Ogata, H. et al. (Eds.) (2015). Proceedings of the 23rd International Conference on Computers in Education. China: Asia-Pacific Society for Computers in Education

Lea's Box: A Competency-oriented Approach to Facilitate Learning Analytics in School Settings

Michael D. KICKMEIER-RUST[a*], Michael BEDEK[a] & Dietrich ALBERT[a]
[a]*Knowledge Technologies Institute, Graz University of Technology, Austria*
*michael.kickmeier-rust@tugraz.at

Abstract: Learning Analytics is a key term of recent research and development in the educational sector. However, the uptake of such techniques is oftentimes limited to online courses and it is sparsely used in schools. The reason is that school-based teaching and learning is still an 'analogue' and personal process that is not producing the digital data that are necessary to conduct in-depth learning analytics. The Lea's Box project is addressing this problem by supporting teachers in their daily practice to collect data as easy and complete as possible to have at least the 'little data' required to make their teaching more individual and more formative. In addition, the project attempts to develop competence-oriented techniques for learning analytics on the basis of solid theories that have been developed in the context of intelligent tutorial systems. In this paper we present a summary about the developments and experiences with the tools and techniques in schools.

Keywords: Learning Analytics, Competence-based Knowledge Space Theory, Formal Concept Analysis, School Teaching

1. Introduction

Educational technologies are advancing rapidly; new solutions and online platforms appear every day. Mobile learning, learning on demand and media rich curricula are recent buzz words describing the "*techno-pedagogical*" state of the art. And not least, the research and development community is encircled by the hovering spirit of "*big data*", *learning analytics* and *educational data mining*. In educational practice, there is an increasing conceptual change towards a formative evaluation and support of learners and a strong orientation to competencies and meta-competencies such as the so-called 21st century skills. There is no doubt, that the pace and mode of learning must adapt to ever fast changing societal challenges.

Well, looking into classroom reality in Europe, we can find a very diverse situation: The most frequent situation in schools is that they are *technology lean*; there is little hardware and software, internet access is often not available, too slow or restricted. There are schools and regions in Europe where the use of the Internet in school is prohibited or strictly limited based on local policy, public opinion and/or parents' consent. It is seen as a source of danger (for example due to well-known cases of cyber mobbing, addictive gaming, etc.) where children need protection rather than the development of digital skills. Of course, there are schools and regions where the opposite is the case and technology is seen as an - still emerging - but already basic literacy skill. The use of (new) technologies is often dependent on the enthusiasm of individual teachers. However, even if teachers are motivated and enthusiastic about using and adapting ICT equipment, they might face obstacles due to mandatory security and organizational policies. Organizational structures usually do not support the use of massive personal devices like laptops, tablets and mobile phones in the classroom. Ultimately, the use of ICT (specifically with the aim of formative assessment) means collecting data on a large scale. With respect to this data collection and assessment the fear exists that assessment results (including data from the Programme for International Student Assessment, PISA) are used to measure the performance of an individual teacher and are thus opposed by teacher unions. Studies show, that if standard assessments like PISA become important, there is more "training for the test" going on and hence less time spent for individual student development. This makes a significant number of teachers' sceptic about the benefits of assessments and analyses in general (Rowlet, 2013).

In conclusion, there is sparsely "big data" and sometimes we do not even find "little data" in European school realities. Even if this perspective is perhaps a little bit larger than life, it nonetheless becomes obviously a long way to widely applied learning analytics with the key goal to make teaching more formatively inspired and more focusing on the individual as opposed to standardized "action - test - outcome" pedagogies.
The key question is how to support teachers in the real life's best. In this paper we introduce a European imitative, the Lea's Box project (www.leas-box.eu), which aims at providing simple and usable and realizable solutions, close to teacher practice, and which aims at bringing all the tiny little bits of data that are available together – for good.

