Theory-based Learning Analytics: Using Formal Concept Analysis for Intelligent Student Modelling

Michael. D. Kickmeier-Rust[1], Michael Bedek[1], Dietrich Albert[1]
[1]Knowledge Technologies Institute, Graz University of Technology, Graz, Austria

Abstract - Learning Analytics is one of the most important fields for driving (educational) artificial intelligence. In this paper we briefly introduce one of Europe's key research initiatives to make a step from primarily statistics and data mining driven ways to do Learning Analytics, towards more theory-grounded, psycho-pedagogically inspired analyses. One approach we suggest is Formal Concept Analysis, which is a strong theoretical backbone in technology-enhanced learning settings For educators, FCA may provide AI-based analyses of 'big data' sets and intuitive representations for a number of pedagogical questions concerning the performance of students on the individual- as well as the class-level.

Keywords: Formal Concept Analysis, Learning Analytics, Visualizations, Learner Modelling

1 Introduction

Learning Analytics (LA) can be considered a best practice in education and it is a key factor for making education more personalized, adaptive, and effective. This, undoubtedly, needs a strong educational AI. Analyzing a variety of available data to uncover learning processes, strengths and weaknesses, competence gaps clearly is a prerequisite for any formative guidance, meaning to assess students in order to provide them with the best possible and most individual support and not just to "rate" them, for changing and adjusting educational measures and teaching, and not least for disclosing and negotiating learner models [1, 2]. On the basis of available data, ideally large scale data sets, smart tools and systems are being developed to provide teachers with effective, intuitive, and easy to understand aggregations of data and the related visualizations. There is a substantial amount of work going on this particular field; visualization techniques and dashboards are broadly available (cf. [3, 4]), ranging from simple meter/gauge-based techniques (e.g., in form of traffic lights, smiley, or bar charts) to more sophisticated activity and network illustrations (e.g., radar charts or hyperbolic network trees). However, LA operates in a delicate and complex area. On the one hand, facing today's classroom realities, we often find environments without a comprehensive availability of technical infrastructure. This, in turn, does not allow or support an easy recording the necessary data, LA and educational data mining (EDM) is always talking about. Also, from a socio-pedagogical perspective, learning must be seen as a process of social interaction that not always occurs in front of some electronic device. Thus, LA must be based on fewer data. Perhaps incomplete data, and accordingly careful must be our approaches to data analyses. On the other hand, it is so very easy to visualize learning on a superficial level using perhaps the aforementioned traffic lights or bar charts. The added value to the teachers is likely of limited utility to them. To provide a deeper and more formatively-inspired insight into the learning history and the current state of a learner and to entire groups of learners (beyond the degree to which a teacher might know it intuitively) requires finding and presenting complex data aggregations. This, most often, bears the significant downside that it is hard to understand.

Challenges for LA and its visualizations, for example, are to illustrate learning progress (including learning paths) and - beyond the retrospective view - to display the next meaningful learning steps and topics.

In this paper we introduce the method of directed graphs, the so-called lattice diagrams, for structuring learning domains and for visualizing the progress of a learner through this domain. This type of visual analytics possibly provides a solid AI foundation for the field of education and LA: In a recent project named Lea's Box (www.leasbox.eu) we propose so-called Formal Concept Analysis (FCA) as a framework for addressing these ideas. We developed a web-based tool which provides a working environment with social and personal open tools to support students in developing their inquiry based learning skills. By this, the tool supports learners by enabling domain and open learner modelling. The fields of application of the FCA in general, and the FCA-tool in particular, have been extended in the course of the LEA´s BOX project (http://leas-box.eu/) which stands for Learning Analytics Toolbox. In the context of LEA´s BOX, the FCA-tool is mainly used by teachers for student modelling and visualization of educational data. By applying formal concept analysis on students´ performance data, a set of pedagogically relevant questions for teachers can be addressed and visualized.

