Conceptual Factors and Fuzzy Data

DISSERTATION

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

M.Sc. (Mathematik) Cynthia-Vera Glodeanu
geboren am 2. Juni 1985 in Satu Mare

Gutachter: Prof. Dr. Bernhard Ganter, Technische Universität Dresden
Prof. Dr. Michal Krupka, Palacký University of Olomouc

Eingereicht am: 09. 10. 2012
Tag der Disputation: 20. 12. 2012

Preface

Complexity reduction is one of the most important techniques in data analysis. With all the huge amounts of large data sets, for instance medical and biological taxonomies, this possibly applies today more than ever before. Moreover, in many fields, data may also be vague or uncertain. Thus, whenever we have an instrument for data analysis, the question of how to apply complexity reduction methods and how to treat fuzzy data arises rather naturally. In this thesis, we take a very successful tool for data analysis, namely Formal Concept Analysis, and discuss precisely these two questions. In fact, we propose different methods for complexity reduction based on qualitative analyses, and we elaborate on various methods for handling fuzzy data. These two issues split the thesis into two parts. The first part mainly deals with the data reduction, whereas we focus on fuzzy data in the second part. Although each chapter may be read almost on its own, each one builds on and uses results from its predecessors. The main crosslink between the chapters is given by the reduction methods and fuzzy data. In particular, we will also discuss complexity reduction methods for fuzzy data, combining the two issues that motivate this thesis.

Formal Concept Analysis ([GW96]) is an instrument for data analysis based on lattice theory. Starting with a set of *formal objects*, a set of *formal attributes* and an *incidence relation* indicating which object has which attribute, one obtains a *formal context* combining these three components. The context, in turn, allows for the computation of the *formal concepts*. These concepts are understood as units with a *conceptual extent* and a *conceptual intent*, an idea that can be already found in the so-called *Logic of Port Royal* ([Duq87]). The extent of a concept contains all the objects shared by the attributes from its intent. Dually, the intent of a formal concept contains all the attributes that the objects from its extent have in common. The order on the concepts is given by the *subconcept-superconcept relation*. Together with this relation the set of all concepts forms a complete lattice, the so-called *concept lattice*, that represents the basis for further data analysis.

In the first part of the thesis we focus on complexity reduction of those kind of qualitative data that can be represented by a formal context. Usually, data reduction is performed by Factor Analysis and related techniques, which are deeply rooted in statistical and numerical methods and offer a quantitative analysis. These methods search for a preferably low number of unobserved underlying "latent" attributes that explain the covariation among the observed attributes. In Chapter 2 we step away from such approaches by performing a non-metrical analysis. First, we build on some already established links between Factor Analysis of binary data and Formal Concept Analysis ([KS04, Kep06, BV10a]). Our "latent" attributes, called *Boolean factors*, correspond to formal concepts and offer optimal factorisations of binary data, i.e., those factorisations with the smallest possible number of elements. However, such Boolean factors may be large in number and may have limited expressiveness due to their unary nature. Thus, it can hardly be expected that much of a complex data set can be captured by only a few Boolean factors. Wishful thinking suggests to group these factors into well-structured families, which then may be interpreted as *many-valued factors*. These are given by the *conceptual standard scales* of Formal Concept Analysis. Here we focus on one-dimensional *ordinal scales* that give rise to *ordinal factors*. Such an ordinal factor represents a chain of Boolean factors. As it will turn out, the Boolean and the ordinal factors are closely connected with the order dimension of the concept lattice, yielding an application of the latter. We also give necessary and sufficient conditions for finding out whether an arbitrary set of factors indeed provides a factorisation. We claim that the newly developed many-valued factors are easy to understand and useful in applications, and support this by analyses we run on real-world data sets. These are chosen in a way such that they cover different areas and data collecting methods. The analyses' results show that these many-valued factorisations are serious competitors to the latent attributes from ordinary Factor Analysis. Indeed, their expressiveness turns out to be similar to that of Factor Analysis based on metric data.

The frequent appearance of triadic data in psychological applications motivated the generalisation of Factor Analysis and related techniques for triadic data. This issue and the fact that Boolean factors yield optimal factorisations lead us to the development of *triadic Boolean factors*. Once again the factors are easy to understand as they correspond to the well-studied *triadic concepts* from *Triadic Concept Analysis* ([LW95]). Our expectation that *triadic Boolean factors* yield optimal factorisations of triadic data confirms to be true in Chapter 3. Given the interpretation of triadic concepts, the triadic Boolean factors are facilely comprehendable and their handling is unlaboured. This is intended to become explicit in our factorisations of real-world data sets. Further, other results from the dyadic setting may also be generalised to the triadic framework. This includes the necessary and sufficient conditions for checking whether an arbitrary set of factors yields a factorisation. Further, we present some tools from ordinary Factor Analysis in our triadic setting in order to show that our triadic factors can keep up with classical approaches. Apart from that, this chapter shall be concerned with mappings that transform a description of a given object in terms of attributes and conditions into a description of the same object in terms of factors. For the dyadic case, such mappings were utilised in [Out10] for improving classification of binary data. In the triadic setting we ask for transformations between the attribute $\times$ condition space and the factor space. As it will turn out, these mappings constitute an isotone Galois connection.

In Chapter 4 we show that the results of *Hierarchical Classes Analysis* ([DR88]), a method developed for applications in personality organisation and implicit belief systems, coincide with the Boolean and triadic Boolean factors. Hence, the formal concept analytical approach to Factor Analysis and the application driven necessity of data reduction meet in the common point of Boolean and triadic Boolean factorisations. We show how this connection allows the two methods to benefit from each other. On the one hand, Hierarchical Classes Analysis opens new doors for the application of Boolean and triadic Boolean factors. On the other hand, Hierarchical Classes Analysis gains structural explanation, graphical representations and algorithmic issues. At the end of the chapter, we develop a fuzzy variant of Hierarchical Classes Analysis providing a novel fundament for further applications. This leads us directly into the world of fuzzy data that is explored in the second part of the thesis.

Fuzzy data is used when there is no sharp boundary between being a member and not being a member of a class. Thus, fuzzy data often occur in real-world data sets. In *Formal Fuzzy Concept Analysis* ([Pol97, Běl02b]) the incidence relation of the formal context is replaced by a fuzzy relation that provides the truth value to which an object has an attribute. The notions “formal concept”, “concept order”, “concept lattice”, etc. are transformed into the fuzzy setting by considering their fuzzy counterparts.

Attribute exploration ([Gan84, GW96]) is a formal concept analytical tool for knowledge discovery by interactive determination of the implications holding between a given set of attributes. The corresponding algorithm questions the user in an efficient way about the implications between the attributes. As a result of the exploration process one obtains a representative set of examples for the entire theory and a set of implications from which all implications that hold between the considered attributes can be deduced. The method was successfully applied in different real-life applications for binary data. In Chapter 5, we show that attribute exploration can be successfully generalised to the fuzzy setting. We also turn our attention to a variant of attribute exploration, the so-called *attribute exploration with background knowledge*. In this case, the exploration process is shortened by taking into account some background knowledge that the user has at the beginning of the exploration. We show that the proposed method can also be generalised to the fuzzy setting. In order to apply the two methods in practice we develop appropriate algorithms.

In Chapter 6 we present a method for complexity reduction of fuzzy data, more precisely of the fuzzy concept lattice of a given formal fuzzy context. This is a different kind of data reduction than the conceptual factorisations dealt with in the first part of the thesis, as we allow the users to express their preferences over the attribute set. Based on these preferences the users only obtain those formal fuzzy concepts that are relevant to them. Our main goal is to allow preferences stated on compounded attributes, i.e., on qualities that include more than one trait. However, we treat attributes with only one trait as well. Since the users are allowed to enter their preferences, it is very likely that their inputs are somewhat redundant. This would make the further handling of these favouritisms more difficult than necessary. Therefore, we also develop techniques for removing these redundancies. Having a non-redundant set of preferences allows the user to review his choices and alter them conveniently.

In the last chapter we combine fuzzy data with a notion we already got acquainted with in the previous part, namely “triadic data”, and come to a framework that we call *Fuzzy-valued Triadic Concept Analysis*. As already mentioned, triadic data is encountered in psychological interviews. When uncertainty comes into play, it may be promising to use fuzzy-valued triadic data. The same applies for image analysis, experimental design, spectroscopy and chromatography, to name but a few. Since Fuzzy-valued Triadic Concept Analysis is a new framework we first present what builds the fundamentals of Formal Concept Analysis in this setting. Afterwards, we study implications in such data sets. These implications can be between attributes under some conditions or between tuples of attributes and conditions. Since we may arbitrarily interchange the roles of objects, attributes and conditions in a triadic context, this gives rise to nine different families of implications in a fuzzy-valued triadic context. As we will see, the results from Chapter 5 play an essential role for this matter. In the last section we bring another notion familiar from the previous part into our new setting and explore the fuzzy-valued triconceptual factorisations. Once again many results from the dyadic and triadic crisp settings hold for the new framework. Thereby, we establish a tight connection between the main topics of this thesis.

Some original results of this thesis have already been published in journals and conference proceedings [Glo10, Glo11b, Glo11a, Glo11d, Glo12b, Glo12a]. We refer to these publications at the beginning of each chapter, but do not cite each specific result, except those that were created in collaboration with one or more coauthors [BGV12, GG12].

Acknowledgements

Foremost I would like to express my profound gratitude to my supervisor Bernhard Ganter for his patience, encouragement, enthusiasm and guidance whilst allowing me the room to work in my own way.

I wish to express my warm and sincere thanks to Christian Săcărea for his support, guidance through my undergraduate studies and for pushing me in the right direction.

I am immensely pleased to have had the current and former members of the Institute of Algebra as my colleagues. The pleasant, friendly, fun and helpful atmosphere they create made every working day enjoyable. In particular, I would like to address special thanks to Sebastian Kerkhoff for his careful proofreading and his numerous remarks that made this thesis smoother to read. Further, I would like to thank Christian Meschke for finding time to proofread some parts of the thesis. I am thankful to Daniel Borchmann for implementing helpful procedures in ConExp-Clj. The time for working out the examples was significantly reduced. Many thanks to Tobias Schlemmer for his LaTeX support.

I am thankful to Radim Bělohlávek, Michal Krupka and Vilém Vychodil for their cooperation and for inviting me to their department for two months.

I am indebted to Deutscher Akademischer Austausch Dienst for the financial support of my PhD thesis and would be indebted to others if it were not for the scholarship.

Last but not least I am deeply grateful to my friends and family for their support, especially to my mother Gabriela who cheered along the whole way.

Contents

1
Preliminaries

The purpose of this chapter is to present the basic notions and results that are needed in the course of this thesis. This preliminary chapter should not be understood as an introduction to the underlying mathematical theories. For this aim we give adequate references in each section. The reader who is already familiar with one or more of the fields from the sections of this chapter may easily skip them.

1.1. Order-theoretic Notions

This section presents the rudimentary bases of order theory. For more details we refer to standard literature such as [DP02, Bir67, Grä03, Ern82].

A binary relation R on a set P is called a **(partial) order** on P if it is reflexive, antisymmetric and transitive, i.e., if the following properties are satisfied for all $x, y, z \in P$:

- $x \mathrm{R} x$, (reflexivity)
- $x \mathrm{R} y$ and $y \mathrm{R} x$ implies $x = y$, (antisymmetry)
- $x \mathrm{R} y$ and $y \mathrm{R} z$ implies $x \mathrm{R} z$. (transitivity)

A binary relation R on a set P is called an **equivalence relation** if it is reflexive, transitive and symmetric, where the last condition is given by

- $x \mathrm{R} y$ implies $y \mathrm{R} x$. (symmetry)

An **ordered set** (**partially ordered set, poset**) is a pair $(P, \leq)$, where P is a set and $\leq$ is an order relation on P. We write $x < y$ to indicate $x \leq y$ and $x \neq y$. Let $x, y \in P$ with $x < y$. Then, x is a **lower neighbour** of y if there is no $z \in P$ with $x < z < y$. In this case y is an **upper neighbour** of x. Two elements $x, y \in P$ are called **comparable** if $x \leq y$ or $y \leq x$ holds, otherwise they are **incomparable**. A subset of $(P, \leq)$ is called a **chain** if any two of its elements are comparable. A subset in which any two elements are incomparable is called an **antichain**. The **width** of a finite poset $(P, \leq)$ is the maximal size of an antichain in it. The **length** is the maximal size of a chain in $(P, \leq)$ minus one.

A mapping $\varphi : P \to Q$ between two ordered sets $(P, \leq)$ and $(Q, \leq)$ is called an **order-embedding** if for all $x, y \in P$ the following condition is satisfied

$$x \leq y \Longleftrightarrow \varphi(x) \leq \varphi(y).$$

If φ satisfies only the $\Rightarrow$-part, then it is called **order-preserving**. A bijective order-embedding is called **(order-)isomorphism**.

Let $(P, \leq)$ be a poset and X a subset of P. A **lower bound** of X is an element $p \in P$ with $p \leq x$ for all $x \in X$. Dually one can define an **upper bound**. An element $p \in P$ is called the **supremum** of X if p is the least upper bound of X. In case it exists, the supremum of X is usually denoted by $\bigvee X$. Dually the greatest lower bound of a subset X is called the **infimum** of X. If it exists, it is denoted by $\bigwedge X$. The poset P is called a **lattice** if for all elements $x, y \in P$ the binary supremum $x \vee y := \bigvee\{x, y\}$ and the binary infimum $x \wedge y := \bigwedge\{x, y\}$ always exist. An ordered set in which every subset has a supremum and an infimum is called a **complete lattice**.

For an element v of a complete lattice $\mathbf{V}$ we define

$$v_\bullet := \bigvee\{x \in V \mid x < v\} \text{ and } v^\bullet := \bigwedge\{x \in V \mid x > v\}.$$

We call v **$\bigvee$-irreducible** (**supremum-irreducible**) if $v \neq v_\bullet$. Dually, v is called **$\bigwedge$-irreducible** (**infimum-irreducible**) if $v \neq v^\bullet$. Further, a set $X \subseteq V$ is called **supremum-dense** in V, if every element from V can be represented as the supremum of a subset of X and, dually, **infimum-dense**, if $v = \bigwedge\{x \in X \mid v \leq x\}$ for all $v \in V$.

A **closure system** on a set P is a set of subsets which contains P and is closed under arbitrary intersections, i.e., $\mathcal{U} \subseteq \mathfrak{P}(P)$ is a closure system if $P \in \mathcal{U}$ and $\mathcal{X} \subseteq \mathcal{U}$ implies $\bigcap \mathcal{X} \in \mathcal{U}$. A **closure operator** φ on P is a mapping $\varphi : \mathfrak{P}(P) \to \mathfrak{P}(P)$ satisfying the following conditions:

- $X \subseteq Y \Longrightarrow \varphi(X) \subseteq \varphi(Y)$, (monotony)
- $X \subseteq \varphi(X)$, (extensivity)
- $\varphi\varphi(X) = \varphi(X)$ (idempotency)

for all $X, Y \subseteq P$. The closure operator assigns to each subset $X \subseteq P$ its *closure* $\varphi(X) \subseteq P$. Given a closure system $\mathcal{U}$ on P one finds the corresponding closure operator on P by

$$\varphi_{\mathcal{U}}(X) := \bigcap\{Y \in \mathcal{U} \mid X \subseteq Y\}.$$

Conversely, given a closure operator φ one finds the corresponding closure system by

$$\mathcal{U}_\varphi := \{\varphi(X) \mid X \subseteq P\}.$$

Further, we always have that $\varphi_{\mathcal{U}_\varphi} = \varphi$ and $\mathcal{U}_{\varphi_{\mathcal{U}}} = \mathcal{U}$, so closure operators and closure systems are in a one-to-one correspondence.

A **Galois connection** is a pair of mappings (φ, ψ) between two ordered sets $(X, \leq)$ and $(Y, \leq)$, i.e., $\varphi : X \to Y$ and $\psi : Y \to X$ that satisfies

- $x_1 \leq x_2 \Longrightarrow \varphi(x_1) \geq \varphi(x_2)$,
- $y_1 \leq y_2 \Longrightarrow \psi(y_1) \geq \psi(y_2)$,
- $x \leq \psi\varphi(x)$ and $y \leq \varphi\psi(y)$

for all $x, x_1, x_2 \in X$ and $y, y_1, y_2 \in Y$. Galois connections can also be characterised by a single condition. They are precisely those pairs of mappings (φ, ψ) for which we have

$$x \leq \psi(y) \Longleftrightarrow y \leq \varphi(x).$$

1.2. Vagueness and the Fuzzy Approach

Consider the predicates “old”, “tall”, “big”, “red”, “heap” etc. These are all vague. Indeed, it is often unclear whether or not they apply to a given object. There are people that are certainly tall and some that are surely not. However, there are people where there is not a definite answer to the question whether they are tall or not. Such considerations violate the classical principle of bivalence. Seemingly, vague predicates, like the ones listed above, lack well-defined extensions. For instance, there is no sharp boundary between the tall people and the others. We refer to [KS97] for the treatment of vague predicates. The work includes different (reprinted) essays on the topic.

One of the most outstanding problems vagueness induces are the *sorites paradoxes*[1]. The interest in them was lost after the antiquity and rose again after Russell’s seminal paper ([Rus23]). Let us consider the example which gave the name to these paradoxes. Suppose we have a big heap of sand. If we remove a sand grain, the heap still remains big. This means, by consecutively removing a sand grain from the heap, it always remains big. But after finitely many steps we obtain a heap consisting of one sand grain, which is obviously not big. The first vague expression in “big heap” is the notion “big”. What does a “big heap” mean? How tall or wide must it be such that we can consider it big? Let us suppose a big heap means as big as a human. This means that by removing just one grain of sand, the heap turns from being big to not being big. This is not intuitive. And why exactly do we compare big heaps to humans? How big are these humans? This is a good example of how vague some frequently used notions actually are. We obtain the paradox of the heap by consecutively applying the rule of inference *modus ponens* which states the following: If P and Q are sentences such that

$$\text{if } P \text{ is true and } P \text{ implies } Q, \text{ then } Q \text{ is true.}$$

In Subsection 1.2.2 we will see that such paradoxes can be easily solved through fuzzy logic. Even further, the latter allows a (more) proper treatment of vagueness.

1.2.1. From Two-valued to Many-valued Logic

As we have seen, vague predicates have borderline cases, fuzzy boundaries, are susceptible to sorites paradoxes, and a bivalent approach to them is not sufficient. However, a many-

[1] The word "sorites" derives from the Greek word for heap. The original characterisation of the paradox is attributed to Eubulides of Miletus.

valued logic/fuzzy logic might be! But before we came to fuzzy logic there had been a long way to go:

The first steps towards fuzzy logic were made by many-valued logics. Although there are some pioneer works about the latter ([Pei85, Mac96]), the first milestone was put in [Łuk18] by Łukasiewicz, where a three-valued logic was mentioned. The first occurrence of an infinite logic was in [Łuk22], which was further elaborated in [ŁT30]. There are many notable names and papers that contributed to the development of many-valued logics. We will mention further just some, those which are often cited as the most important ones. A detailed bibliography can be found for instance in [Háj98] and [KY95]. Heyting introduced in [Hey30] a three-valued propositional calculus related to intuitionistic logic. Gödel presented in [Göd32] an infinite hierarchy of finitely-valued systems, the infinite-valued logic of it is now called the *Gödel logic*. An extensive work about many-valued logics is [Got88] by Gottwald.

