Watanabe, T. and Seta, K. (Eds.) (2015). Proceedings of the 11th International Conference on Knowledge Management.

Applying Formal Concept Analysis to visualize classroom performance

Michael A. BEDEK[a*] & Dietrich ALBERT[a,b]

[a]*Graz University of Technology,* [b]*University of Graz; Austria*

*michael.bedek@tugraz.at

Abstract: Learning analytics and educational data mining are two interrelated research fields which recently gained enormous popularity and started to enter the classroom. Even if these techniques provide a wide range of opportunities, appropriate tools should meet the requirements of its users: Teachers are usually interested in user-friendly tools which enable them to reduce the time necessary for personalized assessment and tailored competence development of their students. The European research project LEA´s BOX aims to provide a Web platform for teachers for activity tracking, domain and student modelling as well as visualizing of educational data. To reach these aims, the project builds upon two powerful set-theoretic frameworks, the Knowledge Space Theory which is well-established in student modelling and the Formal Concept Analysis (FCA) which is usually applied for domain modelling. In this paper, we present new applications of the FCA for educational purposes. In particular, we focus on a set of pedagogically relevant questions addressing the performance of the whole class by intuitive visualizations based on the FCA.

Keywords: Formal Concept Analysis, Knowledge Space Theory, Learning Analytics, Visualizations, Learner Modeling

1. Introduction

Learning analytics and educational data mining are two highly interrelated research fields which became enormously popular in recent years (e.g. Steiner, Kickmeier-Rust & Albert, 2014). When applying learning analytics and educational data mining in schools, it is of high importance to meet the requirements of teachers and students. Teachers usually want to have user-friendly tools which help them to reduce the time required for personalized assessment and tailored competence development of their students.

The European research project LEA´s BOX (http://leas-box.eu/) stands for Learning Analytics Toolbox and aims to provide a Web platform for teachers and students which supports activity tracking, domain modelling, student modelling as well as visualization of educational data. In order to assess a student´s current knowledge and competence state to support personalized competence-centered learning, LEA´s BOX extends existing fields of application of two frameworks: the *Knowledge Space Theory* (KST) and the *Formal Concept Analysis* (FCA). These two frameworks are well established in the fields of student modelling (KST) and domain modelling (FCA) - based on order- and lattice theory. They serve as theoretical basis for structuring, analyzing and visualizing educational data.

The paper is structured as follows: First, a brief introduction on KST and the FCA will be given. It continues with the main part of the paper, which is the application of the FCA for visualizing the students´ performances to answer a set of pedagogical questions for teachers. We will conclude with a discussion on the current findings and an outlook on future activities.

2. Knowledge Space Theory

The KST (Doignon and Falmagne, 1985) suggests that every knowledge domain Q (e.g. descriptive statistics) can be characterized by a set of problems (items). A student´s knowledge state is the set of problems he or she is able to master. In many cases, it is reasonable to assume mutual dependencies, so-called prerequisite relations, between the problems of a given knowledge domain. For example, a student who successfully masters problem *y* (e.g. calculation of standard deviation) presumably masters problem *x* (e.g. calculation of means) too. In this case, problem *x* is a prerequisite of problem *y*.

A *knowledge space* is the ordered set of all reasonable knowledge states. Reasonable in this context means, that a knowledge state which includes a particular problem also includes the problem´s prerequisites (in the example above, all knowledge states which include problem *y* also include problem *x*). A knowledge space also includes the empty set (a student may not master any problems) as well as the set *Q*. For additional properties of knowledge spaces see Doignon and Falmagne (1999).

The KST has a 30-years tradition as powerful framework for learner modelling, adaptive testing and competence development in technology-enhanced learning (for an overview see Falmagne et al., 2013), and thus, the main focus of this paper is on the FCA which hasn't been extensively applied for such kind of purposes so far.