2 Formal Concept Analysis (FCA)

FCA describes concepts and concept hierarchies in mathematical terms, based on the application of order and lattice theory [5]. Insofar, FCA is at the borders between AI research and educational technologies. The starting point is the definition of the formal context *K* which can be described as a triple (*G*, *M*, *I*) consisting of a set of objects *G*, a set of attributes *M* and a binary relation *I* between the objects and the attributes (i.e. "g *I* m" means "the object g has attribute m"). A formal context can be represented as a cross table, with objects in the rows, attributes in the columns and assigned relations as selected cells. An example of a formal context is shown in Fig. 1. This formal context has been created by the FCA-tools *Editor View*. Teachers use the *Editor View* to de-fine the formal context and to add learning resources (URLs or files) which can be assigned to both, to objects and to attributes, respectively. In order to create a concept lattice, for each subset $A \in G$ and $B \in M$, the following derivation operators need to be defined:

$A \mapsto A' := \{m \in M \mid g\, I\, m \text{ for all } g \in A\}$, which is the set of common attributes of the objects in *A,* and

$B \mapsto B' := \{g \in G \mid g\, I\, m \text{ for all } m \in B\}$, which is the set of objects which have all attributes of *B* in common.

A formal concept is a pair (*A*, *B*) which fulfils $A' = B$ and $B' = A$. The set of objects *A* is called the extension of the formal concept; it is the set of objects that encompass the formal concept. The set *B* is called the concept's intension, i.e. the set of attributes, which apply to all objects of the extension. The ordered set of all formal concepts is called the concept lattice $\mathfrak{B}(K)$ (see [6] for details), which can be represented as a labelled line diagram (see Fig. 2). The concept lattice shown in Fig. 2 has been created by the FCA-tool *Lattice View*. Every node represents a for-mal concept. The extension *A* of a particular formal concept is constituted by the objects that can be reached by descending paths from that node. As an example, the node with the label "Goldfish" has the extension {Goldfish, Tree frog}. The intension *B* is represented by all attributes that can be reached by an ascending path from that node. In the example above, the intension consists of {is able to swim, lives in / on the water}.

Fig. 1. *Editor View* for creating objects, attributes, and relations.

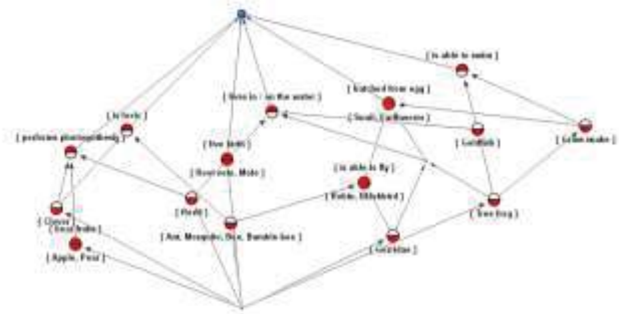

Fig. 2. Concept lattice.

3 Domains Open Learner Models

Once the teacher has created the formal context, students can explore the resulting concept lattice by engaging in interactive graph visualizations (see Fig. 2). By selecting a node, the corresponding concept´s extension and intension are illustrated in a highlighted manner. The concept lattice makes the structure of the knowledge do-main and the interrelations of its concepts explicit. Similar as for concept maps, this kind of graphic organizer aims to facilitate meaningful learning by activating prior knowledge and illustrating its relationship with new concepts [7].

In case the teacher also assigned learning resources to the objects and attributes in the FCA-tool's *Editor View* open learner modelling can be supported. Visualizations of open learner models (for an overview see [8]) are aiming to facilitate reflection on the side of the students and to support teachers to better understand strengths and weaknesses of their students. The FCA-tool's *Lattice View* (cf. Fig. 3) applies the often-used traffic-light analogy (see e.g. [9]) to show the student the extent to which he or she already consumed learning resources.

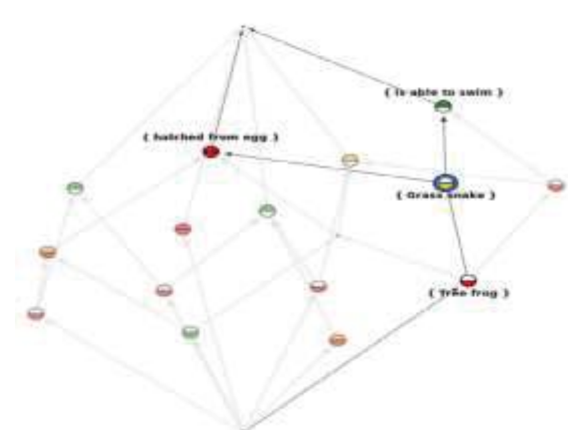

Fig. 3. FCA-tools *Lattice View* for visualizing domain- and learner models.

