An Invitation to Knowledge Space Theory

Bernhard Ganter[1]([✉]), Michael Bedek[2], Jürgen Heller[3], and Reinhard Suck[4]

[1] Technische Universität Dresden, Dresden, Germany
bernhard.ganter@tu-dresden.de
[2] Technische Universität Graz, Graz, Austria
[3] Universität Tübingen, Tübingen, Germany
[4] Universität Osnabrück, Osnabrück, Germany

Abstract. It has been mentioned on many occasions that Formal Concept Analysis and KST, the theory of Knowledge Spaces, introduced by J.-P. Doignon and J.-C. Falmagne, are closely related in theory, but rather different in practice. It was suggested that the FCA community should learn from and contribute to KST. In a recent workshop held at Graz University of Technology, researchers from both areas started to combine their views and tried to find a common language. This article is a partial result of their effort. It invites FCA researchers to understand some ideas of KST by presenting them in the language of formal contexts and formal concepts.

1 Introduction

Knowledge Space Theory (KST) [6] and *Formal Concept Analysis (FCA)* [21] share a range of similarities: Both approaches have been invented in the early 1980s and since then have experienced an astonishing development from a theoretical as well as a practical point of view. Since the very beginning, researchers from both fields were aware of each other, and it was clear that the approaches are also similar in their mathematical foundations (A first translation was worked out by Rusch and Wille [18]). However, KST and FCA developed and grew primarily in parallel, perhaps due to the different fields of application. While FCA always aimed at a broad spectrum of applications, KST from its beginnings has focused on the topic of *learning* and was impressively successful there.

This article is another invitation to the FCA community not only to be aware of KST, but to learn from its success and to contribute, whenever possible, to its theoretical basis. We recommend to study the KST literature, in particular the two monographs by Doignon and Falmagne [6,8]. Here we present some of the basic ideas in FCA language, hoping to spark the interest of FCA researchers.

In a nutshell, KST aims at an adaptive, and thus efficient, assessment of a persons' knowledge state with respect to a given knowledge domain. Such a domain is given by a (finite) set $\mathcal{Q}$ of questions, and a knowledge state is the set of problems the person is capable to master. Details will follow below. We sometimes hear from colleagues in pedagogical research that such an approach is much too simple to model a learning process. One must however admit that KST

is remarkably and measurably successful in such applications. The professional ALEKS software (**A**ssessment and **LE**arning in **K**nowledge **S**paces) is an online tutoring and assessment program with many users that includes course material in mathematics, chemistry, introductory statistics, and business.

There are extensions of KST, for example *Competence-based Knowledge Space Theory (CbKST)* [1,16,17], which aims at an adaptive assessment of a person's skills and competencies that enable him or her to master certain problems. Based on that, CbKST suggests pedagogically reasonable learning paths. Natural applications of (Cb)KST are intelligent tutoring systems, as well as other kinds of technology-enhanced learning applications (e.g., game-based learning).

A main theoretical difference between the two fields is that FCA works with *closure systems*, that is, with families of sets that are closed under intersections, while knowledge spaces are closed under the *union of sets*. The translation seems obvious and is indeed not difficult.

2   Knowledge States and Formal Concepts

We investigate a *knowledge domain*, which is a finite set $\mathcal{Q}$ of **questions** (also called **tasks** or **test items**), which learners may or may not be able to master (see Example 6 for an instance). Each learner has an individual **knowledge state**, characterized by the collection of those questions in $\mathcal{Q}$ which that learner is able to solve. Such **learner-task data** is conveniently recorded in form of a formal context

$$(\mathcal{L}, \mathcal{Q}, \Box),$$

where the $\Box$ symbol denotes the "masters" relation between learners and questions. Symbolically, we abbreviate the statement that "learner l masters question q" to

$$l \, \Box \, q.$$

But beware! The table should not be understood as *empirical* data, which typically would contain *careless errors and lucky guesses* of the learners. Instead, each row of this data is meant to record the respective learner's true knowledge state, formally

$$l^{\Box} := \{q \in \mathcal{Q} \mid l \, \Box \, q\}$$

for each learner $l \in \mathcal{L}$. With a similar notation, we write for each question $q \in \mathcal{Q}$

$$q^{\Box} := \{l \in \mathcal{L} \mid l \, \Box \, q\}$$

to denote the set of those learners who master that question.

Example 1. *Figure 1 shows a very small example of a learner-task context. It lists nine learners $l_1, \ldots, l_9$, and eight tasks $a, \ldots, h$. Learner l_4 masters only tasks e, f, and g, so that the knowledge state of l_4 is*

$$l_4^{\Box} = \{e, f, g\}.$$

Problem g is mastered by learners l_3, l_4, l_6, l_8, and l_9, abbreviated

$$g^{\Box} = \{l_3, l_4, l_6, l_8, l_9\}.$$

□	a	b	c	d	e	f	g	h
l_1								
l_2	×							
l_3	×	×	×	×	×		×	×
l_4					×	×	×	
l_5		×	×	×	×			
l_6		×	×	×	×	×	×	
l_7		×	×			×		
l_8	×	×	×	×	×	×	×	
l_9	×	×	×	×	×	×	×	×

Fig. 1. A learner-task context

Usually not every subset of $\mathcal{Q}$ qualifies as a possible knowledge state, since some of the tasks may be *prerequisites* for others. For example, it seems unlikely that a learner can divide, but not add.

