# Factors and Skills

Bernhard Ganter and Cynthia Vera Glodeanu

Technische Universität Dresden,
01062 Dresden, Germany
{Bernhard.Ganter,Cynthia-Vera.Glodeanu}@tu-dresden.de

**Abstract.** Inspired by Knowledge and Learning Spaces, we present a novel framework for explaining the answering patterns of learners through competences and skills. More precisely, we investigate how a given learner-question data may be ascribed by a set of competences such that a learner masters a question if and only if they have a competence that is sufficient for mastering the question. Each competence is some combination of skills, but there may be restrictions on which skills can be combined. In general a question does not require a unique competence.

**Keywords:** Skills, Competences, Knowledge Spaces, Formal Concept Analysis, Learner-question data, Boolean factorisation.

## 1 Introduction

The theory of *Knowledge Spaces*, as it was introduced by Doignon and Falmagne [1], is closely related to Formal Concept Analysis. Several extensions have been studied, among them the "Competence based Knowledge Space Theory" (CbKST) [2], and, more recently, the theory of *Learning Spaces* [3]. Here we present and extend some ideas from CbKST, using the language of Formal Concept Analysis. We illustrate the basic definitions and results by a small example. Random effects, though important, will not be considered in this basic version.

## 2 Competences and Factors

We consider a formal context $(L, Q, \Box)$ with the following intended interpretation: The elements of $L$ are called **learners**, those of $Q$ are the **questions**, and $l \Box q$ expresses that learner $l$ **masters** question $q$. In the jargon of Formal Concept Analysis the set of questions mastered by learner $l$ then is denoted by $l^{\Box}$, and $q^{\Box}$ is a shorthand notation for the set of learners who master question $q$.

This interpretation should be understood in a very general manner: $Q$ might, for example, be a set of diseases, $L$ a set of therapies and $l \Box q$ indicates that therapy $l$ heals disease $q$. Or $L$ is a set of customers, $Q$ a set of products and $l \Box q$ indicates that product $q$ is a possible choice for customer $l$, et cetera.

We investigate how $(L, Q, \Box)$ may be explained by a set $\mathcal{C}$ of **competences** in such a way that a learner masters a question $q$ if and only if they have a competence that is sufficient for mastering $q$.

C.V. Glodeanu, M. Kaytoue, and C. Sacarea (Eds.): ICFCA 2014, LNAI 8478, pp. 173–187, 2014.

This leads to the well known problem of finding **Boolean factorisations** [4,5] of $(L, Q, \Box)$. Required for such a factorisation are formal contexts $(L, \mathcal{C}, \circ)$ and $(\mathcal{C}, Q, \models)$ such that

$$l \Box q \iff \exists_{C \in \mathcal{C}} \; (l \circ C \text{ and } C \models q),$$

which is symbolised by

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

Of course, $l \circ C$ is interpreted as "learner $l$ has competence $C$" and $C \models q$ reads as "competence $C$ suffices for mastering question $q$".

It is well understood how this problem must be attacked. The factorisations are in 1-1-correspondence with the coverings of the relation $\Box$ by rectangular subrelations. Their smallest number equals the so-called 2-dimension (see [6] for the definition of $k$-dimension for arbitrary integer $k$) of the complementary context $(L, Q, L \times Q \setminus \Box)$. Determining this dimension is known to be $\mathcal{NP}$-complete. Alternatively, one can show that the factorisation problem is hard by reducing the set basis problem to it, see [5].

There is another approach to Boolean factors which is perhaps more intuitive. For a given formal context one may ask if its attributes can be interpreted as disjunctions of attributes of an other, hopefully simpler context. More formally, let us say that a **disjunctive attribute representation** of $(G, M, I)$ over $(G, N, J)$ is a mapping $\delta : M \to \mathfrak{P}(N)$ such that

$$g \; I \; m \iff \exists_{n \in \delta(m)} \; g \; J \; n.$$

The existence of such an attribute representation leads to a factorisation

$$(G, M, I) = (G, N, J) \cdot (N, M, K) \qquad \text{with } n \; K \; m :\iff n \in \delta(m).$$

Conversely any such factorisation leads to a disjunctive attribute representation via $\delta(m) := m^K$ for all $m \in M$.

*Example 1.* The data that we use is from Korossy [2]. It describes how eleven learners performed for a set $Q := \{a, b, c, d, e, f\}$ of six questions. Only seven distinct answering patterns occurred. These are given in Figure 1.