2. Lea's Box – A learning Analytics Toolbox

Learning analytics (LA) and educational data mining (EDM) are more than recent buzz words in educational research: they signify one of the most promising developments in improving teaching and learning. While many attempts to enhance learning with mere technology failed in the past, making sense of a large amount of data collected over a long period of time and conveying it to teachers in a suitable form is indeed the area where computers and technology can add value for future classrooms. However, reasoning about data, and in particular learning-related data, is not trivial and requires a robust foundation of well-elaborated psycho-pedagogical theories. The fundamental idea of learning analytics is not new, of course. In essence, the aim is using as much information about learners as possible to understand the meaning of the data in terms of the learners' strengths and weaknesses, abilities, competences and declarative knowledge, attitudes and social networks, as well as learning progress, with the final goal of providing the best and most appropriate personalized support. Thus, the concept of learning analytics is quite similar to the idea of formative assessment. "Good" teachers of all time have strived to achieve exactly this goal. However, collecting, aggregating, storing and interpreting information about learners that originates from various sources and over a longer period of time (e.g., a semester, a school year, or even in a lifelong learning sense) requires smart technology. To analyze this vast amount of data, give it educational meaning, visualize the results, represent the learner in a holistic and fair manner, and provide appropriate feedback, teachers need to be equipped with the appropriate technology. With that regard, a substantial body of research work and tools already exist. Lea's Box aims to continue and enrich on-going developments and facilitate the broad use of learning analytics in the "real educational world".

Lea's Box concentrates on a competence-centered, multi-source formative assessment methodology based on sound psycho-pedagogical models, such as the *Competence-based Knowledge Space Theory* (CbKST) and the *Formal Concept Analysis* (FCA) which are to the very concrete demands and requirements of teachers and learners.

The tangible result of Lea's Box manifest in form of a Web platform for teachers and learners provide links to the existing components and interfaces to a broad range of educational data sources. Teachers will be able to link the various tools and methods that they are already using in their daily practice and that provide software APIs (e.g., *Moodle* courses, electronic tests, *Google Docs*, etc.) in one central location. More importantly, the platform hosts the newly developed LA/EDM services, empowering educators to conduct competence-based analysis of rich data sets. A key focus of the platform will enable teachers not only to combine existing bits of data but to allow them to "generate" and collect data in very simple forms, not requiring sophisticated hard- or software solutions. Finally, we want to open new ways to display the results of learning analytics - leaving the rather statistical dashboard approach, moving towards structural visualizations and towards opening the internal learner models.

2.1 Generating and Collecting Data

The major difference between typical learning analytics scenarios and school reality is the degree to which an instructor and a learner are supposed to face some sort of digital device. Typical scenarios are e-learning courses, perhaps popular *MOOC*s, where a learner is producing data with each and every mouse click. In school, students are most often required to make their homework the old-fashioned paper pencil style. So, at best, the amount of data that is generated is a final grade for the homework.

Teachers are building their appraisal of students rather intuitive and experience-based instead of a solid, fair, objective data-based and evidence-based approach.

Thus, we developed a tool to allow teachers collecting data using a simple and cheap tablet computer. Based on design workshops with over 100 teachers from Austria, the Czech Republic, Germany, and Turkey, we set up the key needs, the key obstacles, and the key mental models of teachers. The outcome was a tool, named *myClass*, to collect and record data about activities, learning processes, and achievements very easily and independent from Wi-Fi connection. The tool is device-independent and can be used with smartphones, tablet computers, or regular computers. The following figure shows a screen shot of the *myClass* tool.

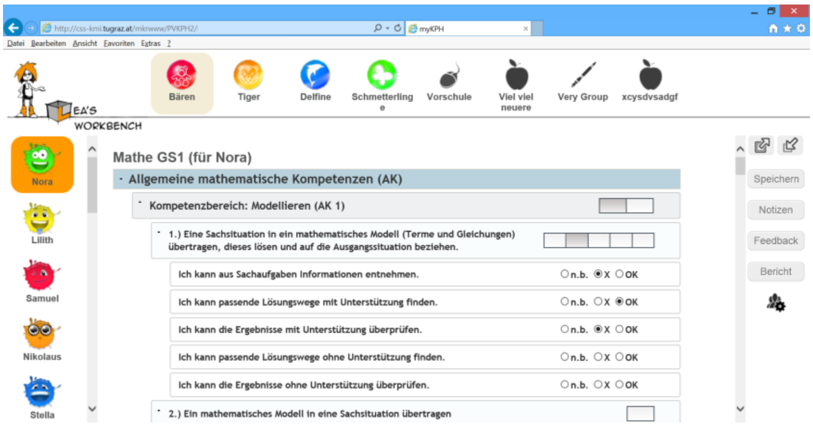

Figure 1. Screenshot of the lightweight classroom tool *myClass*.

The tool is composed of various modules that can be easily arranged by the teachers themselves to find exactly the configuration and feature they need and prefer most and. Now, not only the data collected by a single teacher might be of interest, *myClass* allows bringing together all the data from all teachers of a school; this, in turn, provides extra information for teachers. For example, the appraisal of the regular math teacher with the view of the afternoon tutor can be compared with each other. If there appear severe differences, there would be a clear need for reflect upon the reasons of these different evaluations.