The t-norms play an important role in our setting. They were developed by Menger in [Men42]. His purpose was to construct metric spaces where instead of using numbers for describing distances between two elements one uses probability distributions. The definition of t-norms used nowadays was given by Schweizer and Sklar in [SS60]. In [Lin] it was shown that using the Łukasiwicz, Goguen and Gödel t-norms, one can construct all continuous t-norms.

1965 is the year of birth of fuzzy sets, when Zadeh published [Zad65]. Before that, many-valued logic was mainly considered as a theoretical field. The first work on fuzzy logic seems to be [Gog69] by Goguen. He also studied fuzzy sets with values in a lattice ([Gog67]). Important works of Zadeh include [Zad75a] on linguistic variables (for instance, *age* with possible values *young, medium, old*), [Zad75b] on fuzzy logic, and [Zad96] presenting the generalised version of modus ponens and compositional rule of inference. Very important contributions to the syntax and semantics of fuzzy logic were given by Pavelka in [Pav79].

Let us quote some important statements which show the evolution of logic over time:

> *The same thing cannot at the same time both belong and not belong to the same object and in the same respect. [...] Of any object, one thing must either be asserted or denied.* Aristotle

> *Logic changes from its very foundations if we assume that in addition to truth and falsehood there is some third logical value or several such values.* Łukasiewicz

> *More often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership. [...] Yet, the fact remains that such imprecisely defined "classes" play an important role in human thinking, particularly in the domains of pattern recognition, communication of information, and abstraction.* Zadeh

Zadeh notes that the descriptions used by humans are neither black nor white and that there is a gradual transition from black to white. Further, he points out that classical mathematics is not able to grasp these unsharp notions. Contradicting the principle of bivalence, Zadeh states that there are different cases of belonging to a fuzzy set besides "fully belonging" and "fully not belonging". Hence, being a member of a fuzzy set is a graded matter.

Zadeh has always made the useful distinction between the two different meanings of fuzzy logic. Let us cite from [Mar94]:

> *In a narrow sense, fuzzy logic, FLn, [...] is an extension of multivalued logic. However, the agenda of FLn is quite different from that of traditional multivalued logic. In particular, such key concepts in FLn as the concept of a linguistic variable, canonical form, fuzzy if-then-rule, quantification and defuzzification, the compositional rule of inference, [...] are not addressed in traditional systems. [...] In wide sense, fuzzy logic, FLw, is a fuzzily synonymous with the fuzzy set theory, which is the theory of classes with unsharp boundaries.*
>
> Marks II

Fuzzy theory was successfully used in both theoretical and real-world applications. Considering the latter, its main breakthrough came with the development of a fuzzy controller by Mamdani and Assilian in [MA75]. In [Adl85] the CADIAG-2 system was presented which is a fuzzy expert system making inferences from patient data. We do not aim here to list the various applications of fuzzy theory, extensive references can be found, for instance, in [KY95].

1.2.2. Fuzzy Sets and Fuzzy Logic

First let us come back to the sorites paradoxes. In order to handle them within the framework of fuzzy logic we need the fuzzy version of modus ponens: Let $\mathrm{tv}(P)$, $\mathrm{tv}(Q)$ and $\mathrm{tv}(P \to Q)$ be the truth values of the sentences P, Q and $P \to Q$, respectively. The fuzzy version of modus ponens states:

$$\text{if } \mathrm{tv}(P) \geq a \text{ and } \mathrm{tv}(P \to Q) \geq b, \text{ then } \mathrm{tv}(Q) \geq a\&b \tag{1.1}$$

where a and b are some truth values and & is the conjunction. We apply the fuzzy version of modus ponens, choosing $P(n)$:=*A heap with n sand grains is big.* Suppose n is big enough. We start with $\mathrm{tv}(P(n)) = 1$ and $\mathrm{tv}(P(n) \to P(n-1)) = 0.999$. Thus, $\mathrm{tv}(P(n-1)) = 0.999$. We apply modus ponens once more choosing $\mathrm{tv}(P(n-1) \to P(n-2)) = 0.999$. We obtain $\mathrm{tv}(P(n-2)) = 0.999^2$. This way we no longer come to the conclusion that a heap with one sand grain is big with truth value 1.

In the following we proceed as in [Háj98, Běl02b], illustrating the underlying structures of fuzzy theory. We want to compare truth values and require therefore an order on the set of truth values. The set of truth values is denoted by L and one usually takes for it the real unit interval $[0,1]$ with its natural ordering, where 0 denotes (full) falsity and 1 (full) truth. We would also like to have for two truth values a and b a value that is the least truth value which is greater than both of them and dually a value that is the greatest truth value which is least than both a and b. Therefore, we require the existence of suprema and infima for arbitrary truth values. These requirements lead us to the structure of a (complete) lattice.

Now we are looking for operations on L which shall model the logical connectives. Since fuzzy theory is a generalisation of classical mathematics, these operations should coincide with the classical ones, if we restrict them to the truth values 0 and 1, i.e., $L = \{0,1\}$. We denote the conjunction by $\otimes$ which shall be a binary operation on L. Considering the above requirement, $1 \otimes 1 = 1$ and $1 \otimes 0 = 0 \otimes 1 = 0 \otimes 0 = 0$ should hold. We also want the truth values of $a \otimes b$ and $b \otimes a$ to be equal. This leads us to the commutativity of $\otimes$. In a similar way, we ask for associativity of $\otimes$. Altogether, $(L, \otimes, 1)$ must be a commutative monoid. If $a_1 \leq b_1$ and $a_2 \leq b_2$ hold, it is intuitive to require $a_1 \otimes a_2 \leq b_1 \otimes b_2$, i.e., $\otimes$ shall be non-decreasing. The conjunction $\otimes$ coincides with the notion of a (continuous) *t-norm*. We refer to [KMP00] for properties and constructions of t-norms.

Let us take another look at the fuzzy version of modus ponens. The rule is sound, because through it we cannot assign a greater truth value to Q than it has, see (1.1). We consider a special case of fuzzy modus ponens by taking $tv(P) = a$, $tv(P \to Q) = b$ and $tv(Q) = c$. Then, $\otimes$ gives us:

1. The lower estimation of Q which translates to: $b \leq a \to c$ implies $a \otimes b \leq c$;
2. the highest possible lower estimation of the truth value of Q such that the rule still remains sound. This corresponds to: If $a \otimes b \leq c$, then $b \leq a \to c$.

The two conditions give us

$$a \otimes b \leq c \text{ if and only if } b \leq a \to c. \tag{1.2}$$

This condition is called the **adjointness property**. If $\otimes$ is a continuous t-norm, then there is a unique operation $\to$ satisfying the adjointness property for all $a, b, c \in L$, namely $a \to c := \bigvee\{b \in L \mid a \otimes c \leq b\}$. The algebraic structures which satisfy it are called *residuated lattices*. An early paper about them is [DW39] by Dilworth and Ward. In the following we will consider special kinds of residuated lattices, namely those that have an extra operation on them.

Definition 1.1. A **complete residuated lattice with (truth-stressing) hedge** is an algebra $\mathbf{L} := (L, \wedge, \vee, \otimes, \to, ^*, 0, 1)$ such that:

- $(L, \wedge, \vee, 0, 1)$ is a complete lattice,
- $(L, \otimes, 1)$ is a commutative monoid,
- 0 is the least and 1 the greatest element,
- the adjointness property (1.2) holds for all elements from L.

The hedge $(-)^*$ is a unary operation on L satisfying:

i) $a^* \leq a$,

ii) $(a \to b)^* \leq a^* \to b^*$,

iii) $a^{**} = a^*$,

iv) $1^* = 1$

for every $a, b \in L$. The elements of L are called **truth degrees**, $\otimes$ and $\rightarrow$ are called **multiplication** and **residuum**, respectively. The last two represent the truth functions of "fuzzy conjunction" and "fuzzy implication", respectively.

Until explicitly said otherwise, $\mathbf{L}$ denotes a linearly ordered residuated lattice in this thesis. The hedge $(-)^*$ is a (truth function of) logical connective "very true", see [Háj98, Háj01]. Properties (i)-(iv) have natural interpretations, i.e., (i) can be read as "if a is very true, then a is true", (ii) can be read as "if $a \rightarrow b$ is very true and if a is very true, then b is very true", etc.

As we have already discussed, a common choice of $\mathbf{L}$ is a structure with $L = [0,1]$, $\wedge$ and $\vee$ being minimum and maximum, respectively, and $\otimes$ being a continuous t-norm with its corresponding residuum $\rightarrow$. The three most important pairs of adjoint operations on the unit interval are:

Łukasiewicz: $a \otimes b := \max(0, a + b - 1)$ with $a \rightarrow b = \min(1,\ 1 - a + b)$,

Gödel: $a \otimes b := \min(a, b)$ with $a \rightarrow b = \begin{cases} 1, & a \leq b, \\ b, & a \gneq b, \end{cases}$

Goguen: $a \otimes b := ab$ with $a \rightarrow b = \begin{cases} 1, & a \leq b, \\ b/a, & a \gneq b. \end{cases}$

Typical examples for the hedge are

identity: $a^* := a,\ $ for all $a \in L$,

globalisation: $a^* := \begin{cases} 0, & a \in L \setminus \{1\}, \\ 1, & a = 1. \end{cases}$

Residuated lattices have numerous properties, for an overview see for instance [Běl02b]. We will list some of them which will be needed in the forthcoming chapters. Any residuated lattice $\mathbf{L}$ satisfies the following properties:

$$(x \otimes y) \rightarrow z = x \rightarrow (y \rightarrow z), \tag{1.3}$$

$$x \otimes \bigvee_{i \in I} y_i = \bigvee_{i \in I} (x \otimes y_i), \tag{1.4}$$

$$x \rightarrow \bigwedge_{i \in I} y_i = \bigwedge_{i \in I} (x \rightarrow y_i), \tag{1.5}$$

$$\bigvee_{i \in I} x_i \rightarrow y = \bigwedge_{i \in I} (x_i \rightarrow y), \tag{1.6}$$

$$\bigwedge_{i \in I} \bigwedge_{j \in J} (x_i \rightarrow y_j) = (\bigvee_{i \in I} x_i) \rightarrow (\bigwedge_{j \in J} y_j), \tag{1.7}$$

for all $x, x_i, y, y_i, y_j \in L$.

A **fuzzy set** (**L-set**) A in a universe U is a mapping $A : U \rightarrow L$, where $A(u)$ is interpreted as "the degree to which u belongs to A". We write $u \in A$, if and only if $A(u) = 1$. If $U = \{u_1, \ldots, u_n\}$, then A can be denoted by $A = \{{}^{a_1}/u_1, \ldots, {}^{a_n}/u_n\}$ meaning that $A(u_i)$ equals a_i for each $i \in \{1, \ldots, n\}$. The α**-cut** of A is a subset ${}^{\alpha}A$ of U defined by

$${}^{\alpha}A := \{u \in U \mid A(u) \geq \alpha\}. \tag{1.8}$$

Let $\mathbf{L}^U$ denote the collection of all **L**-sets in U. The operations on **L**-sets are defined component-wise. For instance, the intersection of **L**-sets $A, B \in \mathbf{L}^U$ is an **L**-set $A \cap B$ in U such that $(A \cap B)(u) := A(u) \wedge B(u)$ for each $u \in U$, etc. For $A, B \in \mathbf{L}^U$, the **subsethood degree** is defined by

$$\mathrm{S}(A, B) := \bigwedge_{u \in U} (A(u) \to B(u)),$$

which generalises the classical subsethood relation $\subseteq$. Therefore, $\mathrm{S}(A, B)$ represents the degree to which A is a subset of B. In particular, we write $A \subseteq B$ if and only if $\mathrm{S}(A, B) = 1$, i.e., $A(u) \leq B(u)$ for all $u \in U$. Further, we write $A \subset B$ if and only if $A(u) < B(u)$ for all $u \in U$.

Binary fuzzy relations (**L-relations**) between X and Y can be thought of as **L**-sets in the universe $X \times Y$. A binary **L**-relation $\approx$ on a set X is called an **L-equality** if it satisfies the following conditions:

$$\begin{aligned}
&(x \approx x) = 1,\\
&(x \approx y) = (y \approx x),\\
&(x \approx y) \otimes (y \approx z) \leq (x \approx z),\\
&(x \approx y) = 1 \text{ implies } x = y
\end{aligned}$$

for all $x, y, z \in X$. Further, a binary **L**-relation R between X and Y is **compatible w.r.t. the L-equalities $\approx_X$ and $\approx_Y$** if

$$\mathrm{R}(x_1, x_2) \otimes (x_1 \approx_X x_2) \otimes (y_1 \approx_Y y_2) \leq \mathrm{R}(y_1, y_2)$$

for all $x_1, x_2 \in X$ and $y_1, y_2 \in Y$.

Definition 1.2. An **L-order** on a set X with an **L**-equality relation $\approx$ is a binary **L**-relation $\precsim$ which is compatible w.r.t. $\approx$ and satisfies

$$\begin{aligned}
&x \precsim x = 1,\\
&(x \precsim y) \wedge (y \precsim x) \leq x \approx y,\\
&(x \precsim y) \otimes (y \precsim z) \leq x \precsim z
\end{aligned}$$

for all $x, y, z \in X$. If $\precsim$ is an **L**-order on a set X with an **L**-equality $\approx$, we call the pair $((X, \approx), \precsim)$ an **L-ordered set**.

Fuzzy closure operators (**L***-closure operators) were introduced in [BG96] and studied further by Bělohlávek *at al.*, see for instance [Běl01, BFV05]. The definition given in [BG96] mirrors more a crisp thinking, representing a special case of the one given in [Běl01]. Therefore, we will use the latter.

Definition 1.3. An **L*-closure operator** on a set X is a mapping $\mathrm{C} : \mathbf{L}^X \to \mathbf{L}^X$ satisfying

$$A \subseteq \mathrm{C}(A), \tag{1.9}$$
$$\mathrm{S}(A_1, A_2)^* \leq \mathrm{S}(\mathrm{C}(A_1), \mathrm{C}(A_2)), \tag{1.10}$$
$$\mathrm{C}(A) = \mathrm{C}\,\mathrm{C}(A) \tag{1.11}$$

for all $A, A_1, A_2 \in \mathbf{L}^X$ and where $(-)^*$ denotes the hedge of the residuated lattice $\mathbf{L}$. A system $\mathcal{S} := \{A_j \in \mathbf{L}^X \mid j \in J\}$ is an **$\mathbf{L}^*$-closure system** if for each $A \in \mathbf{L}^X$ it holds that

$$\bigcap_{j \in J} (\mathrm{S}(A, A_j)^* \to A_j) \in \mathcal{S}, \tag{1.12}$$

where

$$\Big(\bigcap_{j \in J} \mathrm{S}(A, A_j)^* \to A_j\Big)(x) := \bigwedge_{j \in J} (\mathrm{S}(A, A_j)^* \to A_j(x))$$

for every $x \in X$.

For the globalisation, (1.10) and (1.12) become

$$A_1 \subseteq A_2 \Longrightarrow \mathrm{C}(A_1) \subseteq \mathrm{C}(A_2) \quad \text{and} \quad \bigcap_{j \in J} (\mathrm{S}(A, A_j)^* \to A_j) = \bigcap_{j \in J, A \subseteq A_j} A_j,$$

respectively. Only these variants of $\mathbf{L}^*$-closure operators and systems were studied in [BG96].

There is a simple yet useful characterisation of $\mathbf{L}^*$-closure systems:

Theorem 1.4 ([Běl01]). *A system $\mathcal{S}$ on $\mathbf{L}^X$ closed under arbitrary intersections is an $\mathbf{L}^*$-closure system if and only if for each $a \in L$ and $A \in \mathcal{S}$ it holds that $(a^* \to A) \in \mathcal{S}$, where $a^* \to A$ is an $\mathbf{L}$-set given by $(a^* \to A)(x) := a^* \to A(x)$ for all $x \in X$.*

In [BFV05] an $\mathbf{L}^*$-kernel operator was studied which is defined as follows:

Definition 1.5. An **$\mathbf{L}^*$-kernel operator** on a set X is a mapping $\kappa : \mathbf{L}^X \to \mathbf{L}^X$ satisfying

$$\begin{aligned} \kappa(A) &\subseteq A, \\ \mathrm{S}(A_1, A_2)^* &\leq \mathrm{S}(\kappa(A_1), \kappa(A_2)), \\ \kappa(A) &= \kappa\kappa(A) \end{aligned}$$

for every $A, A_1, A_2 \in \mathbf{L}^X$. A system $\mathcal{S} := \{A_j \in \mathbf{L}^X \mid j \in J\}$ is an **$\mathbf{L}^*$-kernel system** if for each $A \in \mathbf{L}^X$ it holds that

$$(\bigcup_{j \in J} \mathrm{S}(A, A_j)^* \otimes A_j) \in \mathcal{S},$$

where

$$(\bigcup_{j \in J} \mathrm{S}(A, A_j)^* \otimes A_j)(x) := \bigvee_{j \in J} (\mathrm{S}(A, A_j)^* \otimes A_j(x))$$

for every $x \in X$.

Once again, if we choose for $(-)^*$ the globalisation, we obtain

$$A_1 \subseteq A_2 \Longrightarrow \kappa(A_1) \subseteq \kappa(A_2) \quad \text{and} \quad (\bigcup_{j \in J} \mathrm{S}(A, A_j)^* \otimes A_j) = \bigcup_{j \in J, A \subseteq A_j} A_j.$$

Theorem 1.6 ([BFV05]). *A system $\mathcal{S}$ on $\mathbf{L}^X$ closed under arbitrary unions is an $\mathbf{L}^*$-kernel system if and only if for each $a \in L$ and $A \in \mathcal{S}$ it holds that $(a^* \otimes A) \in \mathcal{S}$, where $a^* \otimes A$ is an $\mathbf{L}$-set given by $(a^* \otimes A)(x) := a^* \otimes A(x)$ for all $x \in X$.*

Further, we are also interested in Galois connections in the fuzzy setting. These have been developed in [Běl02b] and will be introduced in the next definition. Note however, that this definition changes when we use hedges (see Definition 1.38, page 27).

Definition 1.7. An **L-Galois connection** between the sets X and Y is a pair (φ, ψ) of mappings

$$\varphi : \mathbf{L}^X \to \mathbf{L}^Y \text{ and } \psi : \mathbf{L}^Y \to \mathbf{L}^X$$

satisfying the following conditions:

$$\begin{aligned} \mathrm{S}(X_1, X_2) &\leq \mathrm{S}(\varphi(X_2), \varphi(X_1)),\\ \mathrm{S}(Y_1, Y_2) &\leq \mathrm{S}(\psi(Y_2), \psi(Y_1)),\\ X_1 &\subseteq \psi\varphi(X_1),\\ Y_1 &\subseteq \varphi\psi(Y_1) \end{aligned}$$

for every $X_1, X_2 \in \mathbf{L}^X$ and $Y_1, Y_2 \in \mathbf{L}^Y$.

1.3. Formal Concept Analysis

Developed at the beginning of the 80s by a research group around Wille, Formal Concept Analysis is an instrument for data analysis based on lattice theory.