The concept lattice of the complementary relation (the diagram on the right of Figure 1) contains the information about the possible Boolean factorisations. Its length is five, which gives a lower bound for the 2-dimension (and thereby for the number of competences). But the dimension cannot be larger since there are only five join-irreducible elements. Therefore the incidence relation $\Box$ can be covered by five "rectangles", but not by fewer than five. An example of a covering is

| $\square$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
|---|---|---|---|---|---|---|
| $02L$ | × | × | × | | × | × |
| $03L$ | × | × | × | × | × | × |
| $05L$ | × | × | × | × | × | |
| $08L$ | × | | × | | | |
| $11L$ | × | × | × | | × | |
| $13L$ | × | | × | | × | × |
| $20L$ | × | × | × | | | |

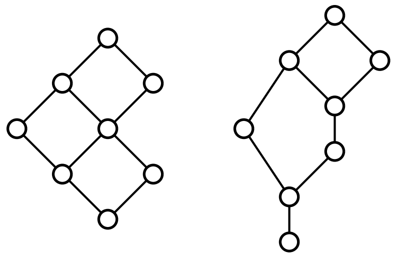

**Fig. 1.** A formal context of learners and questions, and its concept lattice, and the concept lattice of its complementary context (unlabeled)

$$
\begin{aligned}
C_1 &:= \{02L, 03L, 05L, 08L, 11L, 13L, 20L\} \times \{a, c\},\\
C_2 &:= \{02L, 03L, 05L, 11L, 20L\} \times \{a, b, c\},\\
C_3 &:= \{03L, 05L\} \times \{a, b, c, d, e\},\\
C_4 &:= \{02L, 03L, 05L, 11L\} \times \{a, b, c, e\},\\
C_5 &:= \{02L, 03L, 13L\} \times \{a, c, e, f\}.
\end{aligned}
$$

Taking these factors as competences, we get a factorisation of the context in Figure 1 as shown in Figure 2.

| $\circ$ | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ |
|---|---|---|---|---|---|
| $02L$ | × | × | | × | × |
| $03L$ | × | × | × | × | × |
| $05L$ | × | × | × | × | |
| $08L$ | × | | | | |
| $11L$ | × | × | | × | |
| $13L$ | × | | | | × |
| $20L$ | × | × | | | |

$\cdot$

| $\models$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
|---|---|---|---|---|---|---|
| $C_1$ | × | | × | | | |
| $C_2$ | × | × | × | | | |
| $C_3$ | × | × | × | × | × | |
| $C_4$ | × | × | × | | × | |
| $C_5$ | × | | × | | × | × |

**Fig. 2.** A factorisation of the context in Figure 1

The concept lattice of the first factorising context is shown in Figure 3.

In view of a desired interpretation, a result like the one presented in Figure 2 may be somewhat disappointing, because it only produces an (ordered) set of abstract "competences" without further explanation. Moreover, the covering with rectangular subrelations is by no means unique. In the above example, we might take as rectangles the columns of the original context, combining columns $a$ and $c$ to one rectangle, and obtain a different factorisation.

In a second step therefore one can investigate competences which comply with a given theoretical **competence model**.

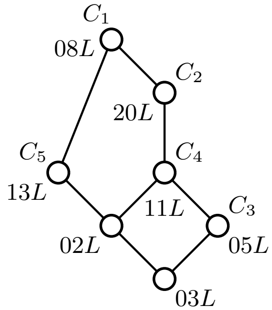


**Fig. 3.** The concept lattice of the first factorising context in Figure 2

Such a model can be given as a formal context $(S, T, *)$, where $S$ is a set of competence "states" which a learner may or may not have, $T$ is a set of competences and $s * t$ indicates that in state $s$ competence $t$ is present.

The basic question then is if the observed learner-question data can be explained by competences from this abstract model. For this, we must associate to every question $q$ the set of those competences from $T$ that are sufficient for mastering the question. Simultaneously, for each learner a suitable competence state from $S$ has to be found that enables the learner to master the questions as observed. A more formal version is given in the following theorem.

**Theorem 1.** *Let formal contexts* $(L, Q, \Box)$ *and* $(S, T, *)$ *be given. Then for every mapping* $\alpha : L \to S$ *the following are equivalent:*

1. *There is a mapping* $\sigma : Q \to \mathfrak{P}(T)$ *such that*

$$l \,\Box\, q \iff \exists_{C \in \sigma(q)}\, \alpha(l) * C.$$

2. *There is a Boolean factorisation* $(L, Q, \Box) = (L, \mathcal{C}, \circ) \cdot (\mathcal{C}, Q, \models)$ *together with a mapping* $\beta : \mathcal{C} \to T$ *such that*

$$l \circ C \iff \alpha(l) * \beta(C).$$

*Proof.* Assuming (1) we let $\mathcal{C} := \bigcup_{q \in Q} \sigma(q)$ and define for $l \in L$, $C \in \mathcal{C}$, and $q \in Q$

$$l \circ C :\iff \alpha(l) * C, \quad \beta := \mathrm{id}, \quad \text{and} \quad C \models q :\iff C \in \sigma(q).$$