3. Formal Concept Analysis

The *F*ormal *C*oncept *A*nalysis (FCA) has been established in the early 80s by Wille and colleagues (Wille 1982, 2005). The FCA aims to describe a domain, i.e. concepts and concept hierarchies in mathematical terms. The starting point is the definition of the formal context. A formal context K is defined as a triple (*G*, *M*, *I*) with *G* as a set of objects (in German: "Gegenstände") and *M* as a set of attributes *M* (in German: "Merkmale"). The relation *I* (incidence-relation) assigns objects and attributes, i.e. *g I m* means the object g has the attribute m. The formal context K can be represented as a cross table, with the objects in the rows, the attributes in the columns and by crosses ("Xs") whenever *g I m* holds for a particular object and attribute (see table 1).

Table 1: Example of a formal context with objects and their attributes

Attributes *M* / Objects *G*	is toxic	is able to fly	is able to swim	hattched from egg
Bee	X	X		X
Bumble-bee		X		X
Tree frog			X	X
Grass snake	X		X	X

For each subset $A \subseteq G$ and $B \subseteq M$ the following derivation operators need to be defined:

$A \to A' := \{m \in M |\ gIm \text{ for all } g \in A\}$ which is the set of common attributes of the objects in *A*, and
$B \to B' := \{g \in G |\ gIm \text{ for all } m \in B\}$ which is the set of objects which have all attributes in *B*.

A formal concept is a pair (*A*, *B*) with the subsets $A \in G$ and $B \in M$ which fulfil $A = B'$ and $B' = A$. The set *A* is called the *extension* of the formal concept; it is the set of objects of the formal concept. The set *B* is called the *intension* of the formal concept; it is 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 Wille, 2005) which can be visualized by a labelled line diagram (see figure 1).

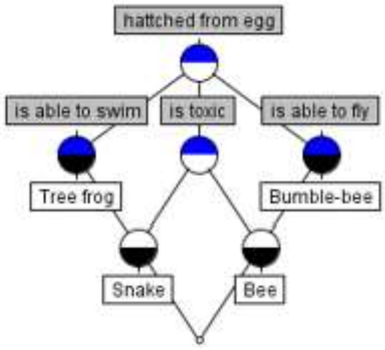

Figure 1. A concept lattice resulting from the formal context in table 1

Every node represents a formal concept. In order to avoid redundancy, all objects and attributes are labelled only once. A concept lattice can be "read" as follows: The extension *A* of a formal concept comprises all objects whose labels can be reached by descending paths. As an example, the node with the label "Tree frog" has the extension {Tree frog, Snake}. The intension *B* of a formal concept can be reached by all attributes whose labels can be reached by ascending paths from that node. In the case of the formal concept in the example above, the intension consists of the attributes {hatched from egg, is able to swim}.

4. Applying the FCA for learner modelling

Rusch and Wille (1996) were the first who applied the FCA with learners and their knowledge states to show the correspondence between the FCA and the KST. They proposed a knowledge context (*S*, *P*, *I)* with students *S*, problems *P* and an incidence-relation which assigns students to problems which they have *not* solved. This rather unintuitive incidence relation leads to formal concepts whose complements of the intensions are knowledge states.

However, for LEA´s BOX such kind of knowledge contexts or concept lattices are not applicable since it is not intuitive for teachers to think in terms of "complements of a formal concept´s intension". They are mainly interested in clear visualizations which directly indicate the set of problems which have been mastered by a student (or which they failed).

We suggest knowledge contexts with student as "attributes" and problems as "objects". An example of such a knowledge context is given in table 2 (the data has been reported by Korrossy, 1999). Such an alternate knowledge context overcomes the above mentioned shortcut since a student´s knowledge state can be directly derived from the according concept´s extension. In addition to that, as it will be outlined in the following sections, the resulting concept lattice allows visualizing answers to a set of pedagogical questions which might be of interest for teachers.