The set $\mathcal{K}$ of all feasible knowledge states over $\mathcal{Q}$ is called a **knowledge structure** and is often denoted as $(\mathcal{Q}, \mathcal{K})$. Although often very large, any knowledge structure $(\mathcal{Q}, \mathcal{K})$ can be written down in the same form as learner-task data, namely as

$$(\mathcal{K}, \mathcal{Q}, \ni).$$

General knowledge structures are difficult to handle, since they are sets of sets and can easily be of enormous size. Reasonable assumptions about structural properties of knowledge structures are therefore welcome, because they may simplify their use. One such assumption relies on the idea that each knowledge state must be obtainable through a step-wise *learning* process, which respects the prerequisites, and that the knowledge states are precisely all subsets of $\mathcal{Q}$ that can be learned. A consequence of this assumption is that only *missing* prerequisites can be obstacles to learning, but the presence of additional knowledge never is. This in turn implies that in theory every learner could acquire the knowledge of each other learner, and that therefore the set union of any two knowledge states is again a knowledge state. A knowledge structure which is closed under arbitrary set unions is called a **knowledge space**.

The assumption that knowledge spaces are closed under unions of states may surprise a little. There are several possible justifications for this. One is that, since knowledge must be acquired step by step, there must be a graded *learning path* to every knowledge state. This leads to an even more specialized notion, that of a *learning space* (or *antimatriod* or *well-graded knowledge space*). We shall not treat this here and refer the reader to Falmagne and Doignon [8] (but we shall mention a characterization of these spaces in Sect. 8). A more simple-minded approach is to consider a knowledge space as consisting of those task sets which can be solved by *groups of learners*, assuming that a group solves a problem if at least one group member does. This obviously leads to union-closed families of states.

The knowledge space *generated* by given knowledge states consists of all set unions of these states, including the empty union, $\emptyset$. The knowledge space generated by the states in given learner-task data $(\mathcal{L}, \mathcal{Q}, \square)$, such as in Fig. 1, is denoted by

$$\mathcal{K}(\mathcal{L}, \mathcal{Q}, \square).$$

In Formal Concept Analysis terms, we find that this knowledge space consists of the intent-complements of $(\mathcal{L}, \mathcal{Q}, \boxslash)$, i.e.,

$$\mathcal{K}(\mathcal{L}, \mathcal{Q}, \square) = \{\mathcal{Q} \setminus B \mid (B', B) \in \underline{\mathfrak{B}}(\mathcal{L}, \mathcal{Q}, \boxslash)\}.$$

Example 2. *The knowledge space generated by the learner-task data in Fig. 1 has 16 elements. A diagram is shown in Fig. 2. We give an examples of how to read this diagram:*

– *Consider the node labeled with d. It represents the knowledge state $\{a, b, c, e, f, g\}$, which is the union of the knowledge states of the learners l_1, l_2, l_4, and l_7. The learners are the ones that can be reached via downward paths from the node. The tasks are the ones which* cannot *be reached via upward paths from the node labeled d.*
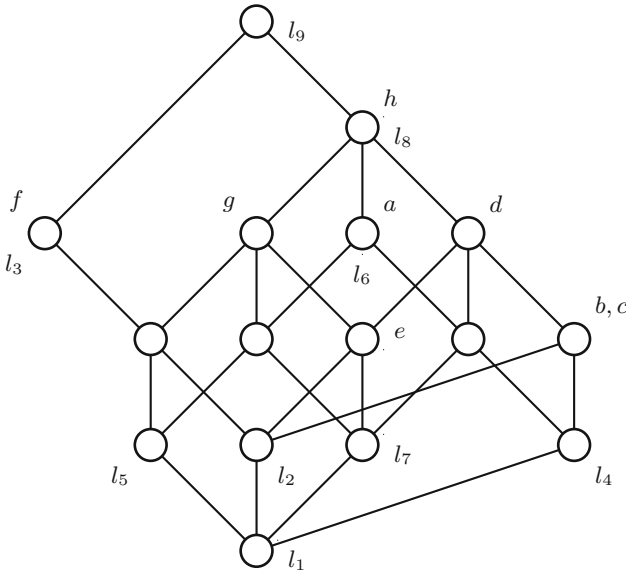

Fig. 2. The knowledge space generated by the learner-task data in Fig. 1, represented as the concept lattice of the complementary context $\underline{\mathfrak{B}}(\mathcal{L}, \mathcal{Q}, \boxslash)$. The concept intents are the complements of the knowledge states. The extents consists of those learners, whose knowledge state is contained in the one represented by the respective node.

Each knowledge space has a unique **basis** $\mathcal{B}$. It consists of the $\cup$-irreducible states, i.e., of those states that cannot be obtained as unions of other states.

It can be obtained from $(\mathcal{K}, \mathcal{Q}, \ni)$ by a process which we call **coreduction**, by removing all rows that can be written as set unions of other rows. This corresponds to (object-) reduction in Formal Concept Analysis, except that set union is used instead of set intersection. In Formal Concept Analysis terms one first goes to the complementary context

$$(\mathcal{K}, \mathcal{Q}, \not\ni),$$

object-reduces it and complements the outcome to obtain the basis.