Certainly, we cannot expect a perfect data source. The policy of Lea's Box is to make the maximum of whatever is available. A nice example from our experiences with applying the solutions in schools is a project named "personal responsibility". In this project, a partner school in Austria attempts to bring all students and teachers together once a month. In the plenum, certain agreements are made on a school-wide basis; for example, to take more personal responsibility in social conflicts. The key question for teachers and the principle is whether such "costly" efforts of bringing all people together and spend a certain amount of time on such projects, pays off in the end. Using *myClass*, a teacher can freely define positive and negative activities. These defined activities are accessible through the *myClass* system and can be counted with a single finger touch on a tablet computer. A key features of the *myClass* application which is of high relevance for teachers is the opportunity to generate report cards automatically and to generate materials and reports for teacher-parents conferences.

2.2 A Focus on Competences rather than on Performance

2.2.1 Competence-based Knowledge Space Theory

While the primary platform of Lea's Box provides needful yet simple tools tailored to the basic demands of teachers, it puts a string emphasis on student´s competencies. The foundation for our work is a conceptual psycho-pedagogical theory named *Competence-based Knowledge Space Theory* (CbKST).

The original *Knowledge Space Theory* (KST), founded by Doignon and Falmagne (1999, 2011), and extensions such as the CbKST, are coming from the genre of autonomous intelligent and adaptive tutoring systems. The idea was to broaden the ideas of the linear Item Response Theory (IRT) scaling, where a number of items are arranged on a single, linear dimension of "difficulty". In essence, KST provided a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. More recent advancements of the theory accounted for a probabilistic view of test results and they introduced a separation of observable performance and the actually underlying abilities and knowledge of a person. Such developments lead to a variety of theoretical, competence-based approaches (cf. Albert & Lukas, 1999 for an overview). An empirically well-validated approach to CbKST was introduced by Korossy (1999); basically, the idea was to assume a finite set of more or less atomic competencies (in the sense of some well-defined, small scale descriptions of some sort of aptitude, ability, knowledge, or skill) and a prerequisite relation between those competences.

In a first step, CbKST attempts to develop a model of the learning domain, e.g. algebra. Examples for such competencies might be the knowledge what an integer is or the ability to add two positive integers and so on. The level of granularity to which a domain is broken down depends on the envisaged application and might range from a very course-grained level on the basis of lessons (for example to plan a school term) to a very fine-grained level of atomic entities of knowledge/ability (for example as the basis of an intelligent problem solving support application). In a second step, CbKST looks into a natural course of learning and development and into logical prerequisites between competencies. Usually, learning and the development of new abilities as well as the stabilization of skills occurs along developmental trajectories. On the basis of a set of competencies and a set of prerequisite relationships between them, we can formally derive a collection of so-called competence states (Figure 2). Due to such prerequisite relations between the competencies, not all subsets of competencies (which would result in the power set) are plausible competence states.

So far, the structural model focuses on latent, unobservable competencies; loosely speaking the model makes hypotheses about the brain's black box. By utilizing interpretation and representation functions the latent competencies are mapped to evidence or indicators relevant for a given domain. Such indicators might be test items but might refer to all sorts of performance or behavior (e.g., the concrete steps when working with a spread sheet application). Due to these functions, latent competencies and observable performance can be linked in a broad form. This means that an entire series of indicators can be linked to underlying competencies. The CbKST accounts for the fact that indictors such as test items cannot be perfect evidence for the latent knowledge or ability. There is always the possibility that a person makes a lucky guess or exhibits a correct behavior/activity just by chance. In turn, a person might fail in a test item although the necessary knowledge/ability is actually available, for example, by being inattentive or careless. Thus, CbKST considers indicators on a probability-based level, this means

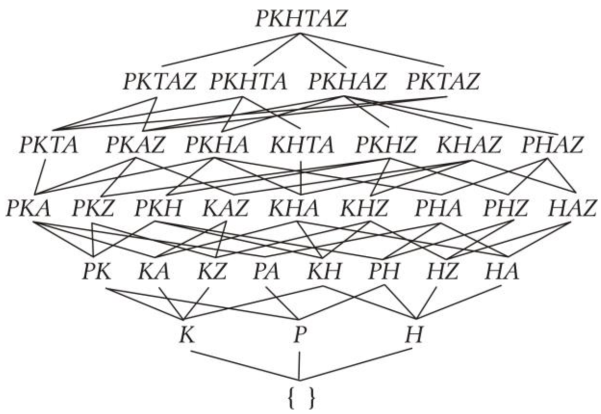

Figure 2. A prototypical competence space.