The conditions of (2) are now easily verified. Conversely when starting from (2) we get (1) by letting

$$\sigma(q) := \{\beta(C) \mid C \models q\}.$$

## 3 Skills and Competences

Several authors (e.g. Korossy [2], Doignon [7]) have investigated if such competences may be explained by a finite set $S$ of **skills**, which learners may have. For this they ask for a **skill function**[1], a mapping[2]

$$\sigma : Q \to \mathfrak{P}(\mathfrak{P}(S))$$

with the property that $\sigma(q)$ is an antichain for each $q \in Q$. It is assumed that the learners have certain skills, as expressed by the **skill context** $(L, S, \bullet)$. The elements of

$$\mathcal{C} := \bigcup_{q \in Q} \sigma(q)$$

then play the role of the competences. They are ordered by set inclusion $\subseteq$. The interpretation is that a learner masters a question if they have the necessary skills, more precisely that

$$l \,\Box\, q \iff \exists_{C \in \sigma(q)}\, C \subseteq l^{\bullet}.$$

The context $(\mathcal{C}, Q, \models)$ is then given by

$$C \models q \iff \exists_{D \in \sigma(q)}\, D \subseteq C.$$

The above mentioned competence model in this case is $(\mathfrak{P}(S), \mathcal{C}, \subseteq)$.

Each skill function $\sigma$ defines a mapping $p_\sigma : \mathfrak{P}(S) \to \mathfrak{P}(Q)$, called the **problem function**, by

$$p_\sigma(T) := \{q \in Q \mid \exists_{C \in \sigma(q)}\, C \subseteq T\}, \qquad T \subseteq S,$$

assigning to each set $T \subseteq S$ the set of problems which can be answered with the skills in $T$. Equivalent to the above condition is that for each learner $l \in L$ it holds that

$$l^{\Box} = p_\sigma(l^{\bullet}),$$

meaning that each learner masters exactly those questions for which they have the necessary skills.

Problem functions are order preserving maps from $(\mathfrak{P}(S), \subseteq)$ to $(\mathfrak{P}(Q), \subseteq)$, and indeed, as Düntsch and Gediga [8] have shown, every order preserving function can be obtained in this way from a unique skill function.

*Example 2.* Continuing the above example we ask how the context in Figure 1 may be explained by skills. It is easier to tackle this problem with respect to a given factorisation. Consider the first factorising context $(L, \mathcal{C}, \circ)$ in Figure 2. It displays which competences the individual learners have. In order to express these competences by subsets of a (yet unknown) set $S$ of "skills" we have to find mappings

$$\alpha : L \to \mathfrak{P}(S) \quad \text{and} \quad \beta : \mathcal{C} \to \mathfrak{P}(S)$$

[1] Skill multiassignment in [7], skill multimap in [3].

[2] We omit some technical conditions which are not necessary for our considerations.

such that a learner $l$ has a competence $C$ if and only if they have all the skills contained in $\beta(C)$, formally

$$l \circ C \iff \alpha(l) \supseteq \beta(C).$$

It is immediate from Proposition 33 in [6] that such mappings can be found if and only if there is an order embedding of $\underline{\mathfrak{B}}(L, \mathcal{C}, \circ)$ into $(\mathfrak{P}(S), \supseteq)$. This is in turn equivalent to the condition that the 2-dimension of $\underline{\mathfrak{B}}(L, \mathcal{C}, \circ)$ is at most the size of $S$, i.e., to

$$\text{fdim}_2(L, \mathcal{C}, \circ) \leq |S|.$$

The 2-dimension of the lattice in Figure 3 obviously is four, and Figure 4 shows an order embedding into the dual of the power set of $S := \{x, y, z, t\}$.

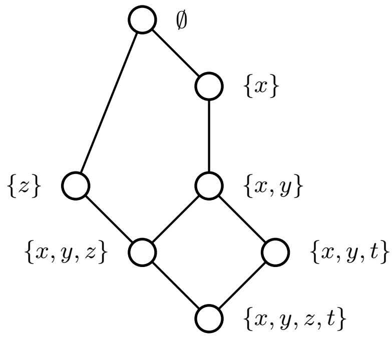


**Fig. 4.** The concept lattice of the first factorising context in Figure 2, embedded into $(\mathfrak{P}(\{x, y, z, t\}), \supseteq)$

A comparison of the labellings in Figures 3 and 4 discloses the skills associated to the learners and to the competences.

| $\bullet$ | $x$ | $y$ | $z$ | $t$ |
|---|---|---|---|---|
| 02$L$ | × | × | × | |
| 03$L$ | × | × | × | × |
| 05$L$ | × | × | | × |
| 08$L$ | | | | |
| 11$L$ | × | × | | |
| 13$L$ | | | × | |
| 20$L$ | × | | | |

$C_1 = \varnothing$
$C_2 = \{x\}$
$C_3 = \{x, y, t\}$
$C_4 = \{x, y\}$
$C_5 = \{z\}.$

**Fig. 5.** The learner-skill context and the competences as sets of skills

The second factorising context in Figure 2 can now be understood as a skill function (Figure 6), however with a slight modification: The attribute intent of each question $q \in \{a, \dots, f\}$ consists of *all* competences which suffice for mastering the question, not only the minimal ones. We call this an *enriched* skill function. Meagering it for each question to the minimal sufficient competences results in a skill function (Figure 7).