In Port Royal Logic a concept is understood as a unit with a conceptual extent and a conceptual intent.[2] Formal Concept Analysis deals with a mathematical formalisation of this notion. The starting point in Formal Concept Analysis is a set of *(formal) objects*, a set of *(formal) attributes* and an *incidence relation* indicating which object has which attribute. These three components are combined into a *formal context* from which the *concept lattice* is computed. A *formal concept* of this lattice contains in its *extent* all the objects shared by the attributes from its intent. Dually, the *intent* of a formal concept contains all the attributes which the objects from its extent have in common. The order on the concepts is given by the *subconcept-superconcept relation*. The concept lattice is the basis for further data analysis. It can be represented graphically in order to facilitate the communication, and it can be explored by algebraic methods. The formal context specifies the frame in which the analysis is valid. If one abandons this frame, the validity of the analysis might get lost. Therefore, the most important step in Formal Concept Analysis is the selection of a suitable formal context.

In this section, almost every result and definition is taken from [GW96] which is the main reference for Formal Concept Analysis.

Definition 1.8. A **formal context** $\mathbb{K} = (G, M, I)$ consists of two sets G and M and a binary relation $I \subseteq G \times M$. The elements of G are called **objects**, the ones of M **attributes** and $(g, m) \in I$ is read "the object g has the attribute m". The relation I is called the **incidence relation** of the context.

Finite contexts can be represented through cross tables, see Figure 1.1. The rows of the table are named after the objects and the columns after the attributes. A cross in row g and column m means $(g, m) \in I$.

[2] See [Duq87] for further explanations and references.

Definition 1.9. For $A \subseteq G$ and $B \subseteq M$ the **derivation operators** are defined by

$$A' := \{m \in M \mid (g, m) \in I \text{ for all } g \in A\},$$
$$B' := \{g \in G \mid (g, m) \in I \text{ for all } m \in B\}.$$

Hence, we have two mappings both labelled with $(-)'$. In general, there should be no confusion which operator maps from where to where. However, in order to distinguish the two mappings, it might be useful to use notations like $(-)^{\uparrow}$ and $(-)^{\downarrow}$. The first operator associates to an object set A the attributes these objects have in common and the second operator associates to a set B of attributes the objects they share. The derivation operators form a Galois connection between the powerset lattices of G and M. Hence, their compounds are closure operators.

Definition 1.10. A pair (A, B) is called a **formal concept** of the context (G, M, I) if $A \subseteq G$, $B \subseteq M$, $A' = B$ and $B' = A$ hold. Then, A is called the **extent** and B is called the **intent** of the concept (A, B). The set of all formal concepts of (G, M, I) is denoted by $\mathfrak{B}(G, M, I)$.

The preposition *formal* in the nomenclature of "formal context" and "formal concept" shall point out that these are just mathematical definitions modelling the issue described above. The preposition *formal* will often be omitted.

Formal concepts represent maximal rectangles filled with crosses in the cross table representation of the context.

One may also be interested in *preconcepts*, as we will see in the upcoming chapters. A **preconcept** is a tuple (A, B) with $A \subseteq G$ and $B \subseteq M$ such that $A' \subseteq B$ and $B' \subseteq A$.

For an object $g \in G$, we write g' instead of $\{g\}'$ for the **object intent** $\{m \in M \mid g\ I\ m\}$ of the object g. Correspondingly, m' denotes the **attribute extent** of the attribute $m \in M$. Further, we denote by γg the **object concept** (g'', g') and by μm the **attribute concept** (m', m'').

Example 1.11. The context from Figure 1.1 ([SW91]) was obtained in a therapy with a patient suffering from anorexia nervosa. The interviewing method used in the therapy was *repertory grid*. Such grids can be easily transformed into formal contexts, as detailed in Section 2.3.

The context from Figure 1.1 has 23 concepts, which can be read from the *concept lattice*, as we will see later on. But first, let us take a look at some concepts. For instance,

$$(\{\text{father, mother, brother-in-law}\}, \{\text{dutiful, hearty, superficial, ambitious}\})$$

is a formal concept. Its meaning is that the father, mother and brother-in-law have exactly the attributes dutiful, hearty, superficial and ambitious in common. Dually, the attributes from the intent are shared only by the persons from the extent. Of course, this holds from the point of view of the patient. In addition to the attributes from the intent of the previous concept, the father and the brother-in-law are also attentive, i.e.,

$$(\{\text{father, brother-in-law}\}, \{\text{dutiful, hearty, attentive, superficial, ambitious}\})$$

is another formal concept.

Anorexia	vulnerable	withdrawn	confident	dutiful	hearty	difficult	attentive	easily offended	not irascible	anxious	talkative	superficial	sensitive	ambitious
myself	×	×	×		×	×	×		×	×			×	×
my ideal	×		×	×	×		×		×				×	×
father	×	×		×	×	×	×	×	×	×		×	×	×
mother	×	×		×	×	×		×	×	×		×	×	×
sister	×	×		×	×	×	×		×	×			×	×
brother-in-law			×	×	×		×				×	×		×

Figure 1.1.: Data from an anorexia nervosa therapy ([SW91])

One may wish to browse between the formal concepts, for instance going from the more general concepts to the more concrete ones. Formal Concept Analysis models the subconcept-superconcept relation in the following self-evident way: For two concepts $(A,B),(C,D)$ of the context (G,M,I) we define

$$(A,B) \le (C,D) :\Longleftrightarrow A \subseteq C \quad (\Longleftrightarrow B \supseteq D),$$

in which case we say that (A,B) is a **subconcept** of (C,D) or that (C,D) is a **superconcept** of (A,B). With this relation the set of all formal concepts $\mathfrak{B}(G,M,I)$ is an ordered set. We call $\underline{\mathfrak{B}}(G,M,I) := (\mathfrak{B}(G,M,I),\le)$ the **concept lattice** of the context (G,M,I). The following *Basic Theorem on Concept Lattices* ensures that this notion is not misleading: The concept lattice actually is a (complete) lattice. But it says even more: For every complete lattice one can find an isomorphic concept lattice.

Theorem 1.12 ([Wil82])**.** *For every context (G,M,I) the ordered set $\underline{\mathfrak{B}}(G,M,I)$ is a complete lattice in which infima and suprema are given by:*

$$\bigwedge_{t\in T}(A_t,B_t) = \left(\bigcap_{t\in T} A_t, \left(\bigcup_{t\in T} B_t\right)''\right),$$
$$\bigvee_{t\in T}(A_t,B_t) = \left(\left(\bigcup_{t\in T} A_t\right)'', \bigcap_{t\in T} B_t\right).$$

A complete lattice $(L,\le)$ is isomorphic to $\underline{\mathfrak{B}}(G,M,I)$ if and only if there are mappings $\widetilde{\gamma}: G \to L$ and $\widetilde{\mu}: M \to L$ such that $\widetilde{\gamma}(G)$ is supremum-dense in $(L,\le)$, $\widetilde{\mu}(M)$ is infimum-dense in $(L,\le)$ and gIm is equivalent to $\widetilde{\gamma}g \le \widetilde{\mu}m$ for all $g \in G$ and all $m \in M$. In particular, $(L,\le) \cong \underline{\mathfrak{B}}(L,L,\le)$.

It is practical to use the *reduced labelling* instead of writing next to each concept its extent and intent, which would overload the diagram. The object concept γg is the least concept containing g in its extent and dually μm is the greatest concept containing m in its intent. Therefore, it is sufficient to label the object concepts with the corresponding objects, and the attribute concepts with the respective attributes. We can find the extent and intent of any concept in the following way: The extent is formed by collecting all objects which are located at the circle of the concept and which can be reached by descending line paths from the concept. The intent consists of all attributes located at the concept and along ascending line paths.

Example 1.13. The concept lattice of the anorexia context is displayed in Figure 1.2. For instance, the first concept from Example 1.11 is the concept labelled "superficial" and the second one is between the concepts labelled "superficial" and "talkative". Further, we can also read from the concept lattice the attributes every object has and the objects to which any attribute belongs. For instance, take the brother-in-law and follow the lines going up from his object concept. We see that he is confident, dutiful, hearty, attentive, talkative, superficial and ambitious.

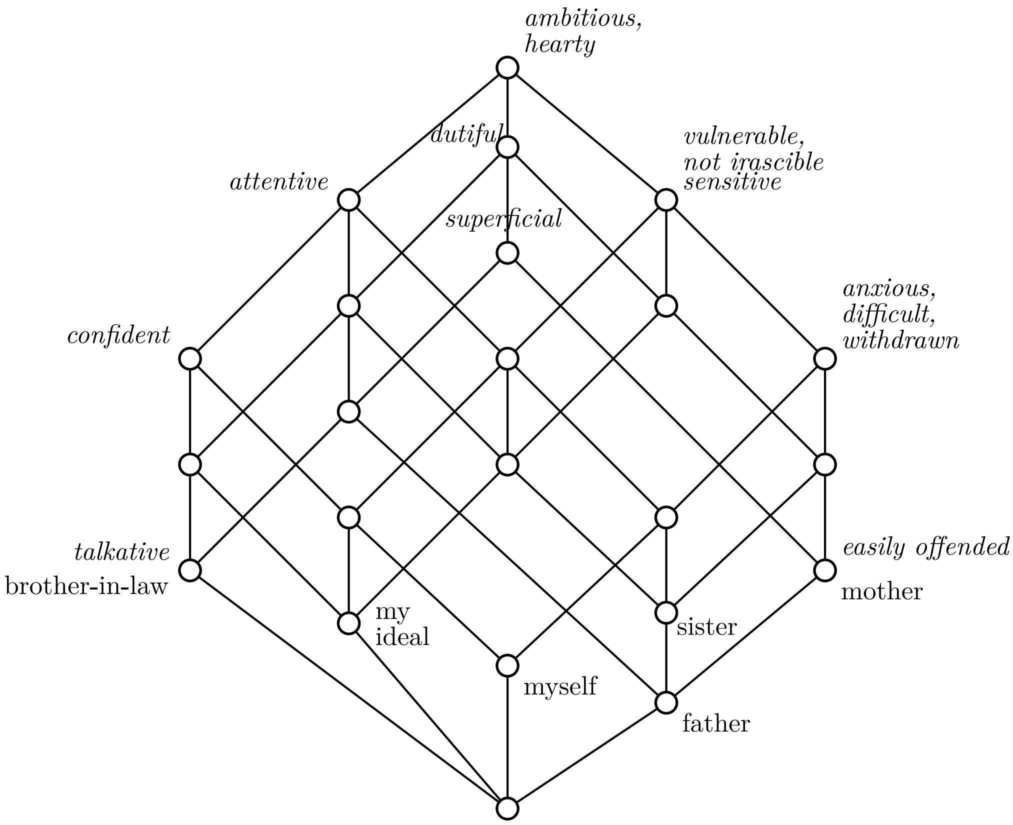

Figure 1.2.: Concept lattice of the anorexia nervosa context

There is an efficient way of computing all concepts of a given formal context, namely by means of the NextClosure algorithm ([Gan84, GW96]). The algorithm enumerates all concepts in the so-called *lectic order* ([Gan84, GW96]). This order is linear on the power set of the attribute set, and it assures that every closure is computed only once. The algorithm is implemented in different freewares, for instance in ConExp[3] and Conexp-clj[4].

[3]http://conexp.sourceforge.net/
[4]http://daniel.kxpq.de/math/conexp-clj/

These also contain numerous tools and methods of Formal Concept Analysis including *attribute implications, attribute exploration*, which we will become acquainted with in the next subsections.

There are context manipulations which do not change the structure of the concept lattice. Usually these operations are carried out before starting the computation of the concept lattice. Such an operation is **clarification**, which reduces the number of objects and attributes. All objects that have equal rows in the context may be removed except one. For attribute clarification we look for equal columns.

Definition 1.14. A context (G, M, I) is called **clarified**, if for any objects $g, h \in G$ from $g' = h'$ it always follows that $g = h$ and, correspondingly, $m' = n'$ implies $m = n$ for all $m, n \in M$.

Another operation which does not change the structure of the concept lattice is the **reducing**. This allows removing attributes which can be written as combination of other attributes, i.e., for $m \in M$ and $X \subseteq M$ such that $m \notin X$ but $m' = X'$. This means that $\mu m = \bigwedge_{x \in X} \mu x$, i.e., the set $\mu(M \setminus \{m\})$ is infimum-dense in $\underline{\mathfrak{B}}(G, M, I)$. The information contained in the reducible attribute m is not lost, because we can reproduce it through the attributes from X. Using the Basic Theorem we get:

$$\underline{\mathfrak{B}}(G, M, I) \cong \underline{\mathfrak{B}}(G, M \setminus \{m\}, I \cap (G \times (M \setminus \{m\}))).$$

Definition 1.15. A clarified context (G, M, I) is called **row reduced**, if every object concept is $\bigvee$-irreducible, and **column reduced**, if every attribute concept is $\bigwedge$-irreducible. A context, which is both row and column reduced, is **reduced**.

Full rows, objects g with $g' = M$, and full columns, attributes m with $m' = G$, are always reducible. When dealing with a finite context, we can remove simultaneously reducible objects and attributes, because in this case every element of the concept lattice is the join of $\bigvee$-irreducible and the meet of $\bigwedge$-irreducible elements. There is another possibility for the reduction, namely via the **arrow relation**:

Definition 1.16. For a context (G, M, I), an object $g \in G$ and an attribute $m \in M$ we write:

$$\begin{aligned} g \swarrow m &:\Longleftrightarrow \begin{cases} g \not{I}\, m \text{ and} \\ \text{if } g^I \subseteq h^I \text{ and } g^I \neq h^I, \text{ then } h\, I\, m, \end{cases} \\ g \nearrow m &:\Longleftrightarrow \begin{cases} g \not{I}\, m \text{ and} \\ \text{if } m^I \subseteq n^I \text{ and } m^I \neq n^I, \text{ then } g\, I\, n, \end{cases} \\ g \swarrow\!\!\!\nearrow m &:\Longleftrightarrow g \swarrow m \text{ and } g \nearrow m. \end{aligned}$$

We may write the arrow relation into the cross table, because it refers to pairs (g, m) that are not contained in the incidence relation. To perform a reducing by means of the arrow relation one proceeds as follows: Insert the arrow relation into the clarified cross table and then remove rows and columns which do not contain a double arrow.

There are various operations which allow the construction of new contexts from given ones. We list just a few of them, others can be found in [GW96]. Let $\mathbb{K} := (G, M, I)$,

$\mathbb{K}_1 := (G_1, M_1, I_1)$ and $\mathbb{K}_2 := (G_2, M_2, I_2)$ be formal contexts. We define the following contexts

$$\begin{aligned}
\mathbb{K}^c &:= (G, M, (G \times M) \setminus I) && \text{the } \textbf{complementary context} \text{ to } \mathbb{K}, \\
\mathbb{K}^d &:= (M, G, I^{-1}) && \text{the } \textbf{dual context} \text{ to } \mathbb{K}, \\
\mathbb{K}_1 \mid \mathbb{K}_2 &:= (G, \dot{M}_1 \cup \dot{M}_2, \dot{I}_1 \cup \dot{I}_2) && \text{the } \textbf{apposition} \text{ of } \mathbb{K}_1 \text{ and } \mathbb{K}_2, \\
& && \text{if } G = G_1 = G_2, \\
\frac{\mathbb{K}_1}{\mathbb{K}_2} &:= (\dot{G}_1 \cup \dot{G}_2, M, \dot{I}_1 \cup \dot{I}_2) && \text{the } \textbf{subposition} \text{ of } \mathbb{K}_1 \text{ and } \mathbb{K}_2, \\
& && \text{if } M = M_1 = M_2
\end{aligned}$$

where $\dot{G}_j := \{j\} \times G_j$, $\dot{M}_j := \{j\} \times M_j$ and $\dot{I}_j := \{((j, g), (j, m)) \mid (g, m) \in I_j\}$ for $j \in \{1, 2\}$.

In practice most data tables do not have the form of cross tables and hence cannot be described directly by formal contexts. In the following we will show how Formal Concept Analysis deals with such data sets.

Definition 1.17. A **many-valued context** (G, M, W, I) consists of sets G, M and W and a relation $I \subseteq G \times M \times W$ such that

$$(g, m, w) \in I \text{ and } (g, m, v) \in I \text{ always imply } w = v.$$

The elements of G are called **objects**, the ones of M **(many-valued) attributes** and those of W **values**. Then, $(g, m, w) \in I$, also denoted by $m(g) = w$, is read "the attribute m has the value w for the object g".

We want to obtain the concepts of the many-valued context. Therefore, we first transform it into a one-valued context through the *conceptual scaling*:

Definition 1.18. A **scale** for the attribute m of a many-valued context is a (one-valued) context $S_m := (G_m, M_m, I_m)$ with $m(G) \subseteq G_m$. The objects of a scale are called **scale values**, the attributes are called **scale attributes**.

The concepts of the derived one-valued context are considered to be the concepts of the many-valued context for this scaling. The scaling process is not uniquely determined and depends on the interpretation of the many-valued attributes. Through scaling we answer question like: Should the values of a many-valued attribute be evaluated as being mutually exclusive? As a hierarchy? The answers are given by the chosen scales. Formally, every context can be used as a scale. However, the notion of "scale" will be used just for contexts that have a clear conceptual structure.

To obtain the corresponding one-valued context for (G, M, W, I), we "somehow" join together the different scales. This can also be done through various methods. The simplest one is the **plain scaling**, where the object set G remains unchanged and every attribute $m \in M$ is replaced by the scale attributes of the scale S_m. In this way, the attribute set of the derived context becomes the disjoint union of the attribute sets of the scales.

Definition 1.19. If (G, M, W, I) is a many-valued context and S_m, $m \in M$ are scale contexts, then the **derived context with respect to the plain scaling** is the context (G, N, J) with

$$N := \bigcup_{m \in M} \dot{M}_m,$$

where $\dot{M}_m := \{m\} \times M_m$ and

$$g\ J\ (m, n) :\Longleftrightarrow m(g) = w \text{ and } w\, I_m\, n.$$

The most frequently used scales are the following elementary ones for $n \in \mathbb{N}$:

- The **nominal scale** $\mathbb{N}_n := (n, n, =)$ is used whenever the attribute values mutually exclude each other. Through the nominal scale there is a partition of the objects into extents. The classes correspond to the values of the attributes. For the special case that $n = 2$, we call $\mathbb{N}_n$ the **dichotomic scale**;
- the **(one-dimensional) ordinal scale** $\mathbb{O}_n := (n, n, \leq)$ scales many-valued attributes, the values of which are ordered and the "stronger" attribute values imply the "weaker" ones. The attribute values form a chain of extents, which can be interpreted as a hierarchy;
- often in questionnaires the possible answers are opposite and one can also choose a statement with intermediate value. The order of the characteristics is bipolar. In such cases using **(one-dimensional) interordinal scales** $\mathbb{I}_n := (n, n, \leq) \mid (n, n, \geq)$ can be very helpful. The concept extents of the interordinal scale are the intervals of the characteristics.

1.3.1. Attribute Implications

The formal context and the corresponding concept lattice are two different representation methods of the same issue. Considering implications gives us a third interesting illustration. By investigating attribute implications we study the possible attribute combinations, the attribute logic, of the respective situation. In addition, we can also reconstruct the concept lattice with the help of these implications.