That mastering a test item suggest having the underlying competencies with a certain probability. Conceptually, this view constitutes a probability distribution over the competence structure. A further significant advantage of such approach is that learning is not only considered a one dimensional course on a linear trajectory, equal for all learners. Learning and development rather occur along one of an entire range of possible learning paths.

Recent advancements of CbKST primarily concern the integration of theories of human problem solving (given that most indictors can be interpreted as solving some sort of problem). This work was essentially driven in the genre of smart, educationally adaptive computer games for learning – loosely speaking for developing an educational AI support the players of the game (Kickmeier-Rust & Albert, 2012).

2.2.2 Formal Concept Analysis

Formal Concept Analysis (FCA) describes concepts and concept hierarchies in mathematical terms, based on the application of order and lattice theory (Wille, 1982). The starting point is the definition of the formal context which can be described as a triple consisting of a set of objects, a set of attributes, and a binary relation between the objects and the attributes (e.g., object *A* has attribute *B*). 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 Figure 3. Teachers use the tool to define the formal context and to add learning resources which can be assigned objects and to attributes.

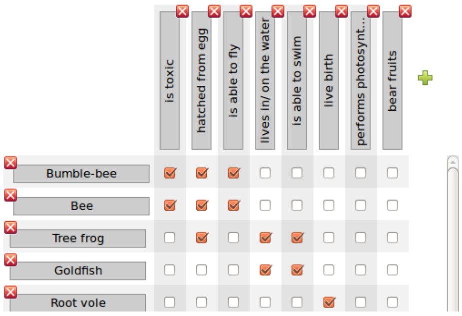

Figure 3. FCA-tool´s Editor View for creating a domain with objects, attributes, and relations.

Theoretically speaking, in order to create a concept hierarchy (called concept lattice), for each subset $A \in G$ and $B \in M$, the following derivation operators need to be defined:

$A \mapsto A' := \{m \in M \; I \; g \; I \; m \text{ for all } g \in A\}$, which is the set of common attributes of the objects in *A*, and $B \mapsto B' := \{g \in G \; I \; g \; I \; m \text{ for all } m \in B\}$, 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 Wille, 2005), which can be represented as a labelled line diagram (see Figure 4).

Every node of the lattice represents a formal concept. The extension *A* of a particular formal concept is constituted by the objects whose labels 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 whose lables can be reached by an ascending path from that node. In the example above, the formal concept´s intension consists of {is able to swim, lives in / on the water}.

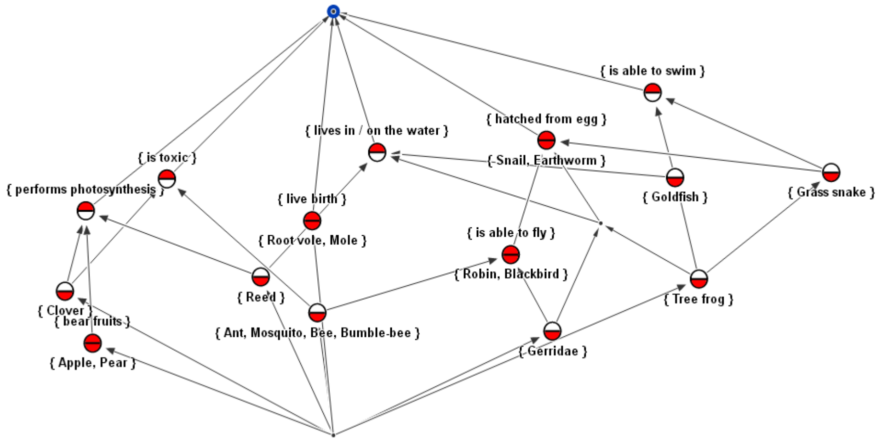

Figure 4. Concept lattice.

Similar as described by Rusch and Wille (1996) who were the first who applied the FCA with students and their performance data, we suggest formal contexts with student as “attributes” and competencies and skills as “objects”. The relation between these two sets means “student *m* holds competency *g*”. By such a concept lattice, a variety of different information can be displayed. Examples are overlaps and differences of students’ skills and competences or the visualization of the learning progress of an entire class over time.