| $\models$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
|---|---|---|---|---|---|---|
| $C_1 = \varnothing$ | × | | × | | | |
| $C_2 = \{x\}$ | × | × | × | | | |
| $C_3 = \{x, y, t\}$ | × | × | × | × | × | |
| $C_4 = \{x, y\}$ | × | × | × | | × | |
| $C_5 = \{z\}$ | × | | × | | × | × |

**Fig. 6.** The attribute intents define an enriched skill function

| $q$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
|---|---|---|---|---|---|---|
| $\sigma(q)$ | $\{\varnothing\}$ | $\{\{x\}\}$ | $\{\varnothing\}$ | $\{\{x, y, t\}\}$ | $\{\{x, y\}, \{z\}\}$ | $\{\{z\}\}$ |

**Fig. 7.** The derived skill function

It can now easily be verified that the skill function in Figure 7 together with the learner-skill context in Figure 5 result in the original learner-question data shown in Figure 1.

We summarise our findings in a theorem. This theorem, as well as the next one, may look a little technical, but their content is easy. The first one says, loosely spoken: Given learner-question data, pick a Boolean factorisation and a representation of the first factorising context by sets. Then a skill function is obtained representing the given data.

**Theorem 2.** *Let*

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

*be a Boolean factorisation and let* $\alpha : L \to \mathfrak{P}(S)$ *and* $\beta : \mathcal{C} \to \mathfrak{P}(S)$*, where* $S$ *is a finite set, be mappings such that*

$$l \circ C \iff \alpha(l) \supseteq \beta(C) \qquad \textit{(for all } l \in L, C \in \mathcal{C}\textit{)}.$$

*Then the mapping* $\sigma : Q \to \mathfrak{P}(\mathfrak{P}(S))$*, defined by*

$$\sigma(q) := \{\beta(C) \mid \beta(C) \textit{ is minimal wrt. } C \models q\},$$

*is a skill function such that*

$$l \,\square\, q \iff \exists_{D \in \sigma(q)} D \subseteq \alpha(l).$$

*Proof.* Because of the minimality condition it is clear that $\sigma$ is a skill function. Since we have a Boolean factorisation we get for $l \in L$ and $q \in Q$

$$\begin{aligned} l \,\square\, q &\iff \exists_{C \in \mathcal{C}}\; l \circ C \text{ and } C \models q \\ &\iff \exists_{C \in \mathcal{C}}\; \alpha(l) \supseteq \beta(C) \text{ and } C \models q \\ &\iff \exists_{D \in \sigma(q)}\; \alpha(l) \supseteq D. \end{aligned}$$

The existence of such a set $D$ follows from the finiteness of $S$.

## 4 From Skills to Factors

In the previous section we have demonstrated how a skill function can be constructed from learner-question data using a two-stage set representation process. It is however not yet obvious that this method always works and, if so, that it leads to a small number of skills.

The latter is indeed not always true. The number of required skills depends on the choice of the Boolean factorisation. In fact, the data of the example can be represented by fewer skills, as we shall show.

Nevertheless is the method general enough to cover all possibilities. Each skill function can be reconstructed, as we shall demonstrate in the next theorem. Informally, it says that when the construction described in Theorem 2 is applied to learner-question data which is based on a skill function, the factorisation and the embedding can be chosen so that this skill function is reconstructed.

**Theorem 3.** *Let finite sets $L$, $Q$, and $S$ (of "learners", "questions", and "skills", respectively) be given together with a mapping $\sigma : Q \to \mathfrak{P}(\mathfrak{P}(S))$ that maps questions to antichains of skill sets (i.e., a skill function) and a mapping $\alpha : L \to \mathfrak{P}(S)$ that assigns to each learner a set of skills. Then for the relation $\square \subseteq L \times Q$, defined by*

$$l \,\square\, q :\iff \exists_{C \in \sigma(q)}\; C \subseteq \alpha(l)$$

*there is a Boolean factorisation $(L, Q, \square) = (L, \mathcal{C}, \circ) \cdot (\mathcal{C}, Q \models)$ and a bijection $\beta : \mathcal{C} \to \bigcup_{q \in Q} \sigma(q)$, such that*

$$l \circ C \iff \alpha(l) \supseteq \beta(C) \qquad \text{and} \qquad C \models q \iff \exists_{D \in \sigma(q)}\; D \subseteq \beta(C).$$

*In particular,*

$$\sigma(q) = \{\beta(C) \mid \beta(C) \text{ is minimal wrt. } C \models q\} \quad \text{for each } q \in Q.$$

