{|RC_I \cap RC_R|}{|RC_I \cup RC_R|}$.

To identify the closest strategy to the one followed by a learner during construction, cumulative similarity measures are used for each of the four similarity types: (a) Numeric attributes - as this metric has a reversed meaning compared to the other ones, i.e. a smaller number means a greater similarity, the following function is used to bring it to the same meaning as the other three similarity measures, i.e. a greater number means greater similarity: $F_1 = z/\sum_{i=1}^{z} D_{I_iR_i}$ if $\sum_{i=1}^{z} D_{I_iR_i} \neq 0$ and $F_1 = z$ if $\sum_{i=1}^{z} D_{I_iR_i} = 0$, where z represents the minimum number of cases among the two compared strategies; (b) Variables: $F_2 = (\sum_{i=1}^{z} V_{I_iR_i})/z$; (c) Relations between attributes: $F_3 = (\sum_{i=1}^{z} A_{I_iR_i})/z$; (d) Relations between cases: $F_4 = (\sum_{i=1}^{z} B_{I_iR_i})/z$. As the four similarity metrics have different ranges, normalisation is applied to have a common measurement scale (details can be found in [1]). The similarity between the current strategy and a stored strategy is defined as the sum of the four measures after they are normalised.

2. Scenario-based Validation

Scenario-based validation was used to assess our approach in the context of an ELE for teaching mathematical generalisation. The scenarios of Table 1 cover some of the most complex situations encountered in the trials with pupils. We use as an example here a task called "footpath", typical in the UK curriculum, which requires finding the number of tiles that surround a pattern of red tiles with gaps in between them. There are severals strategies for constructing the surrounding for that pattern (see Figure 1).

The identification mechanism was validated using scenarios for the following situations: (a) detection of partial constructions; (b) detection of intermediate partial constructions; (c) detection of combination of strategies and (d) detection of specific and

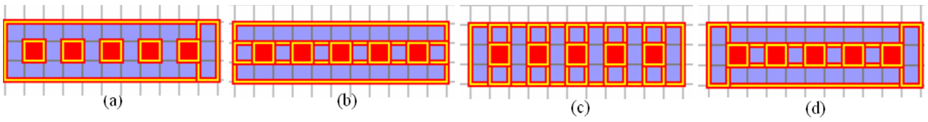

Figure 1. Strategies for footpath task: (a) forward C; (b) HParallel (horizontally parallel); (c) VParallel (vertically parallel); (e) H&VParallel (horizontally and vertically parallel).

general constructions. The constructions for these scenarios are illustrated in Figure 2; their pedagogical rational and the similarity measures for the two most similar strategies are given in Table 1.

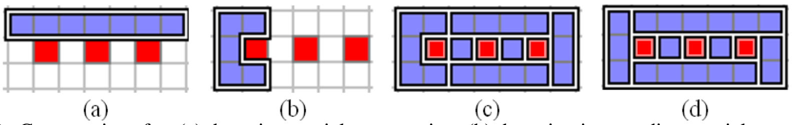

Figure 2. Constructions for: (a) detecting partial construction; (b) detecting intermediate partial construction; (c) mixed strategies; (d) lack of symmetry.

Table 1. Scenarios summary.

Scenario	Pedagogical rational	Construction	Top matching strategies
Detecting partial constructions	Guiding before end of construction	Figure 2a	HParallel: 2.83
			H&VParallel: 1.99
		Figure 2b	Partial forward C: 3
Working with the specific/general	Identify best approach	Figure 1b	HParallel specific: 3.88
			HParallel general: 3.78
Mixed strategies	Guide the learner towards one strategy	Figure 2c	H&VParallel: 2.12
			Forward C: 1.98
	Symmetry as a generalisation principle	Figure 2d	HParallel: 2.22
			H&VParallel: 1.96

3. Conclusion

In this paper an approach for strategy identification in the domain of mathematical generalisation was presented. Some details of knowledge representation and strategy identification were provided together with some examples of pedagogically-driven scenarios.

Acknowledgements

This work is partially funded by the ESRC/EPSRC Teaching and Learning Research Programme (Technology Enhanced Learning), award RES-139-25-0381.

References

[1] M. Cocea, G. Magoulas: Hybrid Model for Learner Modelling and Feedback Prioritisation in Exploratory Learning, to be published in *The International Journal of Hybrid Intelligent Systems* (2009).

[2] J. Kaput: Technology and Mathematics education. In D. Grouws (ed.) *Handbook of Research on Mathematics Teaching and Learning*, New York: Macmillan (1992) 515-556.

[3] T. de Jong and W.R. van Joolingen: Scientific discovery learning with computer simulations of conceptual domains, *Review of Educational Research*, **68** (1998), 179-202.

[4] J.L. Kolodner: *Case-Based Reasoning*, Morgan Kaufmann Publishers, Inc., 2nd edn. (1993).

[5] D. Pearce, E. Geraniou, M. Mavrikis, S. Gutierrez-Santos, K. Kahn: Using Pattern Construction and Analysis in an Exploratory Learning Environment for Understanding Mathematical Generalisation: The Potential for Intelligent Support. In S. Gutierrez-Santos, M. Mavrikis (eds.), *Proceedings of the 1st International Workshop on Intelligent Support for Exploratory Environments*, EC-TEL'08 (2008).