An **attribute implication** in M is denoted by $A \to B$, where $A, B \subseteq M$. First of all we are interested in the implications of a given formal context.

Definition 1.20. A subset $T \subseteq M$ **respects** an implication $A \to B$ if $A \nsubseteq T$ or $B \subseteq T$. Further, T **respects a set** $\mathcal{L}$ of implications if T respects every single implication in $\mathcal{L}$. $A \to B$ **holds** in a set $\{T_1, T_2, ...\}$ of subsets if each of the subsets T_i respects the implication $A \to B$. An implication $A \to B$ **holds in a context** (G, M, I) if it holds in the system of object intents. If this is the case, we also call $A \to B$ **an implication of the context** (G, M, I) or say that, within the context (G, M, I), A is a **premise of** B.

It is easy to observe that an implication $A \to B$ holds in (G, M, I) if and only if $B \subseteq A''$ which is equivalent to $A' \subseteq B'$. Further, the implication $A \to B$ holds if and only if $A \to m$ holds for every $m \in B$.[5] Another way to describe that the implication $A \to m$ holds, is

[5] We write $A \to m$ instead of $A \to \{m\}$.

to say that the infimum of the attribute concepts corresponding to the attributes in A must be less or equal than the attribute concept of m, i.e., $\bigwedge\{\mu n \mid n \in A\} \leq \mu m$. This observation helps us to read the implications from the concept lattice: $A \to m$ holds if the concept denoted by m is above the infimum of all concepts denoted by an n from A.

Example 1.21. Take another look at the concept lattice from Figure 1.2. For instance, we have the following two implications:

$$\begin{aligned} \text{confident} &\to \text{attentive, ambitious, hearty,} \\ \text{attentive, superficial} &\to \text{dutiful, ambitious, hearty.} \end{aligned}$$

One should keep in mind that the implications hold in the context built by the anorexia nervosa patient. Other patients might have different implications in their contexts.

We can observe that there are many implications holding in even a small context like the anorexia nervosa. In the following we will see that it is not necessary to save all the implications of a context. For every context there exists a set of implications from which all the implications of the context can be derived. But first we have to answer the question: When does an implication follow from another implication?

Definition 1.22. An implication $A \to B$ **follows (semantically)** from a set $\mathcal{L}$ of implications if each subset of M respecting $\mathcal{L}$ also respects $A \to B$. A set $\mathcal{L}$ of implications of (G, M, I) is called **complete** if every implication of (G, M, I) follows from $\mathcal{L}$. Further, $\mathcal{L}$ is called **non-redundant** if none of the implications follows from the others.

Guigues and Duquenne ([GD86]) have developed a method for obtaining these complete and non-redundant implication bases, provided the attribute set M is finite. First we need the definition of a *pseudo-intent*:

Definition 1.23. $P \subseteq M$ is called a **pseudo-intent** of (G, M, I) if and only if $P \neq P''$ and $Q'' \subseteq P$ holds for every pseudo-intent $Q \subsetneq P$.

Theorem 1.24. *The set of implications*

$$\mathcal{L} := \{P \to P'' \mid P \text{ pseudo-intent}\}$$

is complete, non-redundant and minimal (w.r.t. its cardinality). We call this set the ***Duquenne-Guigues-base*** *or the* ***stem base****.*

1.3.2. Attribute Exploration

Attribute exploration, as introduced in [Gan84], is a tool for knowledge discovery by interactive determination of the implications holding between a given set of attributes. This method is especially useful when the examples, objects having the considered attributes, are infinite, hardly to enumerate or (partially) unknown. With the examples (possibly none) of the user's knowledge the object set of the context is built step-by-step. The user is asked whether "some" implication holds. If the answer is affirmative, the implication is added to the stem base and the next implication is considered. If, however, the implication is false, the user has to provide a counterexample. This method assumes that the user can distinguish between true and false implications and that he can provide counterexamples for false implications. This is a crucial point since once a decision was taken about the validity of an implication, the choice cannot be reversed. Therefore, the counterexamples may not contradict the so-far confirmed implications. The procedure ends when all implications of the current stem base hold in general. The result of the attribute exploration is a set of implications which are true in general for the attributes under consideration and a representative set of examples for the whole theory.

The following proposition justifies why we do not have to reconsider the already confirmed implications:

Proposition 1.25 ([GW96]). *Let $\mathbb{K}$ be a context and $P_1, P_2, \ldots, P_n$ be the first n pseudo-intents of $\mathbb{K}$ with respect to the lectic order. If $\mathbb{K}$ is extended by an object g the object intent g' of which respects the implications $P_i \to P_i''$, $i \in \{1, \ldots, n\}$, then $P_1, P_2, \ldots, P_n$ are also the lectically first n pseudo-intents of the extended context.*

Attribute exploration was successfully applied in both theoretical and practical research domains. On the one hand it facilitated the discovery of implications between properties of mathematical structures, see for example [Sac01, KPR06, RK10]. On the other hand it was also used in real-life scenarios, for instance in civil engineering ([EKSW99]), chemistry ([BN97]), information systems ([Stu99]), etc.

There are also further variants of attribute exploration, for instance *attribute exploration with background knowledge* for the case that the user knows in advance some implications between the attributes that hold ([Gan99, Stu96]). Another possibility is to perform *concept exploration* as presented in [Wil87]. By replacing the implications with Horn clauses from predicate logic one obtains the so-called *rule exploration* developed in [Zic91].

1.4. Triadic Concept Analysis

The triadic approach to Formal Concept Analysis was introduced by Wille and Lehmann in [LW95]. Triadic Concept Analysis is based upon Peirce's pragmatic philosophy with his three so-called universal categories. These categories can be interpreted as *the quality of feeling, the reaction as an element of the phenomenon* and *the medium, between a Second and a First* or as *representation as an element of the phenomenon*, respectively ([Pei31, LW95]). This raises the question of how can these three categories be formulated in the language of Formal Concept Analysis. Obviously a dyadic relation is not sufficient to express the three categories. However, a triadic relation is!

We start by giving a brief introduction to the mathematical foundation of Triadic Concept Analysis, as presented in [LW95, Wil95], and answer by this the question about Peirce's three categories.

The underlying structure of Triadic Concept Analysis is the *triadic context* which contains the information about the part of the world we aim to analyse. It is defined as follows:

Definition 1.26. A **triadic context** (shortly **tricontext**) is a quadruple (G, M, B, Y) where G, M and B are sets and Y is a ternary relation between them, i.e. $Y \subseteq G \times M \times B$. The elements of G, M and B are called **objects, attributes** and **conditions**, respectively, and $(g, m, b) \in Y$ is read: "the object g has the attribute m under the condition b".

Until explicitly said otherwise, $\mathbb{K}$ denotes a tricontext in this section. For a better handling of the three sets we denote a tricontext by $\mathbb{K} = (K_1, K_2, K_3, Y)$.

Definition 1.27. A **triadic concept** (shortly **triconcept**) of a tricontext (K_1, K_2, K_3, Y) is defined as a triple (A_1, A_2, A_3) with $A_i \subseteq K_i$ $(i = 1, 2, 3)$ that is maximal with respect to component-wise set inclusion in satisfying $A_1 \times A_2 \times A_3 \subseteq Y$. For a triconcept (A_1, A_2, A_3), the components A_1, A_2 and A_3 are called the **extent**, the **intent** and the **modus** of (A_1, A_2, A_3), respectively. We denote by $\mathfrak{T}(\mathbb{K})$ the set of all triconcepts of $\mathbb{K}$.

Small tricontexts can be represented by three-dimensional cross tables. An example can be seen in Figure 1.3. Pictorially, a triconcept (A_1, A_2, A_3) is a maximal rectangular box full of crosses in the three-dimensional cross table representation of (K_1, K_2, K_3, Y). This fact already follows from the definition of triconcepts, as for any triconcept (A_1, A_2, A_3) and for any subsets $X_i \subseteq K_i$ $(i = 1, 2, 3)$ with $X_1 \times X_2 \times X_3 \subseteq Y$, the containments $A_1 \subseteq X_1, A_2 \subseteq X_2$ and $A_3 \subseteq X_3$ always imply $(A_1, A_2, A_3) = (X_1, X_2, X_3)$.

As in the dyadic case we wish for operators that associate the components of triconcepts with each other. For the triadic case these operators are more technical and can be defined in various ways, as we will see in the following.

Definition 1.28. For $\{i, j, k\} = \{1, 2, 3\}$ with $j < k$ and for $X \subseteq K_i$ and $Z \subseteq K_j \times K_k$, the $(-)^{(i)}$**-derivation operators** are defined by

$$\begin{aligned} X \mapsto X^{(i)} &:= \{(a_j, a_k) \in K_j \times K_k \mid (a_i, a_j, a_k) \in Y \text{ for all } a_i \in X\}, \\ Z \mapsto Z^{(i)} &:= \{a_i \in K_i \mid (a_i, a_j, a_k) \in Y \text{ for all } (a_j, a_k) \in Z\}. \end{aligned}$$

These derivation operators correspond to the derivation operators of the dyadic contexts defined by $\mathbb{K}^{(i)} := (K_i, K_j \times K_k, Y^{(i)})$, where

$$a_1 Y^{(1)} (a_2, a_3) :\Longleftrightarrow a_2 Y^{(2)} (a_1, a_3) :\Longleftrightarrow a_3 Y^{(3)} (a_1, a_3) :\Longleftrightarrow (a_1, a_2, a_3) \in Y.$$

Due to the structure of tricontexts further derivation operators can be defined for the computation of triconcepts.

Definition 1.29. For $\{i, j, k\} = \{1, 2, 3\}$ and for $X_i \subseteq K_i$, $X_j \subseteq K_j$ and $A_k \subseteq K_k$ the $(-)^{A_k}$**-derivation operators** are defined by

$$\begin{aligned} X_i \mapsto X_i^{A_k} &:= \{a_j \in K_j \mid (a_i, a_j, a_k) \in Y \text{ for all } (a_i, a_k) \in X_i \times A_k\}, \\ X_j \mapsto X_j^{A_k} &:= \{a_i \in K_i \mid (a_i, a_j, a_k) \in Y \text{ for all } (a_j, a_k) \in X_j \times A_k\}. \end{aligned}$$

These derivation operators correspond to the derivation operators of the dyadic contexts defined by $\mathbb{K}^{ij}_{A_k} := (K_i, K_j, Y^{ij}_{A_k})$ where

$$(a_i, a_j) \in Y^{ij}_{A_k} :\Longleftrightarrow (a_i, a_j, a_k) \in Y \text{ for all } a_k \in A_k.$$

Example 1.30. The data from Figure 1.3 contains the rating of users about different services provided by hostels from Seville. The users are taken from three hostel booking websites, namely *hostelworld, hostels* and *hostelbookers*.[6]

[6]The sites can be found under http://www1.hostelworld.com, http://www.hostels.com, and http://www.hostelbookers.com. The data was extracted in June 2010.

	0						1						2					
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
0			×						×	×					×	×		
1		×	×	×		×		×	×	×		×	×	×	×	×		×
2	×	×	×	×		×	×	×	×	×	×	×		×	×	×	×	×
3	×	×		×		×	×	×	×	×	×	×	×	×	×	×	×	×
4	×	×		×		×	×	×	×	×	×	×	×	×	×	×	×	×
5			×	×		×	×	×	×	×	×	×	×	×	×	×		×

Figure 1.3.: Tricontext “Hostels”

Although the example is small, it is paradigmatic from more points of view. First of all, it reflects a triadic setting as the hostels are rated regarding the quality of their services from different points of view. Second, the example reflects the usual setting in Web 2.0 applications: Users provide data about certain topics interactively. And since we are in the triadic setting, the third aspect is that this example is also in accordance with the idea presented in [DW01], because it arises from the report of different persons about the same situation.

We constructed the corresponding tricontext in the following way:

- the object set contains the hostels *Nuevo Suizo, Samay, Oasis Backpacker, One, Ole Backpacker, Garden Backpacker*;
- the attribute set is given by the services *character, safety, location, staff, fun, cleanliness*;
- the conditions are given by the users of the three websites;
- since we have chosen the hostels with best ratings, the attribute values are ☻ (good) or ☺ (excellent) and are considered as tags. In the tricontext we make a cross in the corresponding line of object, attribute and condition if the hostel's service was considered excellent by the users of the platform.

We number the elements of K_1, K_2 and K_3 consecutively, i.e., $K_i \coloneqq \{0, \dots, |K_i|\}$ with $i = 1, 2, 3$.

This tricontext has 18 triconcepts. We omit listing them all because they can be read from the *trilattice*, the triadic counterpart of the concept lattice, displayed in Figure 1.4. But until we learn how to read such diagrams let us first take a look at some triconcepts. For instance, $(\{0, 1, 2, 5\}, 2, K_3)^7$ is a triconcept and it means that the hostels from its intent were rated as excellent regarding location from all users of the three platforms. Yet another example is $(K_1, \{2, 3\}, \{1, 2\})$, meaning that all hostels were rated excellent for their location and safety from the users of hostelworld and hostels.

We have already mentioned that there exists something called a *trilattice* which is the triadic counterpart of a concept lattice. The “ingredients” of concept lattices are the concepts and an order on them. Hence, let us turn our attention to the so-far missing order on the triconcepts.

[7] For better legibility we omit curly brackets around singletones.

The structure on the set of all triconcepts $\mathfrak{T}(\mathbb{K})$ is the set inclusion in each component of the triconcept. There is for each $i \in \{1,2,3\}$ a quasiorder $\lesssim_i$ and its corresponding equivalence relation $\sim_i$ defined by

$$(A_1, A_2, A_3) \lesssim_i (B_1, B_2, B_3) :\Longleftrightarrow A_i \subseteq B_i \text{ and} \tag{1.13}$$
$$(A_1, A_2, A_3) \sim_i (B_1, B_2, B_3) :\Longleftrightarrow A_i = B_i, \tag{1.14}$$

for all $i \in \{1,2,3\}$. These quasiorders satisfy the *antiordinal dependencies* ([Wil95]): For $\{i,j,k\} = \{1,2,3\}$, $(A_1, A_2, A_3) \lesssim_i (B_1, B_2, B_3)$ and $(A_1, A_2, A_3) \lesssim_j (B_1, B_2, B_3)$ imply $(A_1, A_2, A_3) \gtrsim_k (B_1, B_2, B_3)$ for all triconcepts (A_1, A_2, A_3) and (B_1, B_2, B_3) of $\mathbb{K}$. For $i \neq j$, the relation $\sim_i \cap \sim_j$ is the identity on $\mathfrak{T}(\mathbb{K})$, i.e., if we have two components of the triconcept, then the third one is uniquely determined by them. Further, $[(A_1, A_2, A_3)]_i$ denotes the equivalence class of $\sim_i$ containing the triconcept (A_1, A_2, A_3). The quasiorder $\lesssim_i$ induces an order $\leq_i$ on the factor set $\mathfrak{T}(\mathbb{K})/\sim_i$ of all equivalence classes of $\sim_i$ which is characterised by

$$[(A_1, A_2, A_3)]_i \leq_i [(B_1, B_2, B_3)]_i \Longleftrightarrow A_i \subseteq B_i.$$

Thus, $(\mathfrak{T}(\mathbb{K})/\sim_1, \leq_1), (\mathfrak{T}(\mathbb{K})/\sim_2, \leq_2)$ and $(\mathfrak{T}(\mathbb{K})/\sim_3, \leq_3)$ can be identified with the ordered set of all extents, intents and modi of $\mathbb{K}$, respectively. Generally, every ordered set with smallest and greatest element is isomorphic to the ordered set of all extents, intents and modi, respectively, of some tricontext, as shown in [Wil95]. This means that unlike in the dyadic case, the extents, intents and modi, respectively, do not form a closure system in general.

A **triordered set** is defined as a relational structure $(S, \lesssim_1, \lesssim_2, \lesssim_3)$ for which the relations $\lesssim_i$ are quasiorders on S such that $\lesssim_i \cap \lesssim_j \subseteq \gtrsim_k$ for $\{i,j,k\} = \{1,2,3\}$ and $\sim_1 \cap \sim_2 \cap \sim_3 = id_S$ where $\sim_i := \lesssim_i \cap \gtrsim_i$ $(i = 1,2,3)$. Suprema and infima are defined as: For $\{i,j,k\} = \{1,2,3\}$ and $X_i, X_k \subseteq S$, an element u of S is called an ik**-bound** of (X_i, X_k) if $u \gtrsim_i x$ for all $x \in X_i$ and $u \gtrsim_k x$ for all $x \in X_k$. An ik-bound u of (X_i, X_k) is called an ik**-limit** of (X_i, X_k) if $u \gtrsim_j v$ for all ik-bounds v of (X_i, X_k). There exists at most one ik-limit u of (X_i, X_k) with $u \lesssim_k v$ for all ik-limits v of (X_i, X_k) in $(S, \lesssim_1, \lesssim_2, \lesssim_3)$. This element u is called the ik**-join** of (X_i, X_k) and is denoted by $\mathfrak{b}_{ik}(X_i, X_k)$.

A **complete trilattice** is defined as a triordered set $\underline{L} := (L, \lesssim_1, \lesssim_2, \lesssim_3)$ in which the ik-joins exist for all $i \neq k$ in $\{1,2,3\}$ and all pairs of subsets of L. An example of a complete trilattice is given in Figure 1.4, where the graphical representation is explained in details.

An analogous structure to the concept lattice is given by $\underline{\mathfrak{T}}(\mathbb{K}) := (\mathfrak{T}(\mathbb{K}), \lesssim_1, \lesssim_2, \lesssim_3)$ for the triadic setting. Let $\{i,j,k\} = \{1,2,3\}$, let $\mathcal{X}_i$ and $\mathcal{X}_k$ be sets of triconcepts and let $X_i := \bigcup\{A_i \mid (A_1, A_2, A_3) \in \mathcal{X}_i\}$ and $X_k := \bigcup\{A_k \mid (A_1, A_2, A_3) \in \mathcal{X}_k\}$. The ik-join of $(\mathcal{X}_i, \mathcal{X}_k)$ is defined to be the triconcept

$$\begin{aligned} \mathfrak{b}(\mathcal{X}_i, \mathcal{X}_k) &:= (B_1, B_2, B_3) \text{ with} \\ B_i &:= X_i^{X_k X_k}, \\ B_j &:= X_i^{X_k}, \\ B_k &:= (X_i^{X_k, X_k} \times X_i^{X_k})^{(k)}. \end{aligned} \tag{1.15}$$

In order to introduce the *Basic Theorem of Triadic Concept Analysis* ([Wil95]) we still need to present some order-theoretic notions about complete trilattices. To this end, let $\underline{L} := (L, \lesssim_1, \lesssim_2, \lesssim_3)$ be a complete trilattice. An **order filter** of the quasiordered set $(L, \lesssim_i)$ $(i = 1, 2, 3)$ is a subset F of L such that $x \in F$ and $x \lesssim_i y$ always imply $y \in F$. We denote by $\mathcal{F}_i(\underline{L})$ the set of all order filters of $(L, \lesssim_i)$. A **principal filter** of $(L, \lesssim_i)$ is defined by $[x)_i := \{y \in L \mid x \lesssim_i y\}$. Further, a subset $\mathcal{X}$ of $\mathcal{F}_i(\underline{L})$ is called **i-dense** with respect to $\underline{L}$ if each principal filter of $(L, \lesssim_i)$ is the intersection of some order filters from $\mathcal{X}$. The principal filter generated by the triadic concept (A_1, A_2, A_3) in $(\underline{\mathfrak{T}}(\mathbb{K}), \lesssim_i)$ equals $\bigcap_{a_i \in A_i} \{(B_1, B_2, B_3) \in \underline{\mathfrak{T}}(\mathbb{K}) \mid a_i \in B_i\} \in \mathcal{F}_i(\underline{\mathfrak{T}}(\mathbb{K}))$. Therefore, the i-dense set $\kappa_i(K_i)$ of order filters of $(\underline{\mathfrak{T}}(\mathbb{K}), \lesssim_i)$ $(i = 1, 2, 3)$ is given by

$$\kappa_i(a_i) := \{(B_1, B_2, B_3) \in \underline{\mathfrak{T}}(\mathbb{K}) \mid a_i \in B_i\}$$

for $a_i \in K_i$.