ISBN: 1-60132-438-3, CSREA Press ©

4 Application in School Settings

Similar as [0] who were the first who applied the FCA with students and their performance data we suggest for-mal contexts with student as “attributes” and problems or test-items as “objects”. The relation between these two sets means “student m has solved test item g”. An example of a concept lattice which results from such a formal context is shown in Fig. 4 (the data has been reported by [10]). As briefly outlined in the following sections, such a concept lattice visualizes answers to a set of pedagogical questions which are of high interest for teachers.

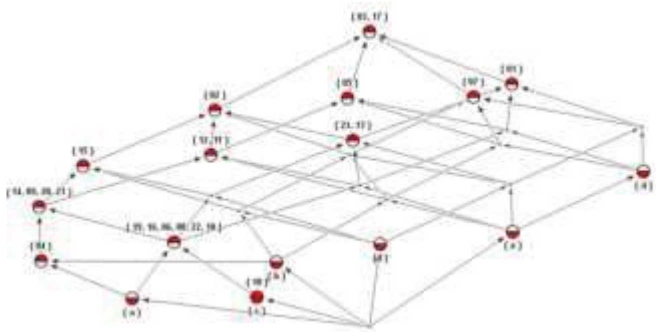

Fig. 4. Concept lattice with students as attributes (numbers from *01* to *23*) and test items (letters *a*, *b*, *c*, *d*, *e*, and *f*) as objects.

4.1 Extensions and Intensions

As mentioned above, the set of test items which have been solved by a particular student can be directly depicted from the extension of the formal concept with the students´ label assigned to it. As an example in Fig. 4, student *10* is the only one who solved only a single test item, *c*, and students *03* and *17* (assigned to the top element of the concept lattice) mastered all problems. When clicking on a particular node the formal concept´s extension and intension is highlighted. As an example shown in Fig. 5 (left side), the student *04* has successfully mastered the test items *a* and *b*.

The intension of a formal concept which has an object-label assigned to it indicates the set of students which have successfully mastered the according test item. As an example, the problem d in Fig. 5 (right side) has been solved by the students 01, 03, 05, 07 and 17. As it can be also seen, this formal concept located above the formal concept with the object-label e assigned to it. This means, that all students who solved item d were also able to solve item e.

4.2 Overlapping and Differences

The performance of two or more students can be compared when examining the intensions of the formal concepts with the according attribute-labels. As an example, the students *07* and *15* mastered different subsets of problems (see Fig. 4): Student

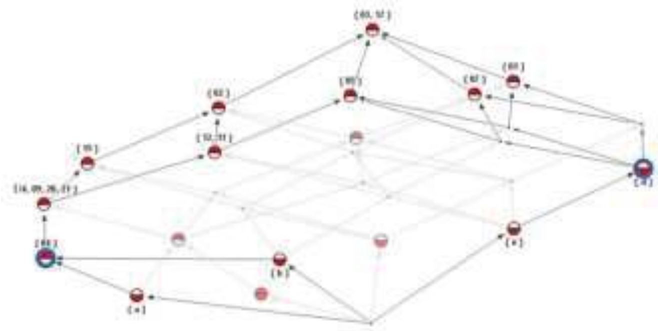

Fig. 5. The extension represents the set of test items solved by a student (see student *04*) and the intension represents the set of students who solved the particular test item (see test item *d*).

07 mastered the items *b*, *d*, *e* and *f* while student *15* mastered the items *a*, *b*, *c*, and *f*. Both students mastered items *b* and *f* (which is the set closure of their intensions) and together they mastered all problems (which is the set union of their intensions).

As a teacher, such kind of information might be of great interest since it helps to effectively arrange groups of students when aiming for collaborative, peer-learning (where students learn together in groups). In the example above, the students *07* and *15* together could be tutors for other students.

4.3 Learning Progress Over Time

The concept lattices shown in Fig. 4 and Fig. 5 are the result of a formal context which is an evaluation of the students´ performances at a certain point in time. However, in some cases it might be of great interest for a teacher to observe the learning progress over a longer period of time. Ideally, all students might be able to master all items at the end of course or the semester. In such a case, all cells in the formal context would be filled with crosses. This would result in a concept lattice with only a single formal concept. Fig. 6 exemplifies such a learning progress over time. The concept lattice in the middle results from adding one solved item to the students´ performance states (except for the students *03* and *17*). The concept lattice on the right results from adding an-other item to the student´s performance states. In general, the visual appearance of the concept lattice gives a first impression of the student’s coherence with respects to their performance: A concept lattice which looks “complex” due to a large amount of formal concepts is a clear indication for a high diversity among the students´ performances.