*Proof.* Let $\mathcal{C} := \bigcup_{q \in Q} \sigma(q)$ and $\beta :=$ id. Then

$$\begin{aligned} l \,\square\, q &\iff \exists_{C \in \sigma(q)}\; C \subseteq \alpha(l) \\ &\iff \exists_{C \in \sigma(q)}\; \exists_{D \in \mathcal{C}}\; \beta(D) = C \subseteq \alpha(l) \\ &\iff \exists_{C \in \sigma(q)}\; \exists_{D \in \mathcal{C}}\; C \subseteq \beta(D) \subseteq \alpha(l) \\ &\iff \exists_{D \in \mathcal{C}}\; l \circ D \text{ and } D \models q. \end{aligned}$$

It remains to show that

$$\sigma(q) = \{\beta(C) \mid \beta(C) \text{ is minimal wrt. } C \models q\} \quad \text{for each } q \in Q.$$

To this end let $q \in Q$ and $C \in \mathcal{C}$. We show the two inclusions:
"$\supseteq$" Let $\beta(C)$ be minimal in $\{\beta(C) \mid C \models q\}$. Then, there is $D \in \sigma(q)$ s.t. $D \subseteq \beta(C)$. Since $\beta(C)$ was chosen minimal wrt. $C \models q$, we have $D = \beta(C)$ and thus $\beta(C) \in \sigma(q)$.
"$\subseteq$" Let $D \in \sigma(q)$. There exists $C \in \mathcal{C}$ s.t. $\beta(C) = D$. Hence, $C \models q$. It remains to show that $\beta(C)$ is minimal in $\{\beta(E) \mid E \models q\}$ for $\beta(E) \subsetneq \beta(C)$. Suppose not. Then, there exists $F \in \sigma(q)$ s.t. $F \subseteq \beta(E) \subsetneq \beta(C) \subseteq D$. Thus, $F \subsetneq D$ yielding a contradiction since $\sigma$ is a skill function.

*Example 3.* The learner-question data in Figure 1 can be based on only three skills, as the following tables show. For the three-element skill set $S := \{u, v, w\}$ they define a skill function $\sigma : Q \to \mathfrak{P}(\mathfrak{P}(S))$ and a learner-skill assignment $\alpha : L \to \mathfrak{P}(S)$.

| $q$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
|---|---|---|---|---|---|---|
| $\sigma(q)$ | $\{\varnothing\}$ | $\{\{v\},\{w\}\}$ | $\{\varnothing\}$ | $\{\{v,w\}\}$ | $\{\{w\},\{u\}\}$ | $\{\{u\}\}$ |

| $l$ | $02L$ | $03L$ | $05L$ | $08L$ | $11L$ | $13L$ | $20L$ |
|---|---|---|---|---|---|---|---|
| $\alpha(l)$ | $\{u,v\}$ | $\{u,v,w\}$ | $\{v,w\}$ | $\varnothing$ | $\{w\}$ | $\{u\}$ | $\{v\}$ |

**Fig. 8.** This skill function leads to the learner-question data in Figure 1 if the learner-skill assignment is as given in the second table

The competences are $\mathcal{C} = \{\varnothing, \{u\}, \{v\}, \{w\}, \{v, w\}\}$, the corresponding factors

$$F_i := \{\text{learners that have } C_i\} \times \{\text{questions that are mastered by } C_i\}$$

are as follows:

$$F_1 = \{02L, 03L, 05L, 08L, 11L, 13L, 20L\} \times \{a, c\}$$
$$F_2 = \{02L, 03L, 05L, 20L\} \times \{a, b, c\}$$
$$F_3 = \{03L, 05L\} \times \{a, b, c, d, e\}$$
$$F_4 = \{03L, 05L, 11L\} \times \{a, b, c, e\}$$
$$F_5 = \{02L, 03L, 13L\} \times \{a, c, e, f\}.$$

These rectangular relations indeed cover the "masters"-relation. The corresponding Boolean factorisation of the learner-question context is shown in Figure 9. It differs only slightly from the one given in Figure 2.

However, the 7-element concept lattice of the first factorising context can easily embedded into $(\mathfrak{P}(\{u, v, w\}), \supseteq)$, as Figure 10 shows.

| $\circ$ | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ |
|---|---|---|---|---|---|
| 02$L$ | × | × | | | × |
| 03$L$ | × | × | × | × | × |
| 05$L$ | × | × | × | × | |
| 08$L$ | × | | | | |
| 11$L$ | × | | | × | |
| 13$L$ | × | | | | × |
| 20$L$ | × | × | | | |

.

| $\models$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ |
|---|---|---|---|---|---|---|
| $C_1$ | × | | × | | | |
| $C_2$ | × | × | × | | | |
| $C_3$ | × | × | × | × | × | |
| $C_4$ | × | × | × | | × | |
| $C_5$ | × | | × | | × | × |

**Fig. 9.** Another factorisation of the context in Figure 1. Here the first factorisation context has 2-dimension three.