3. Visualizing Competence-centered Learning Outcomes

One of the project partners in Lea’s Box is a company named *Scio* (www.scio.cz) from the Czech Republic. This company is responsible for nationwide standardized school entry exams. This gives us one the one hand a rich data basis for research and, on the other hand, access to Czech teachers. In an evaluation study we modelled a standard school entry test for secondary school mathematics. This knowledge domain covers essentially basic mathematical skills (such as solving simple equations) as well as skills like logical thinking or reasoning abilities. For this domain we identified a set of involved skills and competencies and we derived the competency space. Finally, we linked the competency states to the items of the national test. The competence model encompasses the competences and the prerequisite relations between them can represented as a Hasse-diagram as shown in Figure 6. Based on this competence model, the competence space which consists of of the ordered set of all plausible competence states can be derived (see Figure 7).

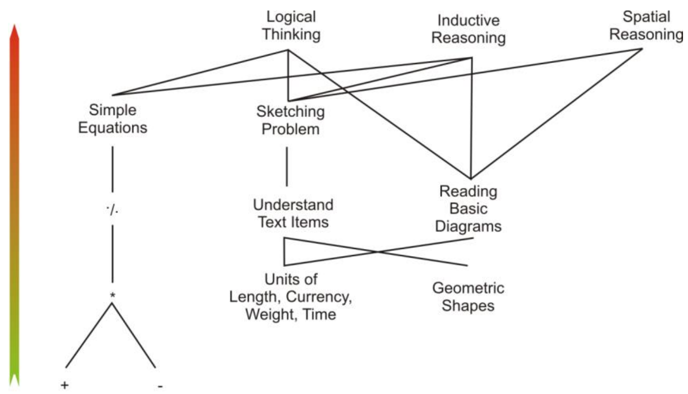

Figure 6. A competence model in the domain of mathematics
(the difficulty level increases from bottom to top)

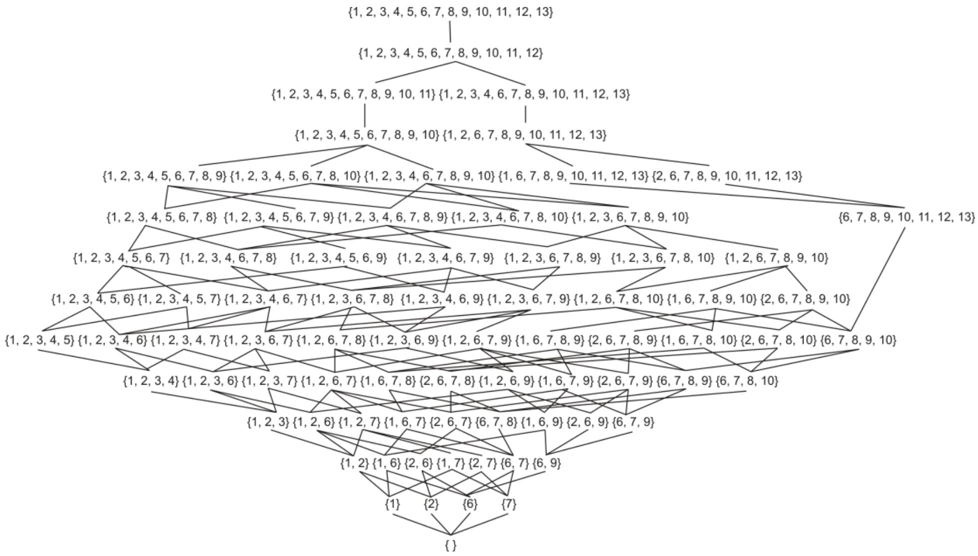

Figure 7. The competence space derived by the relation in Fig. 6.

Twelve teachers were exposed to the model and various visualizations in the context of qualitative design studies. Figure 8 illustrates a weak and a strong performer. As results, we found a reasonable trade-off between information density (complexity) and comprehensibility. Based on the concrete recommendations, we reduced the amount of information and adopted broader color coding features. The new features are presently investigated on a quantitative basis.