5 Conclusions

In this paper we proposed to apply FCA to support students and teachers. Students apply the FCA-tool in order to learn a certain knowledge domain by interacting with the concept lattice which allows uncovering hidden relationships between domain´s concepts. Insofar, FCA is a powerful technique

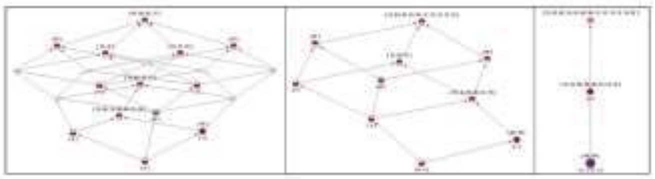

Fig. 6. Concept lattices changing over time reflect the learning progress of the class of students.

coming from AI and data mining research to support education in an intelligent way. In addition to that, a student´s reflection upon his or her learned and still-to-learn concepts is supported by an open learning modelling approach. Students can access the visualizations and compare themselves with peers. Teachers can apply the FCA tool to visualize the answers to a set of pedagogical questions which are of high interest for teachers.

We tried to formulate some of these pedagogical questions; they are the result of focus group studies and inter-views with teachers in the early phase of the LEA´s BOX project. The resulting visualizations as shown above are currently in the spotlight of formative, qualitative evaluation studies with small focused groups of teachers. Current work on the technical side of the project focuses on the development of interactive visualizations which can be easily used by teachers in the classroom. Early feedback of teachers concerns the complexity of the concept lattices, in particular when dealing with a great amount of problems (respectively competences and skills). Conceptual research and the elaboration of ideas on how to reduce this complexity without reducing the amount of information which can be extracted and deduced from the visualizations will be the main focus of our work in the near future.

Acknowledgements. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7) under grant agreement no. 619762 (LEA´s BOX). This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

6 References

[1] Dimitrova, V., McCalla, G. and Bull, S. (2007). Open Learner Models: Future Research Directions (Special Issue of IJAIED Part 2). International Journal of Artificial Intelligence in Education 17(3), 217-226.

[2] Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Buckingham Shum, S:, Ferguson, R., Duval, E., Verbert, K., and Baker, R.S..J.D. (2011). Open Learning Analytics: an integrated & modularized platform: Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques. Available online at http://solaresearch.org/OpenLearningAnalytics.pdf

[3] Ferguson, R., and Buckingham Shum, S. (2012). Social Learning Analytics: Five Approaches. In Proceedings of the 2nd International Conference on Learning Analytics & Knowledge, 29 Apr - 02 May 2012, Vancouver, British Columbia, Canada.

[4] Duval, E., (2011). Attention Please! Learning Analytics for Visualization and Re-commendation. In Proceedings of the 1st International Conference on Learning Analytics & Knowledge, 27 Feb – 1 March 2011, Banff, Alberta, Canada.

[5] Wille, R. (1982). Formal Concept Analysis as Mathematical Theory of Concepts and Concept Hierarchies. In: Ganter, B., Stumme, G., Wille, R. (eds) Formal Concept Analysis, pp 1-34, Berlin: Springer.

[6] Nesbit, J.C., Adesope, O.O. (2006). Learning With Concept and Knowledge Maps: A Meta-Analysis. Review of Educational Research. 76, 413–448.

[7] Bull, S., Kay, J.: Open Learner Models. Nkambou, R., Bordeau, J., Miziguchi, R. (eds.) (2011). Advances in Intelligent Tutoring Systems, pp. 318-388. Berlin: Springer.

[8] Arnold, K.E., Pistilly, M.D. (2012). Course Signals at Purdue: Using learning analytics to increase student success, In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 267-270. ACM, New York.

[9] Rusch, A., Wille, R. (1996). Knowledge spaces and formal concept analysis. In: Bock, H.H., Polasek, W. (eds.) Data analysis and information systems: Statistical and conceptual ap-proaches, pp. 427-436. Springer, Berlin (1996)

[10] Korossy, K. (1999). Modeling Knowledge as Competence and Performance. In: Albert, D., Lu-kas, J. (eds.) Knowledge Spaces: Theories, Empirical Research, and Applications, pp. 103-132. Springer, Mahwah, NJ.