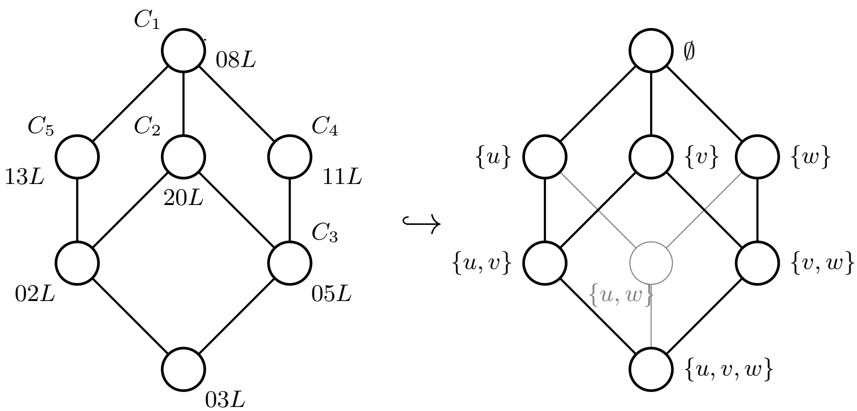


**Fig. 10.** The concept lattice of the first factorising context, embedded into $(\mathfrak{P}(\{u, v, w\}), \supseteq)$

| $\bullet$ | $u$ | $v$ | $w$ |
|---|---|---|---|
| 02$L$ | × | × | |
| 03$L$ | × | × | × |
| 05$L$ | | × | × |
| 08$L$ | | | |
| 11$L$ | | | × |
| 13$L$ | × | | |
| 20$L$ | | × | |

$C_1 = \varnothing$
$C_2 = \{v\}$
$C_3 = \{v, w\}$
$C_4 = \{w\}$
$C_5 = \{u\}$.

**Fig. 11.** The learner-skill context and the competences as sets of skills, for the modified set representation

# 5 Structured Skill Sets

Our approach admits several variations which may be of practical interest. We briefly discuss four of them here.

## 5.1 Graded Skills

The lattice in Figure 4 obviously has an order embedding into a product of two chains, one of size four, the other of size two. This allows to give a more structured interpretation of the four necessary skills: they may be chosen as $\{x, x^+, x^{++}, z\}$, where $x$ is a prerequisite for $x^+$, and $x^+$ a prerequisite for $x^{++}$. The five competences may then be written as

$$C_1 = \varnothing, \quad C_2 = \{x\}, \quad C_3 = \{x^{++}\}, \quad C_4 = \{x^+\}, \quad C_5 = \{z\},$$

with the tacit convention that $x^+$ includes $x$ etc.

This can widely be generalised. A family of competences can be interpreted with skills $\{x_1, x_1^+, x_1^{++}, \ldots, x_k, x_k^+, x_k^{++}\}$ if and only if the 4-dimension of the first factor is at most $k$. But even arbitrarily ordered skill sets can be considered and respective conditions on the factorisations can be formulated.

## 5.2 Propositional Formulae

We may even consider "negative skills". Recall that a skill function encodes that a question $q$ is mastered if and only if at least one competence, i.e., skill combination, from a specified list $\sigma(q)$ is present. So what is required for mastering $q$ is a disjunction of conjunctions of skills, a monotone Boolean term in the language of Propositional Logic.

So why not allow for arbitrary propositional formulae? This can easily be done. Figure 12 shows a representation of our original learner-question data (Figure 1) by propositional formulae in three variables.

But how can this be interpreted? It seems unrealistic that there may be skills which *hinder* a learner mastering a question. However, for other interpretations this may be meaningful. One such case is that of customers selecting goods according to their features. E.g., when buying bread, some customers may prefer one with caraway seeds, while for others this could be a impediment.

## 5.3 The Dichotomic Scale $\mathbb{D}_k$

In the next example we shall make use of the $k$-dimensional **dichotomic scale** $\mathbb{D}_k$, which is one of the standard scales in Formal Concept Analysis (see [6], Lex [9]). It is usually introduced as the $k$-fold semiproduct $\mathbb{D} ⧖ \mathbb{D} ⧖ \cdots ⧖ \mathbb{D}$ of the (one dimensional) dichotomic scale

$$\mathbb{D} := \begin{array}{|c|c|}\hline \cdot & \times \\ \hline \times & \cdot \\ \hline \end{array}\,.$$

| | $\top$ | $\neg(x \wedge (y \vee z))$ | $\top$ | $\neg y \wedge ((\neg x \wedge z) \vee (x \wedge \neg z))$ | $\neg(y \wedge (x \vee \neg z))$ | $\neg(y \vee (\neg x \wedge z))$ |
|---|---|---|---|---|---|---|
| $\bot\ \bot\ \bot$ | × | × | × | | × | × |
| $\top\ \bot\ \bot$ | × | × | × | × | × | × |
| $\bot\ \bot\ \top$ | × | × | × | × | × | |
| $\top\ \top\ \top$ | × | | × | | | |
| $\bot\ \top\ \top$ | × | × | × | | × | |
| $\top\ \bot\ \top$ | × | | × | | × | × |
| $\bot\ \top\ \bot$ | × | × | × | | | |

**Fig. 12.** Truth value assignments and propositional formulae for the context of Figure 1

For our purposes it is convenient to give another (yet equivalent) description based on a set $V := \{v_1, \ldots, v_k\}$ of symbols[3]. The scale $\mathbb{D}_k$ has $2^k$ objects, $2k$ attributes and $3^k + 1$ formal concepts. As objects we may take the set of all maps from $V$ to $\{+, -\}$. The set of attributes is $S := \{+v_1, \ldots, +v_k, -v_1, \ldots, -v_k\}$. An object $\nu : V \to \{+, -\}$ is incident with an attribute $+s$ (where $s \in V$) iff $\nu(s) = +$, and with the "negative" attribute $-s$ iff $\nu(s) = -$.