Table 2: A knowledge context with student as attributes and problems as objects (from Korossy, 1999)

Problem / Students	*01*	*02*	*03*	*04*	*05*	*06*	*07*	*08*	*09*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*
a	X	X	X	X	X	X		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X
b		X	X	X	X		X		X		X	X		X	X		X			X	X		
c	X	X	X		X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
d	X		X		X		X										X						
e	X	X	X		X		X				X	X	X				X						X
f	X	X	X				X						X		X		X						X

4.1. *Depicting knowledge states from formal concepts extensions*

The concept lattice which results from the knowledge context in table 2 is shown in figure 2. As briefly outlined above, the set of problems which have been solved by a particular learner can be directly depicted from the extension of the formal concept with the learners´ label assigned to it. As an example in figure 2 (left side), the student *04* has successfully mastered the problems *a* and *b*. Student *10* is the only one who solved only a single problem, *c*, and students *03* and *17* (assigned to the top element of the concept lattice) mastered all problems.

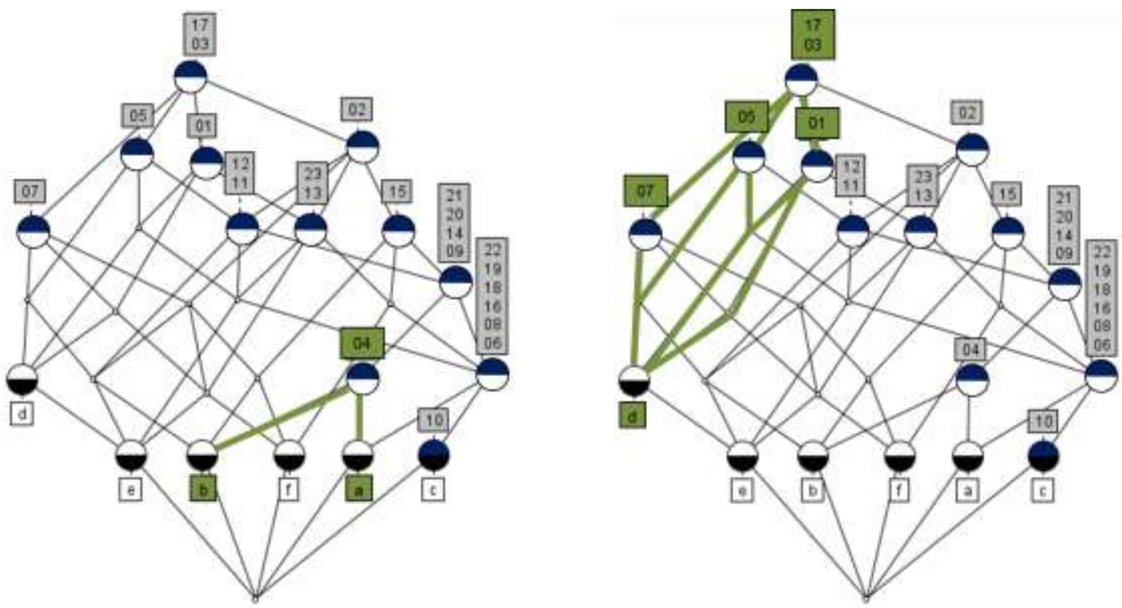

Figure 2. The extension of a formal concept is a knowledge state (left side) and the intensions of a formal concept with an problem-label is the set of students who solved that problem (right side)

4.2. *Depicting the set of students who solved items from formal concepts intensions*

The intension of a formal concept which has an problem-label assigned to it indicates the set of students which have succesfully mastered that problem. As an example, the problem *d* in figure 2 (right side) has been solved by the learners *01*, *03*, 05, *07* and *17*. As it can be also seen, this formal concept located above the formal concept with the problem-label *e* assigned to it. This means, that all students who solved problem *d* were also able to solve problem *e*, i.e problem *e* can be considered as prerequisite for problem *d*.