Now we are prepared to present the *Basic Theorem of Triadic Concept Analysis*:

Theorem 1.31 ([Wil95]). *Let $\mathbb{K} = (K_1, K_2, K_3, Y)$ be a tricontext. Then, $\underline{\mathfrak{T}}(\mathbb{K})$ is a complete trilattice of $\mathbb{K}$ with the ik-joins given by ($\{i, j, k\} = \{1, 2, 3\}$):*

$$\nabla_{ik}(\mathcal{X}_i, \mathcal{X}_k) := \mathfrak{b}_{ik}(\bigcup\{A_i \mid (A_1, A_2, A_3) \in \mathcal{X}_i\}, \bigcup\{A_k \mid (A_1, A_2, A_3) \in \mathcal{X}_k\}).$$

In general, every complete trilattice $\underline{L} := (L, \lesssim_1, \lesssim_2, \lesssim_3)$ is isomorphic to $\underline{\mathfrak{T}}(\mathbb{K})$ if and only if there exist mappings $\widetilde{\kappa}_i : K_i \to \mathcal{F}_i(\underline{L})$ $(i = 1, 2, 3)$ such that $\widetilde{\kappa}_i(K_i)$ is i-dense with respect to $\underline{L}$ and

$$A_1 \times A_2 \times A_3 \subseteq Y \iff \bigcap_{i=1}^{3} \bigcap_{a_i \in A_i} \widetilde{\kappa}_i(a_i) \neq \varnothing$$

for all $A_1 \subseteq K_1, A_2 \subseteq K_2$ and $A_3 \subseteq K_3$. In particular, we have $\underline{L} \cong \underline{\mathfrak{T}}(L, L, L, Y_{\underline{L}})$ with $Y_{\underline{L}} := \{(x_1, x_2, x_3) \in L^3 \mid (x_1, x_2, x_3)$ is joined$\}$.

The trilattice of the Hostel context is displayed in Figure 1.4. The ordered structures of objects, attributes and conditions are given by Hasse diagrams. On the right and upper part of the figure are the Hasse diagrams of objects and conditions, respectively and on the left part the upside-down Hasse diagram of the attributes. The structure of the triconcepts is given by the 3-net in the centre of the diagram. Each circle in the 3-net represents a triconcept which extent, intent and modus can be read through the discontinuous lines connecting the circle with the three Hasse diagrams. A discontinuous line from the circle to an object means that the extent of the triconcept contains that object and all the objects below it. An analogous statement holds for the modus. Since the Hasse diagram of the attributes is upside-down, the intent contains all the attributes attached to and above the discontinuous line. For instance, take the second circle from the left in the third horizontal line from the top. It corresponds to the triconcept $(\{1, 2, 3, 4, 5\}, \{1, 2, 3, 5\}, \{1, 2\})$. For now we ignore that some circles are drawn larger. (As it will turn out later those correspond to the factors of the triconceptual factorisation.)

In [Bie88], a foundation of the theory of trilattices was presented including, among others, an equation theory for trilattices, completion theorems of triordered sets and trilattices and triadic Galois connections.

Attribute implications were first generalised to the triadic setting in [Bie88]. Due to the nature of tricontexts, there are various ways to define implications. Stronger triadic implications were developed in [GO04]. We will study these later on in connection with our

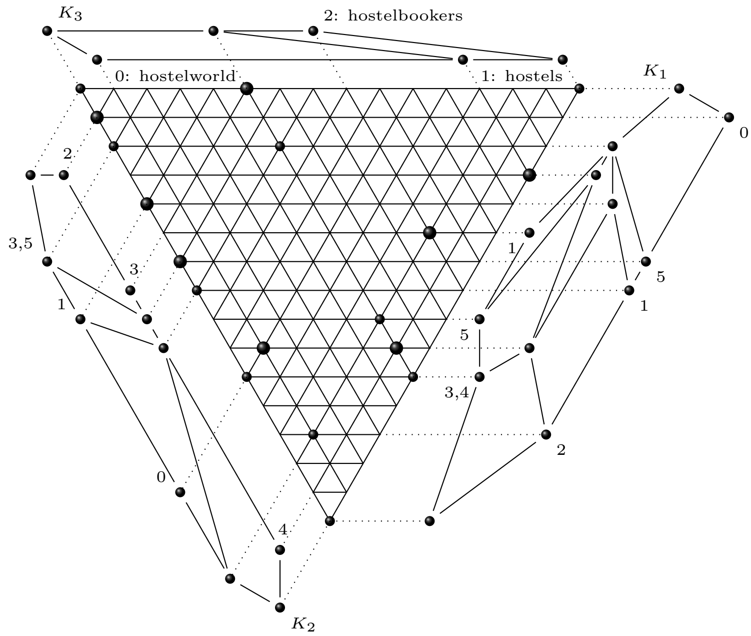

Figure 1.4.: Trilattice of the tricontext "Hostels"

fuzzy-valued triadic setting which we develop in Chapter 7. Here we just give a flavour of the crisp triadic implications developed in [GO04]. For a given discrete tricontext $\mathbb{K} = (K_1, K_2, K_3, Y)$ and for $R, S \subseteq K_2$ and $C \subseteq K_3$, the expression $R \overset{C}{\to} S$ is called *conditional attribute implication*. For $R, S \subseteq K_3$ and $C \subseteq K_2$ the expression $R \overset{C}{\to} S$ is called *attributional condition implication*. Implications of the form $R \to S$ are called *attribute* × *condition implications*, where $R, S \subseteq K_2 \times K_3$.

There are various applications of Triadic Concept Analysis. The most prominent one is the BibSonomy[8] platform which is a social bookmark and publication system. The mining of frequent triconcepts was developed in [JHS+06]. Another important application is Factor Analysis in a triadic setting which we will present in Chapter 3. Pioneer work regarding the investigation of adverse drug reactions with Triadic Concept Analysis can be found in [SBNP11].

1.5. Formal Fuzzy Concept Analysis

There are various approaches to Formal Fuzzy Concept Analysis. A survey on the different methods can be found in [BV05e]. Let us mention just a few. The paper by Burusco and Fuentes-Gonzáles ([BFG75]) seems to be one of the first works that connects Formal Concept Analysis and Fuzzy theory. However, the derivation operators lack some useful properties which hold in the crisp setting. In the independent works of Yahia ([YJ01]) and Krajči ([Kra03]) the extent of the concept is a crisp set and its intent is fuzzy. In [KSVĎ02] the approach is based on alpha-cuts. The author of [Kra04] considers different residuated

[8] http://www.bibsonomy.org/

lattices for objects, attributes and the incidence relation. In [BSZ05] the formal fuzzy concepts are obtained through crisp sets which are evaluated in formal fuzzy contexts.

We will consider the method developed independently by Pollandt ([Pol97]) and Bělohlávek ([Běl02b]) as the standard one. The hedges proved themselves to be useful, as they are able to model fuzzy events both from a fuzzy and a more crisp view. Further, hedges are of avail especially for attribute implications in a fuzzy setting but also for reducing the size of the fuzzy concept lattice. Therefore, we will consider a more general approach to Formal Fuzzy Concept Analysis, namely the one which incorporates residuated lattices and hedges ([BV07]). Definitions and propositions about formal fuzzy contexts, derivation operators and fuzzy concept lattices are taken from [Pol97, Běl02b, BV07].

Definition 1.32. A triple (G, M, I) is called an **L-context** (**formal context with fuzzy attributes**) if $I : G \times M \to L$ is an $\mathbf{L}$-relation between the sets G and M, and L is the support set of some residuated lattice with hedge. Elements from G and M are called **objects** and **attributes**, respectively.

The $\mathbf{L}$-relation I assigns to each $g \in G$ and each $m \in M$ the truth degree $I(g, m) \in L$ to which the object g has the attribute m.

We are interested in the fuzzified version of the derivation operators. In the most general way, one may use different hedges on the objects and on the attributes.

Definition 1.33. For $\mathbf{L}$-sets $A \in \mathbf{L}^G$ and $B \in \mathbf{L}^M$, the **derivation operators** are defined by

$$A^{\uparrow}(m) := \bigwedge_{g \in G} (A(g)^{*_G} \to I(g, m)), \tag{1.16}$$

$$B^{\downarrow}(g) := \bigwedge_{m \in M} (B(m)^{*_M} \to I(g, m)) \tag{1.17}$$

for $g \in G$ and $m \in M$.

Then, $A^{\uparrow}(m)$ is the truth degree of the statement "m is shared by all objects from A", and $B^{\downarrow}(g)$ is the truth degree of "g has all attributes from B". The operators $(-)^{\uparrow}$ and $(-)^{\downarrow}$ form a so-called Galois connection with hedges ([BV07]). If we use for the hedges the identity, then we have the $\mathbf{L}$-Galois connection introduced in Definition 1.7 (page 10). However, if we choose hedges different from the identity, then the situation changes. We will discuss this topic later on.

Definition 1.34. An **L-concept** (**formal fuzzy concept**) is a tuple $(A, B) \in \mathbf{L}^G \times \mathbf{L}^M$ such that $A^{\uparrow} = B$ and $B^{\downarrow} = A$. Then, A is called the **extent** and B the **intent** of (A, B).

We denote the set of all $\mathbf{L}$-concepts of a given context (G, M, I) by $\mathfrak{B}(G^{*_G}, M^{*_M}, I)$. If a hedge is the identity, we omit the superscription. Further, if we consider the same hedge on the objects and attributes, we simply write $(-)^*$ instead of $(-)^{*_G}$ and $(-)^{*_M}$.

Concepts serve for classification. Consequently, the super- and subconcept relation plays an important role. We call (A_1, B_1) a **subconcept** of (A_2, B_2), written $(A_1, B_1) \leq (A_2, B_2)$, if and only if $A_1 \subseteq A_2$ (or, equivalently, $B_1 \supseteq B_2$). Then, (A_2, B_2) is a **superconcept** of (A_1, B_1). The set of all $\mathbf{L}-$concepts ordered by this concept order forms a complete (fuzzy) lattice (with hedge), the so-called **L-concept lattice** (see Theorems 1.36 and 1.39 later on) which is denoted by $\underline{\mathfrak{B}}(G^{*_G}, M^{*_M}, I) := (\mathfrak{B}(G^{*_G}, M^{*_M}, I), \leq)$.

	small (s)	large (l)	far (f)	near (n)
Mercury (Me)	1	0	0	1
Venus (V)	1	0	0	1
Earth (E)	1	0	0	1
Mars (Ma)	1	0	0.5	1
Jupiter (J)	0	1	1	0.5
Saturn (S)	0	1	1	0.5
Uranus (U)	0.5	0.5	1	0
Neptune (N)	0.5	0.5	1	0
Pluto (P)	1	0	1	0

Figure 1.5.: **L**-context about planets ([Běl02b])

Example 1.35. Consider the **L**-context about planets ([Běl02b]) from Figure 1.5. This is our running example on which we will illustrate the notions from Formal Concept Analysis in a fuzzy setting. Using the identity as the hedge, we obtain 38 **L**-concepts with the Łukasiewicz logic and 25 with the Gödel logic. For instance, an **L**-concept is $(\{{}^{0.5}/Ma, {}^{0.5}/U, {}^{0.5}/N, P\}, \{s, f\})$ meaning that Pluto (the example was created, when Pluto was still a planet) is a small and far planet, whereas Mars, Uranus and Neptune belong just partially to the class of small and far planets. Another example of an **L**-concept is $(\{{}^{0.5}/Me, {}^{0.5}/V, {}^{0.5}/E, Ma, {}^{0.5}/U, {}^{0.5}/N, {}^{0.5}/P\}, \{s, {}^{0.5}/f, {}^{0.5}/n\})$. Its meaning is the following: Mars belongs to the class of small, partially far and partially near planets, whereas the other planets from the extent belong just partially to this class. The **L**-concept lattice for the Gödel structure is displayed in Figure 1.6. Such diagrams are to be read in a similar way as their crisp counterparts.

Depending on the chosen hedge, we have different characterisations of the **L**-concept lattice. First, let us use the identity.

Theorem 1.36 ([Pol97, Běl02b])**.** *Let (G, M, I) be an* **L***-context. Then, $\underline{\mathfrak{B}}(G, M, I)$ is a complete lattice in which infima and suprema are given by*

$$\bigwedge_{t\in T}(A_t, B_t) = \left(\bigcap_{t\in T} A_t, \left(\bigcup_{t\in T} B_t\right)^{\downarrow\uparrow}\right),$$

$$\bigvee_{t\in T}(A_t, B_t) = \left(\left(\bigcup_{t\in T} A_t\right)^{\uparrow\downarrow}, \bigcap_{t\in T} B_t\right).$$

A complete lattice $\mathbf{V} = (V, \leq)$ is isomorphic to $\underline{\mathfrak{B}}(G, M, I)$ if and only if there are mappings $\widetilde{\gamma} : G \times L \to V$ and $\widetilde{\mu} : M \times L \to V$ such that $\widetilde{\gamma}(G \times L)$ is supremum-dense in $\mathbf{V}$, $\widetilde{\mu}(M \times L)$ is infimum-dense in $\mathbf{V}$, and $a \otimes b \leq I(g, m)$ is equivalent to $\widetilde{\gamma}(g, a) \leq \widetilde{\mu}(m, b)$ for all $g \in G$, $m \in M$ and $a, b \in L$.

The theorem can be shown either by means of Fuzzy theory ([Běl02b]) or by applying *double-scaling* ([Pol97], we will get there in a minute). In [Běl02b] an even stronger characterisation of **L**-concept lattices was given. In addition to the results of Theorem 1.36

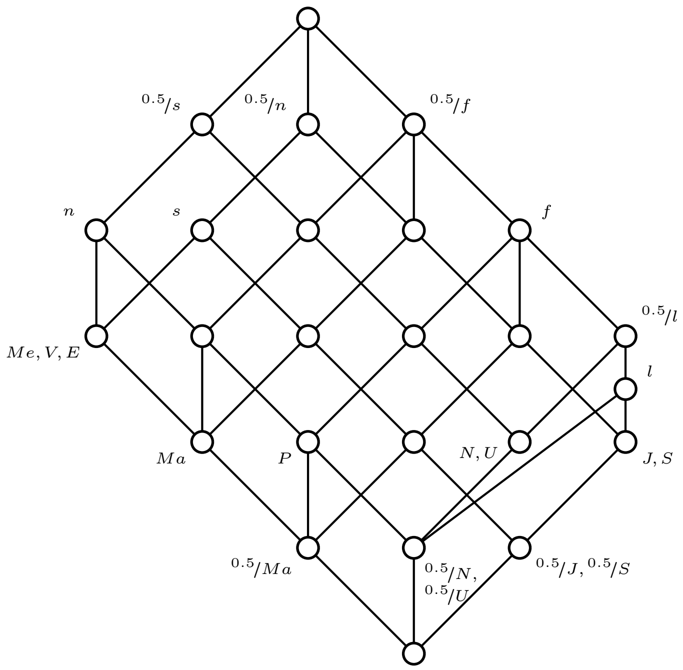

Figure 1.6.: **L**-concept lattice of the context from Figure 1.5 with the Gödel structure

there is a binary **L**-relation $\precsim$ on $\underline{\mathfrak{B}}(G, M, I)$ given by

$$((A_1, B_1) \precsim (A_2, B_2)) := \mathrm{S}(A_1, A_2) \quad (= \mathrm{S}(B_2, B_1)).$$

Together with this relation, $\underline{\mathfrak{B}}(G, M, I)$ is an **L**-ordered set (see Definition 1.2, page 8). For further details we refer the reader to [Běl02b].

Double-scaling is a procedure developed in [Pol97] that transforms an **L**-context into a crisp formal context. The method works as follows: Let (G, M, I) be an **L**-context and define for an **L**-set $A \in \mathbf{L}^G$ the crisp set $A^\square$ by

$$A^\square := \{(g, \nu) \mid g \in G, \nu \in L, \nu \leq A(g)\}. \tag{1.18}$$

Hence, $A^\square \subseteq G^\square := G \times L$. Further, let now $A \subseteq G \times L$ and define for it the **L**-set $A^\diamond$ by

$$A^\diamond(g) := \bigvee\{\nu \in L \mid (g, \nu) \in A\} \tag{1.19}$$

for all $g \in G$. Analogous crisp and **L**-sets can be constructed for attribute sets. For the **L**-relation I between G and M define a crisp incidence relation $I^\square$ between $G^\square$ and $M^\square$ given by

$$(g, \nu)\ I^\square\ (m, \lambda) :\Longleftrightarrow \nu \otimes \lambda \leq I(g, m), \tag{1.20}$$

where $\otimes$ is the multiplication in the residuated lattice **L**. We have the following important result:

Theorem 1.37 ([Pol97]). *Let (G, M, I) be an **L**-context and $(G^\square, M^\square, I^\square)$ the corresponding double-scaled context. Then, $\underline{\mathfrak{B}}(G, M, I) \cong \underline{\mathfrak{B}}(G^\square, M^\square, I^\square)$.*

In the rest of this section we will study the case of general hedges, where the results are taken from [BV07]. If we choose hedges different from the identity, then the size of the **L**-concept lattice is decreasing. The closer we come to the globalisation, the smaller the lattice becomes. For instance, if we replace the identity with the globalisation, then we obtain 17 **L**-concepts with the Łukasiewicz logic for the **L**-context about planets. Some studies about how the size of the **L**-concept lattice changes when applying different hedges can be found in [BV07].

Let us denote by $(-)^{\Uparrow}$ and $(-)^{\Downarrow}$ the derivation operators from (1.16) and (1.17) obtained by using for both hedges the identity, respectively. First note that $A^{\uparrow} = (A^{*_G})^{\Uparrow}$ and $B^{\downarrow} = (B^{*_M})^{\Downarrow}$ hold for arbitrary hedges $(-)^{*_G}$ and $(-)^{*_M}$ and for any **L**-sets $A \in \mathbf{L}^G$ and $B \in \mathbf{L}^M$ in a given **L**-context (G, M, I). However, in general, neither $A \subseteq A^{\uparrow\downarrow}$ nor $B \subseteq B^{\downarrow\uparrow}$ is true. Therefore, we need a different notion of Galois connections for hedges.