Example 3. *The learner-task data in Fig. 1 is not coreduced because, for example, the knowledge state of learner l_6 is the set union of the states of learners l_4 and l_5:*

$$\{b, c, d, e, f, g\} = l_6^\square = l_4^\square \cup l_5^\square = \{e, f, g\} \cup \{b, c, d, e\}.$$

The basis of the knowledge space in Fig. 2 consists of five states, see Fig. 3.

□	a	b	c	d	e	f	g	h
l_2	×							
l_3	×	×	×	×	×		×	×
l_4					×	×	×	
l_5		×	×	×	×			
l_7		×	×			×		

Fig. 3. The basis of the knowledge space generated by the learner-task context in Fig. 1 consists of the complements of the object intents of the object-reduced formal context.

The basis uniquely determines the knowledge space. There are fast algorithms for generating a knowledge space from its basis. For example one can apply the standard NEXT INTENT algorithm to the complementary context and obtain the knowledge states as the complements of the generated intents. Actually, these algorithms can do more. For example, one often has only *incomplete information* about a learner's knowledge state. Some set $P \subseteq \mathcal{Q}$ may be known of questions which that learner has passed, and also some set F of questions which were failed. It then is clear that the learner's state must be some $S \in \mathcal{K}$ satisfying $P \subseteq S$ and $S \cap F = \emptyset$. The above-mentioned algorithm can indeed generate exactly these states, given a basis together with the sets P and F.

But complexity remains an issue, even when fast algorithms are available. This is due to the huge number of possibilities. Consider, for example, the case that only $|\mathcal{Q}| = 7$ questions are investigated. It was shown by Colomb et al. [4] that there are as many as 14 087 648 235 707 352 472 different knowledge spaces over seven questions. The largest one of them has 128 $(= 2^7)$ states and a basis of size seven. It contains all other knowledge spaces over the same seven questions. Among them is one that has only 65 states, but a basis of size 35.

3   Prerequisites and Implications

A knowledge space can be specified by explicitly listing up sufficiently many of its knowledge states, so that the space then can be generated from these via set unions. However, this may be rather tedious and not intuitive. Another approach is to use expert judgments on the dependencies between questions. Such a judgment could say that question q_2 is more difficult than q_1 in the sense that *any learner who masters q_2 also masters q_1*. This is abbreviated to $q_1 \leq q_2$, and a natural assumption is that this relation $\leq$ is reflexive and transitive. Such relations are called **quasi-order**s or, synonymously, **preorder**s.

Example 4. *The quasi-order for the data in Fig. 1 is given by the diagram in Fig. 4.*
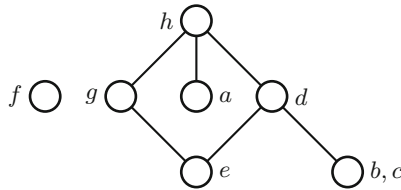

Fig. 4. The quasi-order for the learner-task data in Fig. 1. Whoever masters d also masters b, c, e, but not necessarily a, f, g or h.

A subset of $\mathcal{Q}$ **disrespects** a judgment $q_1 \leq q_2$ if it contains q_2, but not q_1. Otherwise, it **respects** it. Given a collection $\mathcal{J}$ of such judgments, we may consider the set

$$\mathcal{K} := \{S \subseteq \mathcal{Q} \mid S \text{ respects all judgments in } \mathcal{J}\}.$$

It is easy to see that this is a knowledge space: if two sets both respect $q_1 \leq q_2$, then so does their set union. Thus judgments can be used to define knowledge spaces. However, the knowledge spaces obtained in this way are rather special, since they are also closed under set intersections. They are called **quasi-ordinal** knowledge spaces, because their states are precisely the "downsets" or "order ideals" of the quasi-orders $\leq$.

Most spaces are not quasi-ordinal. Of the 14 087 648 235 707 352 472 different knowledge spaces over seven questions only 9 535 241 are quasi-ordinal. To describe all knowledge spaces we therefore need a more general form of expert judgment. It can be shown that judgments of the form

each learner who masters question q also masters at least one of the questions $q_1, \ldots, q_n$

are sufficient for specifying arbitrary knowledge spaces. This is no surprise since these expressions correspond to *implications* between attributes of the complementary context: they can be reformulated in a logically equivalent form

if a learner masters none of the questions $q_1, \ldots, q_n$, then that learner cannot master q.

Symbolically, we abbreviate this to

$$\neg A \rightarrow \neg q \qquad (\text{where } A := \{q_1, \ldots, q_n\}).$$

Mixed approaches are possible using the techniques of **attribute exploration** [11]. Typically information is collected in form of some judgments $A_1 \rightarrow q_1$, $A_2 \rightarrow q_2$, . . . and of some "observations" (P_1, F_1), (P_2, F_2), . . . , as described above. The exploration algorithm then finds, if possible, the *simplest piece of information* missing for a complete description of a knowledge space, and this can be given to an expert as a query.

Example 5. *The knowledge space generated by the learner-task data in Fig. 1 is not quasi-ordinal. It can be checked that its quasi-order, depicted in Fig. 4, is respected by 34 sets, which means that 34 is the number of states in a quasi-ordinal knowledge space with this quasi-order. But as already mentioned, the number of set unions of the states in Fig. 1 is only 16. An example of a "non-ordinal" rule that holds in Fig. 1 is that each learner who masters g also masters one of f and h, i.e., that*

$$\neg\{f, h\} \rightarrow \neg g.$$

4   A Quasi-Ordinal Example

Typically the problems in a knowledge domain $\mathcal{Q}$ are assumed to be dichotomous. Responses are coded as either correct (solved), or incorrect (not solved).