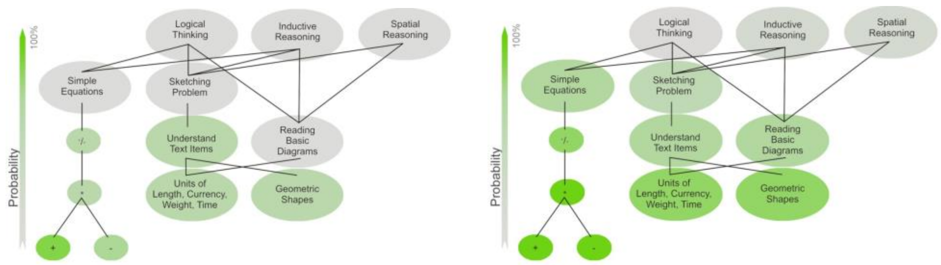

Figure 8. A visualization for teacher to display weak (left image) and good (right image) performes and their competencies.

4. Conclusions

There is little doubt that frameworks, techniques, and tools for LA will increasingly be part of a teacher's professional life in the near future. Some indications for that were already mentioned in section 2.1 of this paper. The benefits are convincing – using the (partly massive) amount of available data from the students in a smart, automated, and effective way, supported by intelligent systems in order to have all the relevant information available just in time and at first sight. The ultimate goal is to formatively evaluate individual achievements and competencies and provide the learners with the best possible individual support and teaching. The idea of formative assessment and educational data mining is not new but the hype over recent years resulted in scientific sound and robust approaches becoming available, and usable software products appeared. However, when surveying the educational landscape, at least that of the EU, the educational daily routines are different. We face technology-lean classrooms and schools, we face a lack of proper teacher education in using ICT in schools – not mentioning of using techniques of LA in schools. We face a certain aloofness to use breaking educational technologies and a well-founded pedagogical view that learning ideally is analogous and socially embedded and doesn't occur in front of some kind of electronic device. These are all experiences and results of a large scale European research project named Next-Tell (www.next-tell.eu) that was looking into educationally practices across Europe and that intended to support teachers where exactly they are today with suitable ICT as effective and as appropriately as possible.

Psychologically-sound frameworks such as the CbKST and new developments and extensions of the FCA which offer a rigorously competence-based, probabilistic, and multi-source approach account for the recent conceptual change in Europe’s educational systems; which is a shift towards a more competence-oriented education including multi-subject competencies and superordinate 21st century (soft) skills.

No matter if data are rich or lean; a teacher is supported to the best possible degree and with a variety of important information about individual and group-based learning processes and performance of learners as well as about the educator’s own performance. The probabilistic dimension enables teachers to have a more cautious view of individual achievements – it might well be that a learner has a competency but fails in a test; vice versa, a student might luckily guess an answer.

From an application perspective, in the context of European projects we developed and evaluated tools that cover the techniques and approaches described in this paper (available through the Lea’s Box website www.leas-box.eu). We piloted various school studies and gathered feedback from teachers. In the end, and this can be considered an outlook for future developments, we had to find out that the ‘massive’ visualizatons (i.e. Hasse diagrams and concept lattices) are overburdening teachers’ understanding and mental models about individual and class-based learning. Moreover, in order to understand the classical Hasse diagrams, it required (too) massive efforts in training teachers to fully utilize the potentials of those diagrams.

Therefore, recent efforts, e.g., in the Lea’s Box project, seek to adjust and advance the classical Hasse diagrams to such visualizations that are intuitively understood by educators and, at the same time, hold the same density of information. In conclusion, the utility of competence-centered approaches to LA, involving a separation of latent competencies and observable behaviors and performances, as well as having a conservative, probabilistic, multi-source approach appears to be a striking classroom-oriented, next-level contribution to LA, learner modelling, and model negotiations.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Program (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.

References

Albert, D., & Lukas, J. (1999). Knowledge spaces: Theories, empirical research, and applications. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Falmagne, J-C., & Doignon, J.-P. (2011). Learning Spaces. Berlin: Springer.

Kickmeier-Rust, M. D., & Albert, D. (Eds.) (2012). An Alien's guide to multi-adaptive educational games. Santa Rosa, CA: Informing Science Press.

Korossy, K. (1999). Modelling knowledge as competence and performance. In D. Albert & J. Lukas (Eds.), Knowledge spaces: Theories, empirical research, and applications (pp. 103–132). Mahwah, NJ: Lawrence Erlbaum Associates.

Rowlett, P. J. (2013). Developing a Healthy Skepticism About Technology in Mathematics Teaching. Journal of Humanistic Mathematics, 3(1), 136-149.

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.

Wille, R. (1982). Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival, I. (ed.) Ordered sets, pp. 445-470. Reidel, Dordrecht-Boston.

Wille, R. (2005). 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. Springer, Berlin.