Definition 1.38. A **Galois connection with hedges** $(-)^{*_G}$ and $(-)^{*_M}$ between sets G and M is a pair (φ, ψ) of mappings $\varphi : \mathbf{L}^G \to \mathbf{L}^M$ and $\psi : \mathbf{L}^M \to \mathbf{L}^G$ satisfying

$$\mathrm{S}(A^{*_G}, \psi(B)) = \mathrm{S}(B^{*_M}, \varphi(A)),$$
$$\varphi\left(\bigcup_{t\in T} A_t^{*_G}\right) = \bigcap_{t\in T} \varphi(A_t),$$
$$\psi\left(\bigcup_{s\in S} B_s^{*_M}\right) = \bigcap_{s\in S} \psi(B_s)$$

for all $A, A_t \in \mathbf{L}^G$ and $B, B_s \in \mathbf{L}^M$ with $t \in T$ and $s \in S$.

The derivation operators form a Galois connection with hedges.

Denote
$$*_G(L) := \{a^{*_G} \mid a \in L\} \text{ and } *_M(L) := \{a^{*_M} \mid a \in L\}.$$

Now we may find an analogous isomorphism to the one given by Theorem 1.37 between crisp concept lattices and concept lattices with hedges: Every concept lattice with hedges $\underline{\mathfrak{B}}(G^{*_G}, M^{*_M}, I)$ is isomorphic to the crisp concept lattice $\underline{\mathfrak{B}}(G \times *_G(L), M \times *_M(L), I^{\square})$, where $I^{\square}$ is given by (1.20). We may characterise the concept lattices with hedges as follows:

Theorem 1.39 ([BV07])**.** *Let (G, M, I) be an **L**-context. Then, $\underline{\mathfrak{B}}(G^{*_G}, M^{*_M}, I)$ is a complete lattice in which infima and suprema are given by*

$$\bigwedge_{t\in T}(A_t, B_t) = \left(\left(\bigcap_{t\in T} A_t\right)^{\uparrow\downarrow}, \left(\bigcup_{t\in T} B_t^{*_M}\right)^{\downarrow\uparrow}\right),$$
$$\bigvee_{t\in T}(A_t, B_t) = \left(\left(\bigcup_{t\in T} A_t^{*_G}\right)^{\uparrow\downarrow}, \left(\bigcap_{t\in T} B_t\right)^{\downarrow\uparrow}\right).$$

A complete lattice $\mathbf{V} = (V, \leq)$ is isomorphic to $\underline{\mathfrak{B}}(G^{*_G}, M^{*_M}, I)$ if and only if there are mappings $\widetilde{\gamma} : G \times *_G(L) \to V$ and $\widetilde{\mu} : M \times *_M(L) \to V$ such that $\widetilde{\gamma}(G \times *_G(L))$ is supremum-dense in $\mathbf{V}$, $\widetilde{\mu}(M \times *_M(L))$ is infimum-dense in $\mathbf{V}$, and $a \otimes b \leq I(g, m)$ is equivalent to $\widetilde{\gamma}(g, a) \leq \widetilde{\mu}(m, b)$ for all $g \in G$, $m \in M$ and $a, b \in L$.

Theorem 1.36 is basically a special case of this one, i.e., take the identity for $(-)^{*_G}$ and $(-)^{*_M}$. Once again, we may use the double-scaling to show this theorem.

We have collected the necessary theory to compute **L**-concepts. But how to do this efficiently? One may either double-scale the **L**-context and use the softwares from the crisp case to compute the **L**-concept lattice, or do it straightforward in the fuzzy setting. For the latter we have a fuzzified version of the NextClosure algorithm which was presented in [Běl02a]. For this we need the *fuzzy lectic order* ([Běl02a]) which is defined as follows: Let $L = \{l_0 < l_1 < \cdots < l_n = 1\}$ be the support set of some residuated lattice and $M = \{1, 2, \ldots, m\}$ the attribute set of the context. For $(x,i),(y,j) \in M \times L$, we write

$$(x, l_i) \leq (y, l_j) :\Longleftrightarrow (x < y) \text{ or } (x = y \text{ and } l_i \geq l_j).$$

For $B \in \mathbf{L}^M$ and $(x,i) \in M \times L$ we define

$$B \oplus (x,i) := ((B \cap \{1,2,\ldots,x-1\}) \cup \{{}^{l_i}/x\})^{\downarrow\uparrow}.$$

Furthermore, for $B, C \in \mathbf{L}^M$ define

$$B <_{(x,i)} C :\Longleftrightarrow B \cap \{1,\ldots,x-1\} = C \cap \{1,\ldots,x-1\} \text{ and } B(x) < C(x) = l_i. \tag{1.21}$$

We say that B is **lectically smaller** than C, written $B < C$, if $B <_{(x,i)} C$ for some (x,i) satisfying (1.21). As in the crisp case, we have that $B^+ := B \oplus (x,i)$ is the least intent which is greater than a given B with respect to $<$ and (x,i) is the greatest with $B <_{(x,i)} B \oplus (x,i)$ (for details we refer to [Běl02a]).

1.5.1. Attribute Implications

Attribute implications in a fuzzy setting were mainly developed and studied in a series of papers by Bělohlávek and Vychodil, for instance in [BV06a, BCV04]. They also appear in [Pol97], however there a crisp-like approach to pseudo-intents is presented. We will comment on this later on.

A **fuzzy attribute implication** (over the attribute set M) is an expression $A \Rightarrow B$, where $A, B \in \mathbf{L}^M$. The verbal meaning of $A \Rightarrow B$ is: "if it is (very) true that an object has all attributes from A, then it also has all attributes from B". The notions "being very true", "to have an attribute", and the logical connective "if-then" are determined by the chosen residuated lattice **L**. For an **L**-set $N \in \mathbf{L}^M$ of attributes, the degree $||A \Rightarrow B||_N \in L$ to which $A \Rightarrow B$ **holds in** N is defined by

$$||A \Rightarrow B||_N := S(A,N)^* \to S(B,N).$$

If N is the **L**-set of all attributes of an object g, then $||A \Rightarrow B||_N$ is the truth degree to which $A \Rightarrow B$ holds for g. For a set $\mathcal{N} \subseteq \mathbf{L}^M$, the degree $||A \Rightarrow B||_{\mathcal{N}} \in L$ to which the implication $A \Rightarrow B$ **holds in** $\mathcal{N}$ is defined by

$$||A \Rightarrow B||_{\mathcal{N}} := \bigwedge_{N \in \mathcal{N}} ||A \Rightarrow B||_N.$$

For an **L**-context (G,M,I), let $I_g \in \mathbf{L}^M$ $(g \in G)$ be an **L**-set of attributes such that $I_g(m) = I(g,m)$ for each $m \in M$. Clearly, I_g corresponds to the row labelled g in (G,M,I).

The degree $||A \Rightarrow B||_{(G,M,I)} \in L$ to which $A \Rightarrow B$ holds in (each row of) $\mathbb{K} = (G, M, I)$ is defined by

$$||A \Rightarrow B||_{\mathbb{K}} = ||A \Rightarrow B||_{(G,M,I)} := ||A \Rightarrow B||_{\mathcal{N}}, \tag{1.22}$$

where $\mathcal{N} := \{I_g \mid g \in G\}$. Denote by

$$\mathrm{Int}(G^*, M, I) := \{B \in \mathbf{L}^M \mid (A, B) \in \mathfrak{B}(G^*, M, I) \text{ for some } A\}$$

the set of all intents of $\mathfrak{B}(G^*, M, I)$. Since $N \in \mathbf{L}^M$ is the intent of some concept if and only if $N = N^{\downarrow\uparrow}$, we have that $\mathrm{Int}(G^*, M, I) = \{N \in \mathbf{L}^M \mid N = N^{\downarrow\uparrow}\}$. Further, the degree $||A \Rightarrow B||_{\mathfrak{B}(G^*,M,I)} \in L$ to which $A \Rightarrow B$ holds in (the intents of) $\mathfrak{B}(G^*, M, I)$ is defined by

$$||A \Rightarrow B||_{\mathfrak{B}(G^*,M,I)} := ||A \Rightarrow B||_{\mathrm{Int}(G^*,M,I)}.$$

Lemma 1.40 ([BV05c])**.** *Let (G, M, I) be an $\mathbf{L}$-context. Then,*

$$||A \Rightarrow B||_{(G,M,I)} = ||A \Rightarrow B||_{\mathfrak{B}(G^*,M,I)} = ||A \Rightarrow B||_{\mathrm{Int}(G^*,M,I)} = S(B, A^{\downarrow\uparrow}),$$

for each fuzzy attribute implication $A \Rightarrow B$.

Example 1.41. As we have already seen, implications in a fuzzy setting do not simply hold or not, they hold with some truth value. To illustrate this fact, consider once again the $\mathbf{L}$-context $\mathbb{K}$ given in Figure 1.5 (page 25). We use the Łukasiewicz logic and the identity for the hedge. Then, $||n \Rightarrow s||_{\mathbb{K}} = 0.5$ and hence $||n \Rightarrow {}^{0.5}/s||_{\mathbb{K}} = 1$. These two implications mean that near implies small only partially, or expressed in other words, near implies partially small. Another implications is $||\{{}^{0.5}/l, {}^{0.5}/n\} \Rightarrow \{l, f\}||_{\mathbb{K}} = 0.5$. If we replace the identity with the globalisation, the truth values of the implications change. For instance, we have $||n \Rightarrow s||_{\mathbb{K}} = 1$ and $||\{{}^{0.5}/l, {}^{0.5}/n\} \Rightarrow \{l, f\}||_{\mathbb{K}} = 1$.

1.5.2. Non-redundant Bases of Fuzzy Attribute Implications

Due to the large number of implications in a fuzzy and even in a crisp formal context, one is interested in the stem base of the implications. The stem base problem for the fuzzy setting was studied in [BCV04, BV05c, BV05b]. Neither the existence nor the uniqueness of it is guaranteed in general for a given $\mathbf{L}$-context. How these problems can be overcome is the topic of the rest of this subsection. For a more detailed description we refer the reader to the before cited papers.

Let T be a set of fuzzy attribute implications. An $\mathbf{L}$-set of $N \in \mathbf{L}^M$ is called a **model** of T if $||A \Rightarrow B||_N = 1$ for each $A \Rightarrow B \in T$. The set of all models of T is denoted by $\mathrm{Mod}(T)$, i.e.,

$$\mathrm{Mod}(T) := \{N \in \mathbf{L}^M \mid N \text{ is a model of } T\}.$$

The degree $||A \Rightarrow B||_T \in L$ to which $A \Rightarrow B$ **semantically follows** from T is defined by $||A \Rightarrow B||_T := ||A \Rightarrow B||_{\mathrm{Mod}(T)}$. The set T of fuzzy attribute implications is called **complete** (in (G, M, I)) if $||A \Rightarrow B||_T = ||A \Rightarrow B||_{(G,M,I)}$ for each $A \Rightarrow B$. If T is complete and no proper subset of T is complete, then T is called a **non-redundant base**.

Theorem 1.42 ([BCV04])**.** *T is complete if and only if* $\mathrm{Mod}(T) = \mathrm{Int}(G^*, M, I)$.

As in the crisp case, a non-redundant base of a given $\mathbf{L}$-context can be obtained through the pseudo-intents.

Definition 1.43. $\mathcal{P} \subseteq \mathbf{L}^M$ is called a **system of pseudo-intents** if for each $P \in \mathbf{L}^M$ we have:

$$P \in \mathcal{P} \Longleftrightarrow (P \neq P^{\downarrow\uparrow} \text{ and } ||Q \Rightarrow Q^{\downarrow\uparrow}||_P = 1 \text{ for each } Q \in \mathcal{P} \text{ with } Q \neq P).$$

In case we choose for $(-)^*$ the globalisation, then, due to the properties of the globalisation, the formalisation of pseudo-intents from Definition 1.43 becomes: $\mathcal{P} \subseteq \mathbf{L}^M$ is a system of pseudo-intents if

$$P \in \mathcal{P} \Longleftrightarrow (P \neq P^{\downarrow\uparrow} \text{ and } Q^{\downarrow\uparrow} \subseteq P \text{ for each } Q \in \mathcal{P} \text{ with } Q \subsetneq P). \tag{1.23}$$

Note that this was the only kind of system of pseudo-intents studied in [Pol97].

Theorem 1.44 ([BV05c])**.** *Let $\mathbf{L}$ be a finite residuated lattice with globalisation. Then, for each (G, M, I) with finite M there is a unique system of pseudo-intents $\mathcal{P}$ given by* (1.23)*.*

Note that this theorem does not hold for general hedges.

Theorem 1.45 ([BV05c])**.** *$T := \{P \Rightarrow P^{\downarrow\uparrow} \mid P \in \mathcal{P}\}$ is complete and non-redundant, called the **stem base**. If $(-)^*$ is the globalisation, then T is unique and minimal.*

For $Z \in \mathbf{L}^M$ and for each natural number n we put

$$Z^{T^*} := Z \cup \bigcup \{B \otimes S(A, Z)^* \mid A \Rightarrow B \in T \text{ and } A \neq Z\},$$
$$Z^{T^*_n} := \begin{cases} Z, & n = 0, \\ (Z^{T^*_{n-1}})^{T^*}, & n \geq 1, \end{cases}$$

where $B \otimes S(A, Z)^*$ is computed component-wise. We define an operator cl_{T} on $\mathbf{L}$-sets in M by

$$\mathrm{cl}_{\mathrm{T}}(Z) := \bigcup_{n=0}^{\infty} Z^{T^*_n}. \tag{1.24}$$

Theorem 1.46 ([BCV04])**.** *If $(-)^*$ is the globalisation, then cl_{T} is an $\mathbf{L}^*$-closure operator and $\{\mathrm{cl}_{\mathrm{T}}(Z) \mid Z \in \mathbf{L}^M\} = \mathcal{P} \cup \mathrm{Int}(X^*, Y, I)$.*

According to this theorem, if $(-)^*$ is the globalisation, then we can obtain all intents and all pseudo-intents of a given $\mathbf{L}$-context by computing the fix points of cl_{T}. In [BCV04] an algorithm for the computation of all intents and all pseudo-intents in lectic order was proposed.

However, there is also an alternative way for obtaining the stem base of an $\mathbf{L}$-context using the globalisation. This method was developed in [BV05d]. Basically one takes the simple-scaled context $(G, M \times L, I^{\square})$ and computes its stem base. Afterwards, the so-obtained crisp implications are transformed into fuzzy ones. These represent a complete set of implications of (G, M, I). Of course, the set is redundant, because there are not just implications between attributes but also between truth values. Note that this method only works in case we are using the globalisation.

globalisation	identity 1	identity 2
$n \Rightarrow s,$ $f, {}^{0.5}/n \Rightarrow l,$ ${}^{0.5}/l \Rightarrow f,$ $l, f \Rightarrow {}^{0.5}/n,$ ${}^{0.5}/s, {}^{0.5}/n \Rightarrow s, n,$ $s, f, {}^{0.5}/l \Rightarrow l, n.$	$n \Rightarrow {}^{0.5}/s,$ $f, {}^{0.5}/n \Rightarrow {}^{0.5}/l,$ $l \Rightarrow f, {}^{0.5}/n,$ $s, {}^{0.5}/l, {}^{0.5}/f \Rightarrow {}^{0.5}/n.$	$n \Rightarrow {}^{0.5}/s,$ $f, {}^{0.5}/n \Rightarrow {}^{0.5}/l,$ $l, {}^{0.5}/n \Rightarrow f,$ $s, {}^{0.5}/l, {}^{0.5}/f \Rightarrow {}^{0.5}/n.$

Figure 1.7.: Different stem bases of the **L**-context given in Figure 1.5

Example 1.47. Consider once again the **L**-context $\mathbb{K}$ given in Figure 1.5 (page 25). In this example we work with the Łukasiewicz logic. If we use the globalisation, then we obtain the unique stem base given in the first column of the table from Figure 1.7. The other two columns show the two different stem bases of the same **L**-context obtained by using the identity.

Now, if we choose a general hedge in the residuated lattice, things get messy. The computation of the systems of pseudo-intents for general hedges was studied in [BV05b]. For an **L**-context (G, M, I) we compute the following:

$$V := \{P \in \mathbf{L}^M \mid P \neq P^{\downarrow\uparrow}\}, \tag{1.25}$$

$$E := \{(P, Q) \in V \times V \mid P \neq Q \text{ and } ||Q \Rightarrow Q^{\downarrow\uparrow}||_P \neq 1\}. \tag{1.26}$$

In case of a non-empty V, $\mathbf{G} := (V, E \cup E^{-1})$ is a graph. For $Q \in V$, $\mathcal{P} \subseteq V$ define the following subsets of V:

$$\begin{aligned}\operatorname{Pred}(Q) &:= \{P \in V \mid (P, Q) \in E\},\\ \operatorname{Pred}(\mathcal{P}) &:= \bigcup_{Q \in \mathcal{P}} \operatorname{Pred}(Q).\end{aligned}$$

Described verbally, $\operatorname{Pred}(Q)$ is the set of all elements from V which are predecessors of Q (in E). Further, $\operatorname{Pred}(\mathcal{P})$ is the set of all predecessors of any $Q \in \mathcal{P}$.

We will compute the systems of pseudo-intents through maximal independent sets. Therefor, the following results are useful:

Theorem 1.48 ([BV05b])**.** (i) *Let $\varnothing \neq \mathcal{P} \subseteq \mathbf{L}^M$. If $V \setminus \mathcal{P} = \operatorname{Pred}(\mathcal{P})$, then $\mathcal{P}$ is a maximal independent set in $\mathbf{G}$.*

(ii) *Let $\mathcal{P} \subseteq \mathbf{L}^M$. Then, $\mathcal{P}$ is a system of pseudo-intents if and only if $V \setminus \mathcal{P} = Pred(\mathcal{P})$.*

It is well-known that the maximal independent sets of a graph can be efficiently enumerated in lexicographic order with only polynomial delay between the output of two successive independent sets ([JYP88]). In [DS11] it was shown that the pseudo-intents cannot be enumerated in lexicographic order with polynomial delay unless P = NP. These two results do not contradict each other because they address different issues. The first one is encountered when we enumerate the maximal independent sets of the graph **G** which is the input of the corresponding algorithm. These sets correspond to the systems of pseudo-intents. Whereas the result from [DS11] takes as input a formal context enumerating its pseudo-intents.

Example 1.49. We start with a very simple example. Let $(\{g\},\{a,b\},I)$ be an **L**-context with $I(g,a) = 0.5$ and $I(g,b) = 0$. Further, we use the three-element Łukasiewicz logic with $(-)^*$ being the identity. First, we compute V as given by (1.25) and obtain

$$V = \{\{^{0.5}/a, ^{0.5}/b\}, \{^{0.5}/b\}, \{\}, \{^{0.5}/a, b\}, \{b\}, \{a\}\}.$$

Afterwards, we compute the binary relation E as given by (1.26). It is displayed in Figure 1.8. Considering the undirected diagram of Figure 1.8 we obtain the graph **G**. There, we have four maximal independent sets, namely

$$\mathcal{P}_1 = \{\{\}, \{^{0.5}/a, b\}, \{a\}\}, \qquad \mathcal{P}_2 = \{\{^{0.5}/b\}, \{a\}\},$$
$$\mathcal{P}_3 = \{\{b\}, \{a\}\}, \qquad \mathcal{P}_4 = \{\{^{0.5}/a, ^{0.5}/b\}, \{a\}\}.$$

The sets $\mathcal{P}_1$ and $\mathcal{P}_3$ do not satisfy condition ii) of Theorem 1.48 and are therefore not

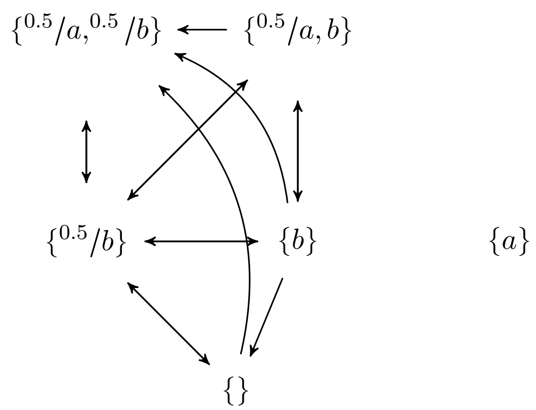

Figure 1.8.: Binary relation E

systems of pseudo-intents. However, $\mathcal{P}_2$ and $\mathcal{P}_4$ do satisfy this condition and hence they are systems of pseudo-intents yielding the stem bases displayed in Figure 1.9.