Example 6. *The knowledge domain $\mathcal{Q} = \{a, b, c, d, e, f\}$ defined in Table 1 is taken from the beginning algebra items in the ALEKS system. Below the table we show a prerequisite quasi-order $(\mathcal{Q}, \leq)$ (which in this case actually is an order), as it is used by ALEKS.*

So ALEKS assumes that a learner which masters Question f also masters all the other questions, while any learner mastering c also masters a, but not necessarily the other ones. The sets respecting this prerequisite order are precisely the downsets (order ideals). So these are the knowledge states if no other prerequisite conditions are imposed. It is well known that the family of order-ideals of a quasi-order is closed under set union and set intersection and therefore forms a distributive lattice. In fact, Birkhoff's representation theorem (Theorem 39 in [12]) shows that every finite distributive lattice is obtained in this way.

It is well known that the lattice of order ideals of a quasi-order $(\mathcal{Q}, \leq)$ is isomorphic to the concept lattice of the corresponding *contra-ordinal scale* $(\mathcal{Q}, \mathcal{Q}, \not\geq)$ (see [12], p. 49). Figure 6 shows this scale (derived from Fig. 5) and its concept lattice.

The *learning paths* of a knowledge space are defined as the maximal chains in the lattice of knowledge states. Figure 7 shows the learning paths for the

Table 1. An example of a knowledge domain from the ALEKS system.

|	Problem type	Instance
a	Word problem on proportions	A car travels on the freeway at an average speed of 52 miles per hour. How many miles does it travel in 5 h and 30 min?
b	Plotting a point in the coordinate plane	Using the pencil, mark the point at coordinates $(1, 3)$
c	Multiplication of monomials	Perform the following multiplication: $4x^4y^4 \cdot 2x \cdot 5y^2$ and simplify your answer as much as possible
d	Greatest common factor of two monomials	Find the greatest common factor of expressions $14t^6y$ and $4tu^5y^8$. Simplify your answer as much as possible
e	Graphing the line through a given point with a given slope	Graph the line with slope -7 passing through the point $(-3, -2)$
f	Writing the equation of the line through a given point and perpendicular to a given line	Write an equation for the line that passes through the point $(-5, 3)$ and is perpendicular to the line $8x + 5y = 11$
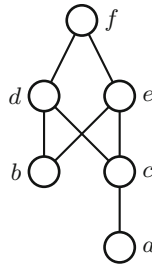
Fig. 5. The prerequisite (quasi-)order for the knowledge domain shown in Table 1

"beginning algebra" domain, depicted as a tree (right hand side). On the left, it is indicated how they can be counted for any given knowledge space. In the case of quasi-ordinal knowledge spaces, the learning paths correspond to the linear extensions of the underlying quasi-order. In general, the number of learning paths can be considerably smaller. The knowledge space shown in Fig. 2 has 18 learning paths, while the ordered set of its questions (shown in Fig. 4) has 198 linear extensions.

5   Probabilistic Knowledge Structures

The probabilistic theory of knowledge spaces is much more elaborate than what is available for FCA. This is due to the applications: it is desirable to guess a
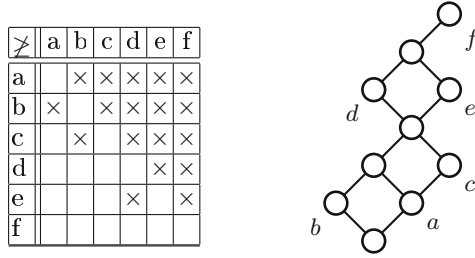
Fig. 6. The contra-ordinal scale and the knowledge space for the data in Table 1. The 10 knowledge states can be read from the reduced labeling as usual in Formal Concept Analysis. For example, the knowledge state corresponding to the "waist" node in the middle is $\{a, b, c\}$, since these are the labels at downward paths starting from this node.
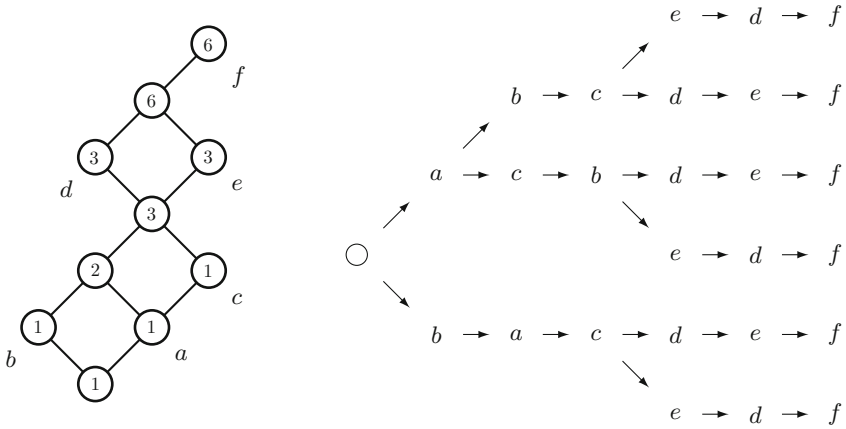

Fig. 7. The learning paths for the "beginning algebra" domain. Each number in the lattice diagram indicates how many learning paths lead to the respective knowledge state. They are obtained recursively by assigning 1 to the least element, and to each other element the sum of the numbers assigned to its lower neighbours. The right diagram shows the paths as a tree.

learner's knowledge state from only a few observations. Once the knowledge state is known, one can determine its "outer fringe", consisting of those problems that the learner "is ready to learn" (i.e., for which all prerequisites are already there), and this can be used with a tutorial system. The parameters of the stochastic models can be adjusted to observed learners' behaviour. Tutorial systems based on probabilistic KST are therefore able to improve upon usage. For a systematic presentation we refer the reader to Falmagne and Doignon [8]. Here we sketch some first ideas.