A subset of $S$ is called *feasible* if it does not contain a symbol $v$ both in its positive form $+v$ and in its negative form $-v$. The only concept intent that is not feasible is the set $S$, and the corresponding extent is $\emptyset$. Apart from this exception, the concept intents of the dichotomic scale are exactly the feasible subsets of $S$. The concept extent corresponding to a feasible set $T \subseteq S$ consist of those mappings $\nu : V \to \{+, -\}$ that satisfy the condition

$$\text{if } +v \in T \text{ then } \nu(v) = +, \qquad \text{and if } -v \in T \text{ then } \nu(v) = -.$$

The concept extents, apart from the smallest one, therefore can be identified with the partial mappings $\nu : V \to \{\bot, \top, ?\}$.

### 5.4 Incompatible Skills

The propositional approach in Subsection 5.2 is based on the *negation* of skills. In practice however it seems unlikely that a skill is the negation of another one. A more realistic assumption is that skills may be mutually exclusive, but not

[3] We avoid naming the elements of $V$ *variables*, because $-v$ is not the negation of $-v$. As a consequence, we later shall work with a modified notion of disjunction.

necessarily exhaustive. In other words: such two skills cannot occur together, but may both be missing. For example, good jockeys usually are not very good high jumpers, because jockeys need to be small, high jumpers to be tall. But most people neither are jumpers nor jockeys.

As in Subsection 5.3 we start with a set $V := \{v_1, \ldots, v_k\}$ of symbols and define $S := \{+v_1, \ldots, +v_k, -v_1, \ldots, -v_k\}$. The elements of $S$ will be the skills, with the intention that for each $i$ the skills $+v_i$ and $-v_i$ are mutually exclusive.

The competence model in this case is introduced as follows: All feasible sets of skills are competences, and all mappings from $V$ to $\{+, -, ?\}$ are possible learner states. The concept intents of the dichotomic scale then are in 1-1-correspondence to the competences, with one exception, which we artificially add: We allow for the set $S$ of *all* skills, though not admissible, as a competence, the *"Chuck Norris competence"*. Similarly, the possible learner states correspond to the concept extents of the dichotomic scale, when we artificially add the possibility of an "almighty" learner that has all skills, negative and positive. The incidence relation in the competence model is the natural one, the one that was discussed in the previous subsection.

Applying Theorem 1 in the case of this slightly artificial competence model yields the following:

**Corollary 1.** *A learner-question context can be interpreted using skills*

$$+v_1, \ldots, +v_k, -v_1, \ldots, -v_k$$

*(where $+v$ and $-v$ are incompatible), iff there is a Boolean factorisation and an order embedding of the concept lattice of the first factorising context into the concept lattice of the $k$-dimensional dichotomic scale.*

*Example 4.* Again we demonstrate this by an example. The concept lattice in Figure 10 (left) can also be embedded into the concept lattice of the 2-dimensional dichotomic scale $\mathbb{D}_2$, see Figure 13, in which we use symbols $x, y$ instead of $v_1, v_2$. Actually, there are several embeddings.

According to the corollary, the learner-question data can be interpreted using two pairs $+x, -x, +y, -y$ of incompatible skills. The competences are mapped to feasible skill sets as follows:

| $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ |
|---|---|---|---|---|
| $\emptyset$ | $\{+x\}$ | $\{+x, +y\}$ | $\{+y\}$ | $\{-y\}$ |

The observed learner states are the following:

| $02L$ | $03L$ | $05L$ | $08L$ | $11L$ | $13L$ | $20L$ |
|---|---|---|---|---|---|---|
| $\{+x, -y\}$ | almighty | $\{+x, +y\}$ | $\emptyset$ | $\{+y\}$ | $\{-y\}$ | $\{+x\}$ |