$\mathcal{T}_2$		$\mathcal{T}_4$	
(1)	$^{0.5}/b \Rightarrow a$	(3)	$^{0.5}/a, ^{0.5}/b \Rightarrow a$
(2)	$a \Rightarrow ^{0.5}/b$	(4)	$a \Rightarrow ^{0.5}/b$

Figure 1.9.: Stem bases

It is still an open question which conditions have to be satisfied in order to have a system of pseudo-intents for a given **L**-context using a general hedge. Further, if it exists, when is it unique? And finally, when is the system of pseudo-intents minimal?

Part I.

Factor Analysis

Introduction to Part I

Factor Analysis, in particular Principal Component Analysis, is a popular technique for analysing metric data. It allows for complexity reduction, representing a large part of the given data by a preferably low number of unobserved "latent" attributes. In the forthcoming chapters we will translate this statistical method into the language of Formal Concept Analysis, leaving the statistical part aside.

In Chapter 2 we first build on some already established connections between Factor Analysis of binary data and Formal Concept Analysis. The factors of such an analysis are called *Boolean factors* and are actually formal concepts. Afterwards, we develop the so-called *many-valued factorisations* that group the Boolean factors into well-structured families. These families are given by the conceptual standard scales of Formal Concept Analysis. Since this is a new technique, we apply it on different real-world data sets. The results show that these many-valued factorisations are serious competitors to the latent attributes from ordinary Factor Analysis.

In Chapter 3 we turn our attention to the factorisation of triadic data with *triadic Boolean factors*. Such a generalisation is wishful due to two aspects. First, triadic data has numerous applications and various data reduction methods were generalised to the triadic case. Second, as we will see in Chapter 2, Boolean factors yield factorisations with the smallest possible number of elements and thus we expect the same from their triadic counterparts. This speculation comfirms to be true in Chapter 4.

Up until Chapter 4 we deal with the formal concept analytical formalisation of the Factor Analysis problem presenting different mathematical results. We claim that those factorisations are easy to understand and may keep up with ordinary data reduction techniques. This claim confirms to be true in Chapter 4. There we show that the so-called *Hierarchical Classes Analysis*, that was developed for applications in personality organisation and implicit belief systems, is reducible to our conceptual factorisation both for dyadic and triadic data. Further, we develop the fuzzy variant of Hierarchical Classes Analysis.

2 Conceptual Factorisations

Factor Analysis is a commonly used complexity reduction technique for metric data. Recently a factor analytical approach was discussed for qualitative data, for data that can be represented by a formal context, and a nice strategy for finding so-called *Boolean factors* was found. The further development of this area is the topic of Section 2.1.

However, such Boolean factors have limited expressiveness due to their unary nature. It can hardly be expected that much of a complex data set can be captured by only a few Boolean factors. But even a large factorisation may be useful provided the factors are conceptually “well behaved” and can be grouped into well-structured families, which then may be interpreted as *many-valued factors*. These are given by the conceptual standard scales of Formal Concept Analysis. In Section 2.2 we focus on the case of (one-dimensional) ordinal scales. Note that Section 2.1 and 2.2 are mainly based on [GG12].

Since the many-valued factors are newly developed, they have to be tested. It is yet unclear whether there is more behind them than a robust mathematical theory. Thus, in Section 2.3 we apply the method on three different real-world data sets. We find out that the many-valued factors are serious competitors to latent attributes, being easily interpretable and relatively small in number.

In [Běl08] and [BV09a] the Boolean factors from Section 2.1 were generalised to the fuzzy setting. Since in the upcoming chapters we will build on these results, we briefly overview them in Section 2.4.

An overall conclusion of the chapter is given in its last section.

2.1. Boolean Factors

The starting point of a Boolean matrix factorisation/Boolean Factor Analysis ([MME90]) is an $n \times m$ binary matrix X which is decomposed into the *Boolean matrix product* $A \circ B$ of an $n \times k$ binary matrix A and a $k \times m$ binary matrix B such that k is as small as possible. Such k is called by some authors the **Schein rank** ([Kim82]) of X and is denoted by $\rho(X)$.

The **Boolean matrix product** $A \circ B$ is defined by

$$(A \circ B)_{ij} := \bigvee_{l=1}^{k} A_{il} \cdot B_{lj} \tag{2.1}$$

for $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$, where $\bigvee$ denotes the maximum and $\cdot$ the usual product.

Decompositions of binary matrices, not necessarily through the Boolean matrix product, have been widely investigated with different methods and approaches. For a brief survey see for instance [BV10a]. Among them, there are also approaches based on methods that were initially developed for real-valued matrices, for example [SSU03, TT06, ZV06]. There one starts with a binary matrix and ends up with two real-valued matrices as its decomposition. However, it is well-known (see for instance [MMG⁺08, TMGM06]) that applying decomposition methods that were designed for real-valued data to binary data distorts the meaning of the data and yields results that are difficult to interpret. On the other hand, decomposition methods based on the Boolean matrix product are interpreted in a straightforward way and are therefore preferable ([MMG⁺08]).

Apparently, the first to link Formal Concept Analysis with the decomposition of binary matrices were Snášel and Keprt ([KS04, Kep06]). Their work was pushed forward by Bělohlávek and Vychodil (see e.g. [BV10a]). However, as we will see in Chapter 4, there is also another approach, *Hierarchical Classes Analysis* by De Boeck and Rosenberg, that is reducible to the formal concept analytical one. For better compatibility with the language of Formal Concept Analysis we slightly deviate from these authors' terminology.

In this section we mainly present the results from [GG12] and recall some from [BV10a]. We start by introducing the factorisation of formal contexts through formal concepts. Thereafter, we show how such a factorisation can be found, and we study some of its properties. Having the theoretical background, we focus afterwards on algorithmic issues and on so-called *approximate factorisations*, which explain only roughly the data. Such factorisations are useful due to their small number of factors that explain a large portion of the data. The developed notions are illustrated on the data from the anorexia nervosa therapy.

Definition 2.1. A **factorisation** of a formal context (G, M, I) consists of formal contexts (G, F, I_{GF}) and (F, M, I_{FM}) such that

$$g\, I\, m \iff g\, I_{GF}\, f \text{ and } f\, I_{FM}\, m \quad \text{for some } f \in F.$$

The elements of F are called **Boolean factors**, (G, F, I_{GF}) and (F, M, I_{FM}) are the **factorisation contexts**. We write

$$(G, M, I) = (G, F, I_{GF}) \circ (F, M, I_{FM})$$

to indicate a factorisation.

As one can easily see, the object-attribute relation of (G, M, I) is expressed through the factorisation contexts: An object g has an attribute m if and only if there is a factor f such that f *applies* to g and m is a *particular manifestation* of f. Thus, the first factorisation context represents the relationship between objects and factors and the second one between factors and attributes.

To each factorisation there corresponds a **factorising family**

$$\{(A_f, B_f) \mid f \in F\}$$

given by

$$A_f := \{g \in G \mid g\ I_{GF}\ f\} \quad \text{and} \quad B_f := \{m \in M \mid f\ I_{FM}\ m\}.$$

Such families are easy to characterise: A family $\{(A_f, B_f) \mid f \in F\}$ is a factorising family of (G, M, I) if and only if

$$I = \bigcup_{f \in F} A_f \times B_f. \tag{2.2}$$

Expressed in words this says that the factorising families are precisely those families of preconcepts of (G, M, I) that cover all incidences. As an obvious consequence we get that each factorisation is uniquely determined by its factorising family.

Each preconcept can be enlarged to a formal concept (though not uniquely). Each factorising family therefore may, without increasing the number of Boolean factors, be made to a factorising family of concepts. Such a factorisation will be called **conceptual**.

A key question of the abovementioned investigations concerned finding *optimal* factorisations, i.e., those with the number of factors being the Schein rank of (G, M, I). The task of filling a given relation $I \subseteq G \times M$ by as few as possible "rectangles" $A \times B \subseteq I$ had been studied earlier under the name **set dimension** (see [GW96] and the literature cited there), and is known to be difficult. There is a close connection to the **2-dimension** of the complementary context, which is the number of atoms of the smallest Boolean algebra that admits an order embedding of the concept lattice of the complementary context $\underline{\mathfrak{B}}(G, M, G \times M \setminus I)$. Indeed, the following proposition is an easy consequence of [BV10a] and the dimension theory in [GW96]:

Proposition 2.2 ([GG12]). *The smallest possible number of Boolean factors of* (G, M, I) *equals the 2-dimension of* $\underline{\mathfrak{B}}(G, M, G \times M \setminus I)$.

Thus, we find the connection between the conceptual factorisation and Boolean Factor Analysis with its matrix product defined in (2.1). By replacing the crosses with ones and the blanks with zeros we obtain a Boolean matrix from (G, M, I). We denote this matrix by I. Its size is $n \times m$, where $n := |G|$ and $m := |M|$. In a similar fashion we build the corresponding binary matrices A_F and B_F to the factorisation contexts, i.e., we set

$$(A_F)_{il} := \begin{cases} 1, & i \in f_l^{I_{GF}}, \\ 0, & i \notin f_l^{I_{GF}}, \end{cases} \qquad (B_F)_{jl} := \begin{cases} 1, & j \in f_l^{I_{FM}}, \\ 0, & j \notin f_l^{I_{FM}}, \end{cases}$$

for all $l \in \{1, \ldots, k\}, i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$. Choosing $k := |F|$ indeed provides us with the Schein rank of I, as shown in the previous proposition.

Example 2.3. Consider the formal context about the anorexia nervosa therapy from Section 1.3 that we display here once again in Figure 2.1. For now ignore that some crosses are larger. One may find a factorisation of it with five Boolean factors in several ways. One way is displayed in Figure 2.2. The first two factors contain the characteristics with negative connotation. The third is a general factor as it contains the fewest attributes and

Anorexia	m_1 : vulnerable	m_2 : withdrawn	m_3 : confident	m_4 : dutiful	m_5 : hearty	m_6 : difficult	m_7 : attentive	m_8 : easily offended	m_9 : not irascible	m_{10}: anxious	m_{11}: talkative	m_{12}: superficial	m_{13}: sensitive	m_{14}: ambitious
g_1 : myself	×	**×**	**×**		×	**×**	×		×	**×**			×	×
g_2 : my ideal	**×**		×	**×**	**×**		×		**×**				**×**	×
g_3 : father	×	×		×	×	×	×	×	×	×		×	×	×
g_4 : mother	×	×		×	×	×		**×**	×	×		×	×	×
g_5 : sister	×	×		×	×	×	×		×	×			×	×
g_6 : brother-in-law			×	×	×		×				×	×		×

Figure 2.1.: Anorexia nervosa context from Figure 1.1 with tight crosses

	f^1	f^2	f^3	f^4	f^5
g_1		×			×
g_2			×		×
g_3	×	×	×		
g_4	×	×			
g_5		×	×		
g_6			×	×	

$\circ$

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}
f^1	×	×		×	×	×		×	×	×		×	×	×
f^2	×	×			×	×			×	×			×	×
f^3				×	×		×							×
f^4			×	×	×		×				×	×		×
f^5	×		×		×		×		×				×	×

Figure 2.2.: A conceptual factorisation of the formal context from Figure 2.1.

applies to almost each family member. It may express the caring of the family members towards other persons and duties. The fourth factor is typical for the brother-in-law. He has his “own factor” because he is the only talkative person. The last factor shows the common ground between the patient and her ideal.

We said before that the optimal factorisation is of size five. This can be seen from the concept lattice of the complementary context to the anorexia nervosa context displayed in Figure 2.3. Indeed, the black shaded substructure does not fit into a Boolean algebra with four atoms. Thus, the whole lattice cannot be embedded into such an algebra. Therefore, the smallest possible number of Boolean factors is five.

Such a factorisation is not very pleasing, since there is a *trivial* factorisation with six Boolean factors, as detailed in a minute.

Note that in a conceptual factorisation of (G, M, I) the second factorisation context is determined by the first. Indeed, we get from $B_f = A_f^I$ that

$$f \ I_{FM} \ m \iff m \in A_f^I = (f^{I_{GF}})^I.$$

Proposition 2.4 ([GG12]). *For any conceptual factorisation with factor set F the dual attribute order of (G, F, I_{GF}) is the same as the object order of (F, M, I_{FM}).*

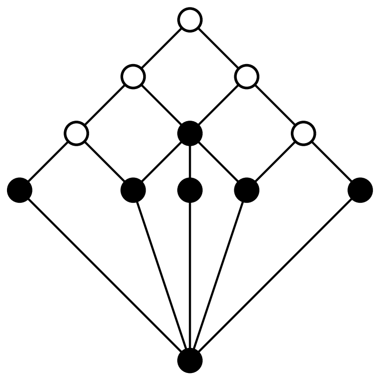

Figure 2.3.: Concept lattice of the complementary context to the anorexia nervosa context from Figure 2.1

Proof. If $f_1^{I_{GF}} \subseteq f_2^{I_{GF}}$, then $A_{f_1} = f_1^{I_{GF}} \subseteq f_2^{I_{GF}} = A_{f_2}$ and thus $B_{f_1} = A_{f_1}^I \supseteq A_{f_2}^I = B_{f_2}$. Therefore, $f_1^{I_{FM}} \supseteq f_2^{I_{FM}}$. The converse is similar. □

In general the condition given in this proposition is not sufficient. The next theorem characterises the conceptual factorisation contexts. It turns out that a conceptual factorisation links each factor context to the complementary context of the other.

Theorem 2.5 ([GG12]). *(G, F, I_{GF}) and (F, M, I_{FM}) are the factorisation contexts of a conceptual factorisation if and only if*

1. *all intents of (G, F, I_{GF}) are extents of $(F, M, F \times M \setminus I_{FM})$, and*
2. *all extents of (F, M, I_{FM}) are intents of $(G, F, G \times F \setminus I_{GF})$.*

Proof. Start with a conceptual factorisation, and recall that

$$g\, I\, m \iff g^{I_{GF}} \cap m^{I_{FM}} \neq \emptyset,$$

which is equivalent to

$$g \not\!I\, m \iff g^{I_{GF}} \subseteq F \setminus m^{I_{FM}}.$$

For arbitrary $g \in G$ we ask if the object intent $g^{I_{GF}}$ is the intersection of attribute extents of $(F, M, F \times M \setminus I_{FM})$. Suppose not. Then there must be some $f \in F$ which is contained in all attribute extents of $(F, M, F \times M \setminus I_{FM})$ that contain $g^{I_{GF}}$, but which does not belong to $g^{I_{GF}}$. From $g \notin f^{I_{GF}} = A_f$ we infer that $g \notin B_f^I = (f^{I_{FM}})^I$. Consequently, there must be some $m \in f^{I_{FM}}$ with $g \not\!I\, m$. This, as stated above, is equivalent to $g^{I_{GF}} \subseteq F \setminus m^{I_{FM}}$. But then $F \setminus m^{I_{FM}}$ is an attribute extent of $(F, M, F \times M \setminus I_{FM})$ containing $g^{I_{GF}}$ and not f, a contradiction.

For the converse direction suppose that the conditions are satisfied. In order to show that the factorisation is conceptual, we need to prove that $(f^{I_{GF}})^I \subseteq f^{I_{FM}}$ and dually $(f^{I_{FM}})^I \subseteq f^{I_{GF}}$ hold for all $f \in F$. Now assume that $m \notin f^{I_{FM}}$, which is the same as $f \notin m^{I_{FM}}$. Since $m^{I_{FM}}$ is an extent of (F, M, I_{FM}), there must be, according to the second condition, an object intent of $(G, F, G \times F \setminus I_{GF})$ containing $m^{I_{FM}}$, but not f. In other words, there must be an object $g \in f^{I_{GF}}$ such that $g^{I_{GF}} \cap m^{I_{FM}} = \varnothing$. Thus $g \not I m$, i.e., $m \notin (f^{I_{GF}})^I$. $\square$

Note that every formal context (G, M, I) is a conceptual factorisation context in its **trivial factorisations**, one of them being

$$(G, M, I) = (G, M, I) \circ (M, M, \rightarrow),$$

where $m \rightarrow n :\Longleftrightarrow n \in m^{II}$. The other is defined dually. Theorem 2.5 therefore imposes no restriction on single factorisation contexts.

The following two results are straightforward. However none of them holds generally in ordinary Factor Analysis due to the statistical tools used in that framework. Thus, these propositions might be interpreted as robustness results for the conceptual factorisations. They say that the conceptual factorisation is invariant under dualisation and clarification of the formal context in the sense detailed in the propositions.

Proposition 2.6. *The conceptual factorisation contexts* (G, F, I_{GF}) *and* (F, M, I_{FM}) *of* (G, M, I) *are the dual contexts to the conceptual factorisation contexts* (F, G, I_{GF}^{-1}) *and* (M, F, I_{FM}^{-1}) *of* (M, G, I^{-1})*, respectively.*

Proof. According to the *Duality Principle of Concept Lattices* ([GW96]) for any formal context (G, M, I) it holds that $\underline{\mathfrak{B}}(M, G, I^{-1}) \cong \underline{\mathfrak{B}}(G, M, I)^d$, where the isomorphism is given by $(A, B) \mapsto (B, A)$. Hence, if we find a conceptual factorisation of (G, M, I) with the corresponding factorising family $\{(A_f, B_f) \mid f \in F\}$, then $\{(B_f, A_f) \mid f \in F\}$ is the factorising family of (M, G, I^{-1}). Indeed, from $I = \bigcup_{f \in F} A_f \times B_f$ it trivially follows that $I^{-1} = \bigcup_{f \in F} B_f \times A_f$. $\square$

Proposition 2.7. *The clarified conceptual factorisation contexts of* (G, M, I) *are the same as the conceptual factorisation contexts of the clarified context of* (G, M, I).