In practical applications we cannot assume that a student's response to an item is correct if and only if the student masters it (i.e. if the item is an element of the respective knowledge state). There are two types of response errors, namely

– careless errors, i.e. the response is *incorrect* although the item is an element of the learner's knowledge state (and could potentially be *mastered*),
– lucky guesses, i.e. the response is *correct* although the item is not an element of the learner's knowledge state (and could regularly *not* be *mastered*).

In any case, we need to dissociate knowledge states and response patterns. Let R denote a response pattern, which is a subset of $\mathcal{Q}$, and $\mathcal{R} = 2^{\mathcal{Q}}$ the set of all response patterns.

In this approach the knowledge state K is a latent construct, and the response pattern R is a manifest indicator to the knowledge state. The conclusion from the observable response pattern to the unobservable knowledge state can only be of a stochastic nature. This requires to introduce a probabilistic framework capturing the facts that knowledge states will not occur with equal probability, and that response errors will occur with a certain probabilities.

A *probabilistic knowledge structure* $(\mathcal{Q}, \mathcal{K}, P)$ is defined by specifying a (marginal) distribution $P_{\mathcal{K}}$ on the states of $\mathcal{K}$, and the conditional probabilities $P(R \mid K)$ for all $R \in \mathcal{R}$ and $K \in \mathcal{K}$. The marginal distribution $P_{\mathcal{R}}$ on $\mathcal{R}$ then is predicted by

$$P_{\mathcal{R}}(R) = \sum_{K \in \mathcal{K}} P(R \mid K) \cdot P_{\mathcal{K}}(K).$$

The probabilistic knowledge structure that received most attention is the *basic local independence model* (BLIM), which satisfies the following condition: For each $q \in \mathcal{Q}$ there are real constants $0 \leq \beta_q < 1$ and $0 \leq \eta_q < 1$ such that for all $R \in \mathcal{R}$ and $K \in \mathcal{K}$

$$P(R \mid K) = \left(\prod_{q \in K \setminus R} \beta_q \right) \cdot \left(\prod_{q \in K \cap R} (1 - \beta_q) \right) \cdot \left(\prod_{q \in R \setminus K} \eta_q \right) \cdot \left(\prod_{q \in \mathcal{Q} \setminus (R \cup K)} (1 - \eta_q) \right).$$

The constants β_q and η_q are interpreted as the probabilities of a careless error and a lucky guess, respectively, on item q.

This condition entails two important properties. First, it assumes local stochastic independence. Given the knowledge state K, the responses are stochastically independent across the items. This means that any dependence between the responses to different items is mediated by the knowledge state K. Second, the response errors only depend on the respective item, but not on the knowledge state (beyond the fact that the item may be an element of the state, or not). This is a strong constraint which reduces the number of free parameters drastically and renders the BLIM testable. Procedures for estimating its parameters were proposed by Heller and Wickelmaier [14].

6   Competence-Based Knowledge Space Theory

The notion of a knowledge structure was formulated explicitly without reference to cognitive or mental concepts, like "latent traits", or specific problem

components. Concepts like this were only introduced later on. There are parallel, but only partially overlapping approaches: Falmagne et al. [9], Doignon [5], Doignon and Falmagne [6], Albert and Held [1], as well as Korossy [15], and finally Düntsch and Gediga [7] and Gediga and Düntsch [13]. Korossy's approach is in a sense more general, as it allows for having a structure on the underlying cognitive abilities (a so-called *competence structure*).

We start by repeating an instructive example that was given by Falmagne and Doignon ([8], Example 6.1.1):

Example 7. *The question which could be included in a test of proficiency in the UNIX operating system is*

*How many lines of the file **lilac** contain the word 'purple'? (Only one command line is allowed.)*

The skills *required for solving this question include the use of basic UNIX commands, such as*

$$S := \{\texttt{grep, wc, cat, more}, |, \ldots\}.$$

There are several possible correct solutions, each one usually involving several of the basic commands, for example $\{\texttt{cat, grep}, |, \texttt{wc}\}$.