4.3. *Highlighting overlaps and differences of students performances*

The performances of two or more students can be compared when examining the intensions of the formal concepts with the according attribute-labels. As exemplified in figure 3, the students *07* and *15* mastered different subsets of problems. The knowledge state of student *07* encompasses the problems solved *b*, *d*, *e* and *f* while the knowledge state of student *15* encompasses the problems *a*, *b*, *c*, and *f*. Both students mastered problems *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.

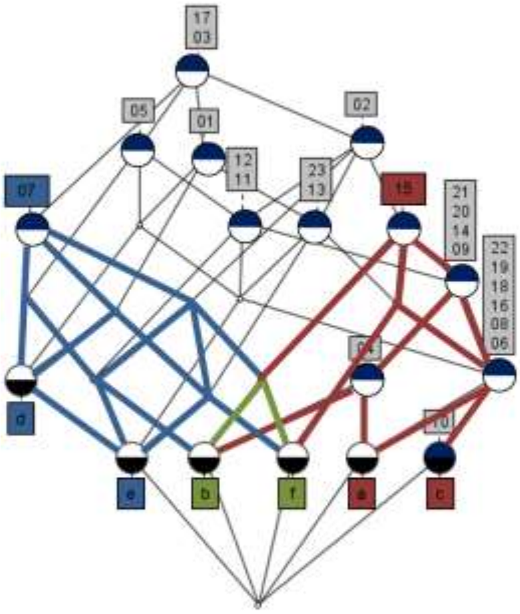

Figure 3. Comparing performance patterns between subsets of learners

4.4. Visualizing a classrooms´ learning progress over time

The concept lattice in figures 2 and 3 results from 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. In the perfect case, all students should finally end up (e.g. at the end of the semester) with the knowledge state *Q*. In such a case, all cells in the knowledge context would be filled with crosses. This would result in a concept lattice with only a single formal concept. Figure 4 exemplifies such an ideal learning progress. The concept lattice in the middle results from adding one solved item to the students´ knowledge states (except for the students *03* and *17*). The concept lattice on the right side results from adding another item to all knowledge states smaller than *Q*.

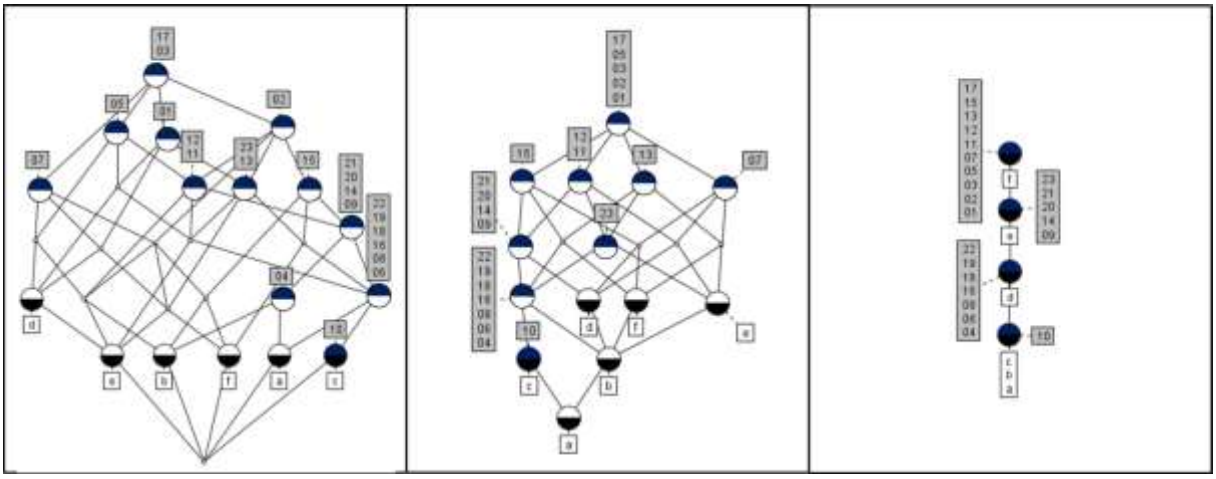

Figure 4. Changing concept lattices over time reflects learning progress of the whole class of students