We can also give a skill function based on these skills. It is tempting to do this in propositional form, similar as in Figure 12. However the meaning of disjunction has to be modified, the expression $+v \vee -v$ should not evaluate to $\top$. Instead, we introduce a new symbol depending on $v$ by

$$\delta(v) := +v \vee -v.$$

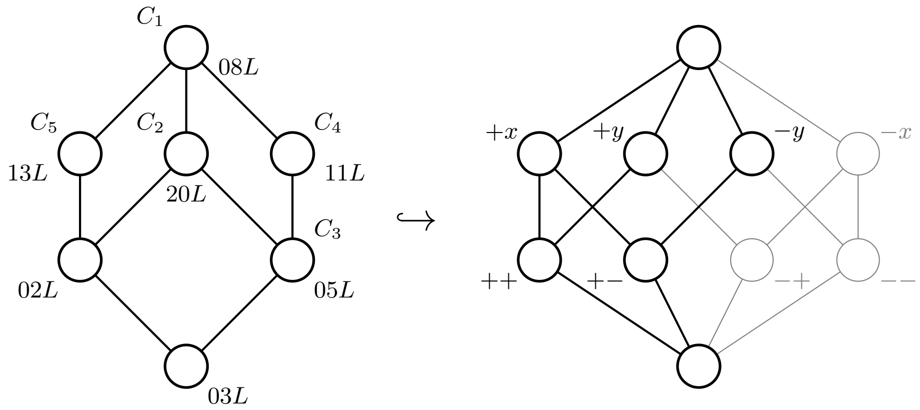


**Fig. 13.** The concept lattice of the learner-competence context in Figure 9 embedded into $\underline{\mathfrak{B}}(\mathbb{D}_2)$

| $\square$ | $a : \top$ | $b : +x \vee +y$ | $c : \top$ | $d : +x \wedge +y$ | $e : \delta(y)$ | $f : -y$ |
|---|---|---|---|---|---|---|
| 02L: $\{+x, -y\}$ | × | × | × | | × | × |
| 03L: *almighty* | × | × | × | × | × | × |
| 05L: $\{+x, +y\}$ | × | × | × | × | × | |
| 08L: $\emptyset$ | × | | × | | | |
| 11L: $\{+y\}$ | × | × | × | | × | |
| 13L: $\{-y\}$ | × | | × | | × | × |
| 20L: $\{+x\}$ | × | × | × | | | |

**Fig. 14.** A representation of the learner-question data in Figure 1 using two pairs of mutually exclusive skills. $\delta(y)$ is an abbreviation for $+y \vee -y$

The reason is this: If mastering a problem requires $+v$ or $-v$, then one of the two skills $+v$ and $-v$ must be present. This is not necessarily the case, and replacing $+v \vee -v$ by $\top$ therefore leads to errors.

With this notation we obtain from the second factorising context in Figure 10

$$\sigma(a) = C_1 \vee C_2 \vee C_3 \vee C_4 \vee C_5 = \top$$
$$\sigma(b) = C_2 \vee C_3 \vee C_4 = +x \vee +y$$
$$\sigma(c) = C_1 \vee C_2 \vee C_3 \vee C_4 \vee C_5 = \top$$
$$\sigma(d) = C_3 = +x \wedge +y$$
$$\sigma(e) = C_3 \vee C_4 \vee C_5 = \delta(y)$$
$$\sigma(f) = C_5 = -y.$$

Figure 14 finally shows that the combination of these findings indeed represents the learner-question context in Figure 1.

## 6 Conclusion

The combination of Boolean factorisations and of embeddings into standard concept lattices gives promising results for the analysis of learner-question data, in particular for the construction of skill functions according to given competence models. In our presentation we have worked out a few examples. A more general and versatile theory seems possible.

## References

1. Doignon, J.P., Falmagne, J.C.: Spaces for the assessment of knowledge. International Journal of Man-Machine Studies 23(2), 175–196 (1985)
2. Korossy, K.: Modeling knowledge as competence and performance. In: Albert, D., Lukas, J. (eds.) Knowledge Spaces: Theories, Empirical Research, and Applications. Lawrence Erlbaum Associates (1999)
3. Doignon, J.P., Falmagne, J.C.: Learning Spaces. Springer, Heidelberg (2011)
4. Keprt, A., Snásel, V.: Binary factor analysis with help of formal concepts. In: Snásel, V., Belohlávek, R. (eds.) CLA. CEUR Workshop Proceedings, vol. 110. CEUR-WS.org (2004)
5. Belohlávek, R., Vychodil, V.: Formal concepts as optimal factors in boolean factor analysis: Implications and experiments. In: Eklund, P.W., Diatta, J., Liquiere, M. (eds.) CLA. CEUR Workshop Proceedings, vol. 331. CEUR-WS.org (2007)
6. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, Heidelberg (1999)
7. Doignon, J.P.: Knowledge spaces and skill assignments. In: Fischer, G.H., Laming, D. (eds.) Contributions to Mathematical Psychology, Psychometrics, and Methodology, Recent Research in Psychology, pp. 111–121. Springer, New York (1994)
8. Düntsch, I., Gediga, G.: Skills and knowledge structures. British Journal of Mathematical and Statistical Psychology 48, 9–27 (1995)
9. Lex, W.: Eine Darstellung zur maschinellen Behandlung von Begriffen. In: Ganter, B., Wille, R., Wolff, K.E. (eds.) Beiträge zur Begriffsanalyse, Wissenschaftsverlag, pp. 141–160 (1986)