Example 7 suggests the following general approach: it is assumed that the ability to answer questions depends on the presence of certain *skills*, more precisely, on combinations of skills. Competence based Knowledge Space Theory (CbKST) is an extension of the KST approach explaining a learner's knowledge state by the competencies and skills that learner has. It is assumed that for each question there are one or several competencies that enable mastering the question, and that each of these competencies is a combination of skills. More precisely, one assumes for a given set $\mathcal{Q}$ of questions a set $\mathcal{C}$ of **competencies**, each of which is a subset of the set $\mathcal{S}$ of **skills**. For each question $q \in \mathcal{Q}$ there is a set $\mathcal{C}_q \subseteq \mathcal{C}$ of those competencies that suffice for mastering q. In FCA language one might write this in the form of a formal context

$$(\mathcal{C}, \mathcal{Q}, \models),$$

where the $\models$-relation expresses that a competency suffices for mastering a question:

$$C \models q :\Leftrightarrow \text{ competency } C \text{ suffices for mastering question } q.$$

It is assumed that each learner has some, but not necessary all of these competencies. Abbreviating by $l \circ C$ the fact that learner l has competency C, we obtain a learner-competency context

$$(\mathcal{L}, \mathcal{C}, \circ).$$

The crucial model assumption then is that *a learner masters a question if and only if (s)he has a competency that suffices for mastering the question.* This can formally be expressed as

$$l \,\square\, q \Leftrightarrow \exists_{C \in \mathcal{C}} \ l \circ C \text{ and } C \models q,$$

which mathematically can be understood as a **Boolean product**

$$(\mathcal{L}, \mathcal{Q}, \Box) = (\mathcal{L}, \mathcal{C}, \circ) \cdot (\mathcal{C}, \mathcal{Q}, \models)$$

of formal contexts. Of course, the three formal contexts in this equation are usually not given, and it is a goal of *Learning Analytics* to provide reasonable models. Doignon and Falmagne write that their example (Example 7 above)

> ...may lead one to believe that the skills associated with a particular domain could always be identified easily. In fact, it is by no means obvious how such an identification might proceed in general.

Ganter and Glodeanu [10] have suggested that FCA could contribute to this important problem, and have made first steps into this direction.

Things get even more complicated when skills come into play. As said above, each competency corresponds to a skill combination, so that, sloppily spoken, the ability of mastering a question corresponds to a disjunction of conjunctions of skills. However, for reasons of space we shall not discuss the role of skills here in detail. Instead, we sketch the special case that skills and competencies coincide.

7   The Disjunctive Skill Model

In the disjunctive model each competency consists of a single skill, so that competencies and skills essentially may be identified. A disjunctive model thus is essentially described by the formal context

$$(\mathcal{C}, \mathcal{Q}, \models),$$

which is assumed to fulfill the Boolean matrix equation

$$(\mathcal{L}, \mathcal{Q}, \Box) = (\mathcal{L}, \mathcal{C}, \circ) \cdot (\mathcal{C}, \mathcal{Q}, \models).$$

A first research question is to identify competencies for given learner-task data $(\mathcal{L}, \mathcal{Q}, \Box)$. Is it possible to determine the minimum size of an appropriate set $\mathcal{C}$ of competencies? This question has two different answers.

Recall that the knowledge space $\mathcal{K}$ corresponding to $(\mathcal{L}, \mathcal{Q}, \Box)$ is generated by the object intents:

$$\mathcal{K} = \{\bigcup_{l \in L} l^{\Box} \mid L \subseteq \mathcal{L}\}.$$

Falmagne and Doignon [8] request that this knowledge space $\mathcal{K}$ is *delineated* by $(\mathcal{C}, \mathcal{Q}, \models)$, meaning that the knowledge states are precisely the unions of object intents. Written in terms of Boolean products, this reads as

$$\mathcal{K} = \{t \cdot (\mathcal{C}, \mathcal{Q}, \models) \mid t \in 2^{\mathcal{C}}\}.$$

More comprehensively this means that to every combination $T \subseteq \mathcal{C}$ of competencies, the set of questions that can be mastered having these competencies is a knowledge state. This set is

$$\bigcap_{C \in T} C^{\models},$$

which corresponds to

$$t \cdot (\mathcal{C}, \mathcal{Q}, \models),$$

where t is the Boolean vector for T.

Falmagne and Doignon show ([8], Theorem 6.3.8) that every finite knowledge space is delineated by some *minimal* skill map, which is unique up to isomorphism. In particular, they show that the minimal possible number of competencies equals the size of the basis of the knowledge space, which in turn is equal to the number of attributes of the complementary context of $(\mathcal{L}, \mathcal{Q}, \square)$. As an example, consider the knowledge space in Fig. 2, which corresponds to the learner-task data in Fig. 1. The unique minimal skill map is given, in form of a formal context, in Fig. 3, with the learners l_2, l_3, l_4, l_5, l_7 representing the competencies. Learner l_6 in Fig. 1 has the combined competencies of the learners l_4, l_5, and l_7, and so on.

The condition that $(\mathcal{C}, \mathcal{Q}, \models)$ delineates the knowledge space is, however, quite restrictive. It assumes that *every* combination of competencies may occur. In practice one may expect that some competencies are more advanced than others, so that the set of admissible competency combinations itself is structured, similarly to the questions. With this assumption, solutions with fewer competencies are possible, and uniqueness is lost. The formal problem statement under these conditions is to solve the Boolean matrix equation

$$(\mathcal{L}, \mathcal{Q}, \square) = (\mathcal{L}, \mathcal{C}, \circ) \cdot (\mathcal{C}, \mathcal{Q}, \models)$$

for given $(\mathcal{L}, \mathcal{Q}, \square)$, but unknown relations $\circ$ and $\models$, with a smallest possible set $\mathcal{C}$.