Such a kind of interactive visualization (which could be manipulated for example with a slider) might be of particular interest when dealing with competences rather than on a rather behaviorist performance level (i.e. solved or failed problems; see for example competence-based extensions of the KST, e.g. Albert & Lukas, 1999; Heller, Steiner, Hockemeyer, & Albert, 2006). In general, the visual appearance of the concept lattice gives a first impression on the coherence among the students: A concept lattice which looks “complex” due to a large amount of formal concepts is an indication for a high diversity among the students´ performance- and competence states. On the other side, a concept lattice with a relatively small amount of formal concepts indicates that the students with respect to the knowledge or competence states are more coherent.

5. Discussion and Outlook

In the previous sections, we suggested to apply the FCA in a classroom to visualize the answers to a set of pedagogical questions which are of high interest for teachers. Even if the focus of this paper was on the FCA, however, for other pedagogical questions, such as *which learning trajectories were most common in the classroom*? or *which learning path should be taken given a certain knowledge or competence state?*, the KST and its competence-based extensions seem to be more adequate (see for example Heller, Steiner, Hockemeyer, & Albert, 2006; Nakamura et al., 2011). In LEA´s BOX it is foreseen to exploit the strengths of both frameworks.

The pedagogical questions described above are the result of small focus groups and interviews 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. The results of the above mention formative evaluation studies will feed back to these developments. Early feedback 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 work presented in this paper is supported by the European Commission (EC) (7th Framework Programme contract 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.

References

Albert, D., & Lukas, J. (Eds.) (1999). *Knowledge Spaces: Theories, Empirical Research, Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Doignon, J.-P., & Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, *23*, 175–196.

Doignon, J.-P., & Falmagne, J.-C. (1999). *Knowledge Spaces*. Berlin: Springer.

Falmagne, J.-C., Albert, D., Doble, D., Eppstein D., Hu X. (2013). *Knowledge Spaces: Applications in Education*. Berlin: Springer.

Heller, J., Steiner, C., Hockemeyer, C., & Albert, D. (2006). Competence-Based Knowledge Structures for Personalized Learning. *International Journal on E-Learning*, *5*, 75-88.

Korossy, K. (1999). Modeling Knowledge as Competence and Performance. In D. Albert & J. Lukas (Eds.), *Knowledge Spaces: Theories, Empirical Research, and Applications* (pp. 103-132). Mahwah, NJ: Springer.

Nakamura, Y., Tsuji, H., Seta, K., Hashimoto, K., & Albert, D. (2011) Visualization of Learner's State and Learning Paths with Knowledge Structures. In A. Konig et al. (Eds.), *Knowledge-Based and Intelligent Information and Engineering Systems*. KES 2011, Part IV, LNAI 6884 (pp. 261-270). Berlin: Springer

Rusch, A., & Wille, R. (1996). Knowledge spaces and formal concept analysis. In H. H. Bock & W. Polasek (Eds.), *Data analysis and information systems: Statistical and conceptual approaches* (pp. 427-436). Berlin: Springer.

Steiner, C., Kickmeier-Rust, M., & Albert, D. (2014). Learning Analytics and Educational Data Mining: An Overview of Recent Techniques. In: *Proceedings of the EC-TEL 2014 workshop ´Learning analytics for and in serious games´*, 6 – 15.

Wille, R. (1982). Restructuring lattice theory: An approach based on hierarchies of concepts. In I. Rival (Ed.), *Ordered sets* (pp. 445–470). Dordrecht: Reidel.

Wille, R. (2005). Formal Concept Analysis as Mathematical Theory of Concepts and Concept Hierarchies, In B. Ganter , G. Stumme and R. Wille (Eds.), *Formal Concept Analysis* (pp. 1-34). Berlin: Springer.