Again, the solution is known, but it may be hard to determine. It turns out that the smallest possible number of competencies equals the *2-dimension* of the knowledge space, which is the number of atoms of the smallest Boolean algebra that contains an order-embedding of the knowledge space. Bělohlávek and Vychodil [3] observed that for every competency C in a solution the set $C^{\circ} \times C^{\models}$ must be a "rectangular" subset of the relation $\square \subseteq \mathcal{L} \times \mathcal{Q}$, and that a solution is obtained if and only if these rectangles cover the relation, i.e., iff

$$\square = \bigcup_{C \in \mathcal{C}} C^{\circ} \times C^{\models}.$$

There is a more intuitive formulation, using the fact that every rectangle is contained in a maximal one. *The smallest possible number of competencies for a given formal context equals the smallest number of formal concepts which together cover all incidences of the incidence relation.* According to [12], Theorem 47,

this number equals the 2-dimension of the complementary context, i.e., the 2-dimension of the knowledge space. The problem of determining, for given $\mathcal{K}$ and n, if $\dim_2(\mathcal{K}) \leq n$, is known to be $\mathcal{NP}$-complete.

8   Skills First

Explaining a knowledge space by a small number of skills is only one of several interesting challenges. Another task is finding a set of questions for testing given skills. Here we can make use of a result by R. Suck, which also has an interpretation for concept lattices.

Theorem 1 (Suck [20]). *Given an ordered set* $(S, \leq)$, *a finite set* $\mathcal{Q}$, *and an order embedding* $\varphi : (S, \leq) \to (\mathfrak{P}(\mathcal{Q}), \subseteq)$, *the set*

$$\{\varphi(s) \mid s \in S\}$$

is the basis of a knowledge space if and only if φ *satisfies*

$$|\varphi(s)| > \left| \bigcup_{t < s} \varphi(t) \right|. \tag{1}$$

Corollary 1. *Given an ordered set* $(G, \leq)$, *a finite set* M, *and an order embedding* $\varphi : (G, \leq) \to (\mathfrak{P}(M), \subseteq)$, *the formal context* (G, M, I) *defined by*

$$g \, I \, m :\Leftrightarrow m \notin \varphi(g)$$

is object-reduced if and only if φ *satisfies Condition (1) of Theorem 1.*

We prove the corollary:

Proof. By definition of the incidence relation we have $g' = M \setminus \varphi(g)$ for all $g \in G$. An object $g \in G$ could be reducible only if there was some subset $X \subseteq G$, $g \notin X$, with

$$g' = \bigcap_{x \in X} x', \tag{2}$$

which implies $g' \subset x'$ and therefore $\varphi(x) \subset \varphi(g)$ for all $x \in X$. Since φ is an order embedding, we conclude $x < g$ for all $x \in X$. Equation (2) transforms to

$$M \setminus \varphi(g) = \bigcap_{x \in X} (M \setminus \varphi(x)) = M \setminus \bigcup_{x \in X} \varphi(x),$$

which implies

$$\varphi(g) = \bigcup_{x \in X} \varphi(x).$$

But that contradicts Condition (1).

Suck's theorem provides us with an easy method for constructing formal contexts with a preset order of the object concepts. One way to achieve this is given by the contra-ordinal scale $(S, S, \not\geq)$ (see Fig. 6 for an example). In that case, the concept lattice is distributive and Condition (1) is fulfilled in the most parsimonious way:

$$|\varphi(s)| = \left| \bigcup_{t < s} \varphi(t) \right| + 1. \tag{3}$$

Theorem 4 in Suck [19] shows that this condition is characteristic for the well-graded knowledge spaces which were mentioned in Sect. 2.

Theorem 1 gives an answer to the following questions: Suppose we start with a set S of skills (or competencies, since we are in the disjunctive case), together with an order relation $\leq$ on S, where $s_1 \leq s_2$ expresses that *learners with skill s_2 master all questions which learners with skill s_1 master*. How can a set $\mathcal{Q}$ of questions be designed that tests all these skills?

The answer given by Theorem 1 is that for every skill $s \in S$ a set $\varphi(s)$ of questions must be given in such a way, that two conditions are fulfilled: φ must be an order embedding, and Condition (1) must be satisfied.

9   Conclusions and Outlook

Put simply, the main aim of this paper was to describe Knowledge Space Theory in the terminology of Formal Concept Analysis. Apart from theoretical insights on mathematical overlaps and differences between both theories, the main advantage of such a "translation" is that it may facilitate further developments of applications. We see possibilities in the field of Technology-Enhanced Learning (TEL):

As mentioned above, knowledge- and competence spaces can become very large. This may be problematic for applications which make use of (Cb)KST for learner modelling purposes, due to the required computational power. Time consuming tasks are the adaptive assessment of the learner's knowledge state or personalized recommendations of suitable learning paths. As an on-vogue example, Game-based Learning applications need to be smoothly running, because otherwise the envisaged flow experience of the learners might diminish. This requires ongoing updates of the learner model in real time, based on the learner's actions and interactions with the TEL-application. The FCA community developed a bundle of remarkably fast algorithms which could be exploited.

But also more "traditional" TEL applications could benefit from applying FCA to learner modelling. The recently finished European research project LEA's BOX (**Le**arning **A**nalytics Tool**box**, see http://leas-box.eu/), aimed at providing a Web platform for teachers and students for activity tracking, domain modelling, student modelling as well as visualization of educational data. (Cb)KST and FCA have been applied for domain and student modelling.

Visualizations of concept lattices resulting from learner-task data as well as learner-competence matrices have been used by teachers to get an overview of

the strengths and weaknesses of the whole classroom. After a short training on how to interpret and "read" the concept lattices, teachers were able to make use of interactive lattices by getting insights on several pedagogically relevant questions (Bedek et al. [2]). Consider a lattice similar to the one shown in Fig. 1. Pedagogically relevant questions which can be answered by reading the lattice are for example: "What are the best and weakest students?", "which students could be tutors for others in a collaborative, peer-learning scenario?", or "how coherent is the classroom performance" (As a "rule of thumb", a concept lattice with a large number of formal concepts is an indication for a high diversity among the students' performance- and competence states). Finally, in some cases it might be of great interest for a teacher to observe the learning progress over a longer period of time by comparing the lattices generated at different points in time with the empirical learner-competence matrices. In the perfect case, all students should finally end up (e.g. at the end of the semester) with the knowledge state $\mathcal{Q}$.

The attempts of the LEA's BOX project to introduce such kind of visualizations into the classroom are just the starting point. So far, rather small learner-competence matrices have been used (i.e. around 30 students and 10 competencies). Further research is required to provide more intuitive and user-friendly interactive visualizations—even with larger lattices. In addition to that, empirical investigations on how such interactive visualizations are used in the classroom by teachers and students should and will be carried out in the near future.

References

1. Albert, D., Held, T.: Establishing knowledge spaces by systematical problem construction. In: Albert, D. (ed.) Knowledge Structures, pp. 78–112. Springer, New York (1994)
2. Bedek, M.A., Kickmeier-Rust, M.D., Albert, D.: Formal concept analysis for modelling students in a technology-enhanced learning setting. In: ARTEL@ EC-TEL, pp. 27–33 (2015)
3. Bělohlávek, R., Vychodil, V.: Discovery of optimal factors in binary data via a novel method of matrix decomposition. J. Comput. Syst. Sci. **76**(1), 3–20 (2010)
4. Colomb, P., Irlande, A., Raynaud, O.: Counting of moore families for n=7. In: Kwuida, L., Sertkaya, B. (eds.) ICFCA 2010. LNCS, vol. 5986, pp. 72–87. Springer, Heidelberg (2010). doi:10.1007/978-3-642-11928-6_6
5. Doignon, J.-P.: Knowledge spaces and skill assignments. In: Fischer, G.H., Laming, D. (eds.) Contributions to Mathematical Psychology, Psychometrics, and Methodology, pp. 111–121. Springer, Heidelberg (1994)
6. Doignon, J.-P., Falmagne, J.-C.: Knowledge Spaces. Springer, Heidelberg (1999)

7. Düntsch, I., Gediga, G.: Skills and knowledge structures. Br. J. Math. Stat. Psychol. **48**(1), 9–27 (1995)
8. Falmagne, J.-C., Doignon, J.-P.: Learning Spaces. Springer, Heidelberg (2011)
9. Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P., Johannesen, L.: Introduction to knowledge spaces: how to build, test, and search them. Psychol. Rev. **97**(2), 201 (1990)
10. Ganter, B., Glodeanu, C.V.: Factors and skills. In: Glodeanu, C.V., Kaytoue, M., Sacarea, C. (eds.) ICFCA 2014. LNCS, vol. 8478, pp. 173–187. Springer, Cham (2014). doi:10.1007/978-3-319-07248-7_13
11. Ganter, B., Obiedkov, S.: Conceptual Exploration. Springer, Heidelberg (2016)
12. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, Heidelberg (1999)
13. Gediga, G., Düntsch, I.: Skill set analysis in knowledge structures. Br. J. Math. Stat. Psychol. **55**(2), 361–384 (2002)
14. Heller, J., Wickelmaier, F.: Minimum discrepancy estimation in probabilistic knowledge structures. Electron. Notes Discrete Math. **42**, 49–56 (2013)
15. Korossy, K.: Modellierung von Wissen als Kompetenz und Performanz. Unpublished doctoral dissertation, see [15]. Universität Heidelberg, Germany (1993)
16. Korossy, K.: Extending the theory of knowledge spaces: a competence-performance approach. Zeitschrift für Psychol. **205**(1), 53–82 (1997)
17. Korossy, K., et al.: Modeling knowledge as competence and performance. In: Albert, D., Lukas, J. (eds.) Knowledge Spaces: Theories, Empirical Research, and Applications, pp. 103–132. Lawrence Erlbaum Associates, Mahwah (1999)
18. Rusch, A., Wille, R.: Knowledge spaces and formal concept analysis. In: Bock, H.H., Polasek, W. (eds.) Data Analysis and Information Systems: Statistical and Conceptual Approaches, pp. 427–436. Springer, Heidelberg (1996)
19. Suck, R.: Parsimonious set representations of orders, a generalization of the interval order concept, and knowledge spaces. Discrete Appl. Math. **127**(2), 373–386 (2003)
20. Suck, R.: Knowledge spaces regarded as set representations of skill structures. In: Dzhafarov, E., Perry, L. (eds.) Descriptive and Normative Approaches to Human Behavior, pp. 249–270. World Scientific, Singapore (2012)
21. Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival, I. (ed.) Ordered Sets, pp. 445–470. Reidel, Dordrecht-Boston (1982)