Proof. From [GW96] we know that for each context (G, M, I) its clarified context can be associated as follows:

$$(G/\ker\gamma, M/\ker\mu, I^\circ),$$

where $\ker\gamma$ is the equivalence relation on G given by

$$(g, h) \in \ker\gamma :\Longleftrightarrow \gamma g = \gamma h.$$

The definition of $\ker\mu$ is correspondingly. The equivalence classes of $\ker\gamma$ are the objects of the context $(G/\ker\gamma, M/\ker\mu, I^\circ)$ and those of $\ker\mu$ are the attributes. The incidence relation I° is given by

$$([g]_{\ker\gamma}, [m]_{\ker\mu}) \in I^\circ :\Longleftrightarrow g\, I\, m.$$

Let F be the factor set of a conceptual factorisation of the formal context (G, M, I). For any $g, h \in G$ we trivially obtain $g^{I_{GF}} = h^{I_{GF}}$ from $g^I = h^I$. An analogous remark holds for attributes. Further, let F_{ker} be the factor set and let $\{(A_f/\ker\gamma, B_f/\ker\mu) \mid f \in F\}$ be the factorising family of $(G/\ker\gamma, M/\ker\mu, I^\circ)$ given by

$$A_f/\ker\gamma := \{g \in G/\ker\gamma \mid g\, I_{GF}\, f\} \quad \text{and} \quad B_f/\ker\mu := \{m \in M/\ker\mu \mid f\, I_{FM}\, m\}.$$

Now suppose that $F \setminus F_{\text{ker}} \neq \emptyset$, i.e., there is some factor $l \in F$ such that $l \notin F_{\text{ker}}$. Thus, we have

$$I^\circ = \bigcup_{f \in F \setminus \{l\}} A_f / \ker\gamma \times B_f / \ker\mu,$$

which implies

$$I = \bigcup_{f \in F \setminus \{l\}} A_f \times B_f,$$

a contradiction. The converse is similar. □

We have already mentioned that finding an optimal factorisation is difficult. Indeed, it has been shown in [BV10a] that the set-basis problem is reducible to conceptual factorisation. As it is well-known that the set basis problem is NP-complete, it follows that conceptual factorisation is NP-hard. Let us present this result briefly.

In the set basis problem we have a collection $S = \{S_1, \ldots, S_n\}$ of sets $S_i \subseteq \{1, \ldots, m\}$ and a natural number k. The problem consists in answering the question whether there is a collection $C = \{C_1, \ldots, C_k\}$ of subsets $C_l \subseteq \{1, \ldots, m\}$ such that for every S_i there is a subset $D_i \subseteq \{C_1, \ldots, C_k\}$ with $\bigcup D_i = S_i$. The corresponding optimisation problem lies in finding a C of least cardinality, that satisfies the above conditions, for a given S. The set basis problem is easily reducible to the factorisation problem: Given S, one defines an $n \times m$ binary matrix I by $I_{ij} = 1$ if and only if $j \in S_i$. One can show that I can be decomposed into the Boolean matrix product $A \circ B$ of an $n \times k$ and a $k \times m$ binary matrix A and B, respectively, if and only if C_l $(l = 1, \ldots, k)$ and D_i, defined by $j \in C_l$ iff $B_{lj} = 1$ and $C_l \in D_i$ iff $A_{il} = 1$, are a solution to the set basis problem given by S. Then, I corresponds to a formal context and the matrices A and B to its factorisation contexts. Therefore, the problem of finding a conceptual factorisation with $|F| = k$ as small as possible is NP-hard and the corresponding decision problem is NP-complete.

In our setting the universe U to be covered corresponds to the incidence relation of the context and the family $\mathcal{S}$ of subsets of the universe U that is used for finding a cover corresponds to the set of all concepts. We are looking for $\mathcal{C} \subseteq \mathcal{S}$ with the smallest number of sets such that $\bigcup \mathcal{C} = U$.

Although the factorisation problem is difficult, there exists a greedy approximation algorithm for the set covering optimisation problem which achieves an approximation ratio $\leq \ln(|U|) + 1$, see [CLRS01]. In [BV10a] a greedy approximation algorithm for the conceptual factorisation was presented. It selects formal concepts which cover the most part of the incidence relation until there is nothing left to be covered.

In [BV10a] it was shown that mandatory factors, i.e., those which are present in each factorisation, correspond to concepts which are both object and attribute concepts. In [GG12] we took a step forward: An incident object-attribute pair $(g, m) \in I$ is called **tight** if and only if there is no pair $(h, n) \in I$ such that

$$[\gamma h, \mu n] \subsetneq [\gamma g, \mu m].$$

It is easy to see that for a conceptual factorisation it suffices to cover the tight incidences by formal concepts, because the other incidences will then automatically be covered. This is sometimes useful for the computation of the set dimension, since it reduces the size of the corresponding set covering problem. The tight incidences correspond to the double arrows of the complementary context. The larger crosses in Figure 2.1 match the tight incidences.

As we have already seen, even for a small formal context numerous factors are needed in a conceptual factorisation. Thus, generally, Factor Analysis seeks for an *approximate* factorisation. Such a factorisation fits just partially the data and is adequate whenever the the user is roughly interested in the information contained in the data. The approximate factorisations are of two kinds: with *negative discrepancies*, where not the entire incidence relation of the context is covered by the factorisation; and with *positive discrepancies*, where some blank entries are "covered" with crosses, i.e., one explains more than there is. Of course, one may combine the two kinds of approximate factorisations.

Conceptual factorisations with negative discrepancies were studied in [BV10a]. It turned out that in practice it is often the case that exact factorisations may require a large number of factors, however a relative small number of them covers most part of the incidence relation.

Example 2.8. We have already found out that the anorexia nervosa context can be factorised using five Boolean factors that are displayed in Figure 2.2. However, the factor set $\mathcal{F}_1 := \{f^2\}$ already covers $55,17\%$ of the incidence relation! So, by using just one factor we can explain more than a half of the data. The factor sets $\mathcal{F}_2 := \{f^2, f^3\}$ and $\mathcal{F}_3 := \{f^2, f^3, f^5\}$ cover $75,86\%$ and $86,21\%$ of the incidence relation, respectively. The factors f^1 and f^4 tighten also the factorisation, however not as drastically as the other ones. Thus, the more factors we add to the factorisation, the tighter it becomes.

A *dense rectangle* in a formal context (G, M, I) is a tuple (A, B) with $A \subseteq G$ and $B \subseteq M$ such that $A \times B \nsubseteq I$, i.e., it is a rectangle in the cross table which allows blank entries. The computation of dense rectangles was studied in [GV93, BV06b], however the factorisation problem was not considered. Here we will use the terminology from [GV93] and denote by z_d the number of blank entries in a dense rectangle $d := (A, B)$, i.e., $z_d := |(A \times B) \setminus I|$. Further, we denote by $\rho_{z_d} := z/(|A| \cdot |B|)$ the *density* of d. Now, a **factorisation with positive discrepancies** is a set $\mathcal{D}$ of dense rectangles such that

$$I \subseteq \bigcup_{(A,B) \in \mathcal{D}} A \times B.$$

Obviously, the lower the density is that we allow for the dense rectangles, the more precise the factorisation becomes.

Example 2.9. The following three dense rectangles yield an approximate factorisation with positive discrepancies of the formal context from Figure 2.1:

$$\begin{aligned}
&(\{g_1, g_3, g_4, g_3\}, M \setminus \{m_8, m_{11}, m_{12}\}),\\
&(\{g_2, g_6\}, \{m_1, m_3, m_4, m_5, m_7, m_9, m_{13}, m_{14}\}),\\
&(\{g_3, g_4, g_6\}, \{m_4, m_5, m_8, m_{11}, m_{12}, m_{14}\}).
\end{aligned}$$

Their densities are $0.11, 0.19$ and 0.17, respectively, and they cover 11 blank entries.

Now we come to the last issue of this section, which addresses the transformations between the space of attributes and the space of factors. Given a formal context (G, M, I) and a conceptual factorisation of it with factorisation contexts (G, F, I_{GF}) and (F, M, I_{FM}), one is naturally interested in how to transform a description of a given object in terms of attributes into a description of the same object in terms of factors. That is, one asks for transformations between the attribute space and the factor space. Such mappings were studied in [BV10a] and were utilised in [Out10] for improving classification of binary data. For a better compatibility with the language of Formal Concept Analysis, we deviate once again from the notation used by the authors and define $\varphi : \mathfrak{P}(M) \to \mathfrak{P}(F)$ and $\psi : \mathfrak{P}(F) \to \mathfrak{P}(M)$ by

$$\varphi(P) := \{f \in F \mid f^{I_{FM}} \subseteq P\}, \tag{2.3}$$

$$\psi(S) := \bigcup_{f \in S} f^{I_{FM}}, \tag{2.4}$$

for $P \in \mathfrak{P}(M)$ and $S \in \mathfrak{P}(F)$. In [BV10a] it was shown that

$$\varphi(g^I) = g^{I_{GF}} \quad \text{and} \quad \psi(g^{I_{GF}}) = g^I$$

for any $g \in G$. Thus, φ maps the rows of (G, M, I) to the rows of (G, F, I_{GF}), and ψ maps the rows of (G, F, I_{GF}) to the rows of (G, M, I). One may show that these mappings form an isotone Galois connection.

2.2. Ordinal Factors

The set F of Boolean factors may be large and should then, for the sake of better interpretability, be divided into conceptually meaningful subsets. An ordinal factor, for instance, simply represents a chain of Boolean factors.

Proposition 2.10 ([GG12])**.** *If (G, F, I_{GF}) and (F, M, I_{FM}) are conceptual factorisation contexts and $E \subseteq F$, then $(G, E, I_{GF} \cap (G \times E))$ and $(E, M, I_{FM} \cap (E \times M))$ are also conceptual factorisation contexts.*

Proof. Let

$$(G, M, I_E) := (G, E, I_{GF} \cap (G \times E)) \circ (E, M, I_{FM} \cap (E \times M)).$$

Each (A_e, B_e), $e \in E$, is a formal concept of (G, M, I) and, since $I_E \subseteq I$, also of (G, M, I_E). □

Definition 2.11. If (G, F, I_{GF}) and (F, M, I_{FM}) are conceptual factorisation contexts of (G, M, I) and $E \subseteq F$, then $(G, E, I_{GF} \cap (G \times E))$ is called a **(many-valued) factor** of (G, M, I).

Many-valued factors are closely related to the **scale measures** described in [GW96]:

Definition 2.12. Let $\mathbb{K} := (G, M, I)$ and $\mathbb{S} := (G_{\mathbb{S}}, M_{\mathbb{S}}, I_{\mathbb{S}})$ be formal contexts. An $\mathbb{S}$**-measure** is a map

$$\sigma : G \to G_{\mathbb{S}}$$

with the property that the preimage $\sigma^{-1}(E)$ of every extent E of $\mathbb{S}$ is an extent of $\mathbb{K}$. An $\mathbb{S}$**-measure** is called **full**, if every extent of (G, M, I) is the preimage of some $\mathbb{S}$-extent.

Proposition 2.13 ([GG12]). $\mathbb{S} := (G, F, I_{GF})$ *is a factor of* (G, M, I) *if and only if the identity map is an* $\mathbb{S}$*-measure.*

Proof. Clearly $\mathbb{S} := (G, F, I_{GF})$ is a factor of (G, M, I) if and only if each attribute extent $f^{I_{GF}}$ is an extent of (G, M, I). □

Definition 2.14. A factor (G, F, I_{GF}) of (G, M, I) is called an $\mathbb{S}$**-factor** if it has a surjective full $\mathbb{S}$-measure. If $\mathbb{S}$ is an elementary ordinal, or nominal, etc., scale, we speak of an **ordinal** or **nominal factor**, etc. Moreover, we say that (G, M, I) has an **ordinal** (**nominal**, etc.) **factorisation** if and only if it has a first factorising context that can be written as an apposition of ordinal (nominal, etc.) factors.

In other words: The first factorisation context of an ordinal factorisation must be a derived context of a many-valued context with respect to some ordinal scaling.

Example 2.15. Consider once again our running example from Figure 2.1. It can be ordinally factored, using eight Boolean factors, as shown in Figure 2.4. Obviously, the

	f_1^1	f_2^1	f_3^1	f_4^1	f_1^2	f_2^2	f_3^2	f_4^2
g_1	×	×			×	×		
g_2	×	×	×		×			
g_3	×				×	×	×	×
g_4					×	×	×	×
g_5	×				×	×	×	
g_6	×	×	×	×				

$\circ$

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}
f_1^1					×		×							×
f_2^1			×		×		×							×
f_3^1			×	×	×		×							×
f_4^1			×	×	×		×				×	×		×
f_1^2	×				×				×				×	×
f_2^2	×	×			×	×			×	×			×	×
f_3^2	×	×		×	×	×			×	×			×	×
f_4^2	×	×		×	×	×		×	×	×		×	×	×

Figure 2.4.: An ordinal factorisation of the formal context from Figure 2.1

first factorisation context is a derived context of an ordinally scaled many-valued context with two many-valued attributes, and the second factorisation context is the dual of such a derived context, but with reverse scaling. Such many-valued contexts are given in Figure 2.5. The conceptual scales for the first and the second factorisation context are shown in Figure 2.6. For Figure 2.5 we used the following symbolic notation of the ordinal factorisation defined in Figure 2.4:

$$g_i\, I_{GF}\, f_j^k \iff f^k(g_i) \geq j \quad \text{and} \quad f_j^k\, I_{FM}\, m_i \iff f^k(m_i) < j.$$

Thus, we have

$$g\, I\, m \iff f^k(g) > f^k(m) \text{ for some } k.$$

Expressed differently, it holds that

$$g\, \not{I}\, m \iff f^k(g) \leq f^k(m) \text{ for all } k.$$

$\mathbb{K} =$

	f^1	f^2
g_1	2	2
g_2	3	1
g_3	1	4
g_4	0	4
g_5	1	3
g_6	4	0

$\circ$

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}
f^1	4	4	1	2	0	4	0	4	4	4	3	3	4	0
f^2	0	1	4	2	0	1	4	3	0	1	4	3	0	0

Figure 2.5.: Many-valued factorisation contexts for the ordinal factorisation from Figure 2.4

	≥ 1	≥ 2	≥ 3	≥ 4
0				
1	×			
2	×	×		
3	×	×	×	
4	×	×	×	×

	< 1	< 2	< 3	< 4
0	×	×	×	×
1		×	×	×
2			×	×
3				×
4				

Figure 2.6.: Conceptual scales for the two many-valued factorisation contexts in Figure 2.5.

It is somewhat tempting, but highly experimental, to plot the two factors as it is usual in (numerical) Factor Analysis. Such a diagram is shown in Figure 2.7. Note that we did not include the attributes “hearty” and “ambitious” as they apply to each family member. A representation like this may however be misleading, since it displays purely ordinal data in a metric fashion. An additional source of misinterpretation is that the two “dimensions” represent ordinal, not interordinal (“bipolar”) data. However, the diagram indicates that ordinal factor analysis, when interpreted correctly, has some expressiveness similar to Factor Analysis based on metric data.

Analysing the many-valued factorisation contexts or the “biplot” we observe that one factor contains the attributes with positive connotations, whereas the other one contains the attributes with negative connotations. Further, we can see that, as it is the case in ordinary Factor Analysis, the objects “load high” on only one factor, i.e., the objects have many attributes of only one of the factors. The sole exception is “myself” that is located in the middle of both factors.

Obviously, Proposition 2.6 and Proposition 2.7 also hold in this setting, i.e., the ordinal factorisation is “invariant” under the dualisation and the clarification of formal contexts. The first however does not hold in case we transform the factorisation contexts into many-valued ones.

The following proposition is evident:

Proposition 2.16 ([GG12]). *A formal context is an ordinal factor of (G, M, I) if and only if its attribute extents are a linearly ordered family of concept extents of (G, M, I).*

For an ordinal factorisation there must be a partition $\{F_d \mid d \in D\}$ of the set F of

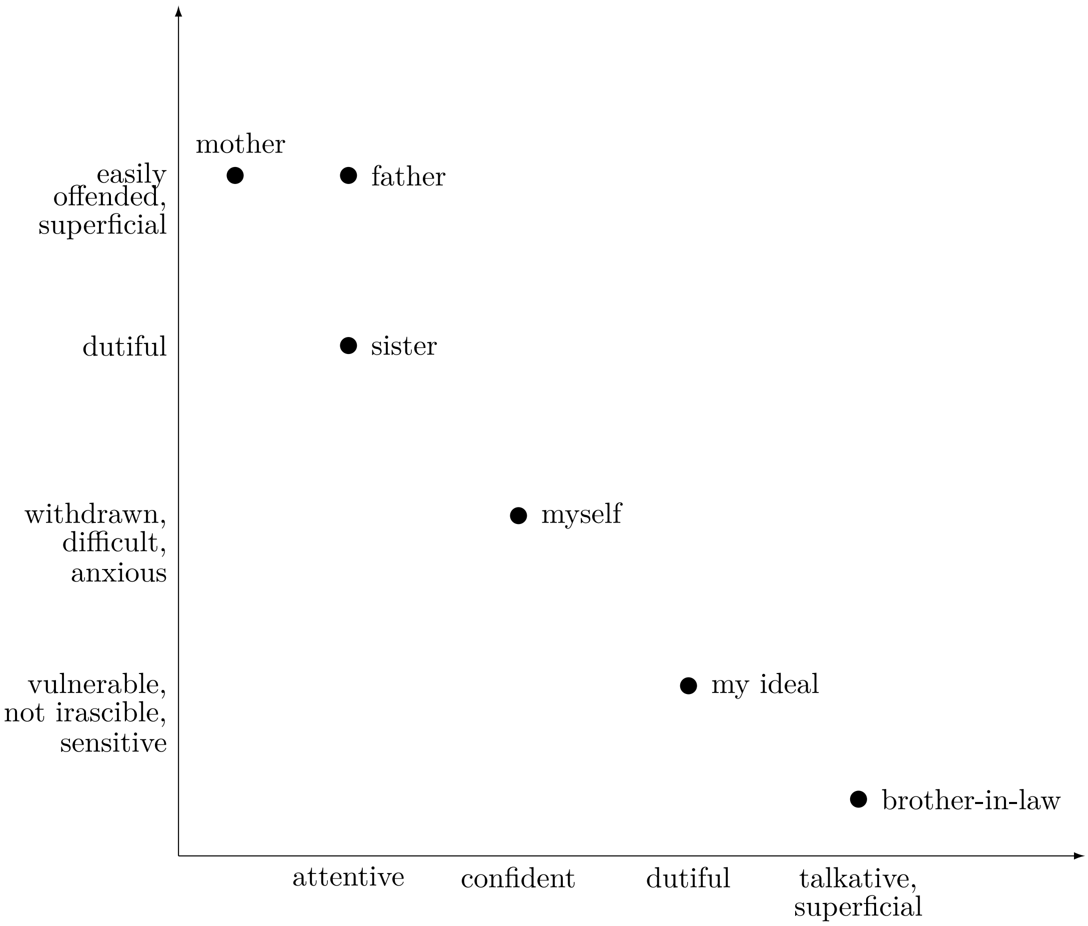

Figure 2.7.: A “biplot” of the data in Figure 2.1, based on the ordinal factorisation in Figure 2.4. Note that no metric information is encoded here. The diagram is based on ordinal data only.

factors such that within each class the attribute order of (G, F, I_{GF}) is linear. According to Proposition 2.4 the attribute order is dual to that of $(M, F, {I_{FM}}^d)$. This gives the following proposition:

Proposition 2.17 ([GG12])**.** *For any ordinal factorisation the dual of the second factorisation context is also a derived context of the same many-valued context, but with reversely ordered ordinal scales.*

Definition 2.18. A relation $R \subseteq G \times M$ is called a **Ferrers relation** if and only if there are subsets $A_1 \subset A_2 \subset A_3 \ldots \subseteq G$ and $M \supseteq B_1 \supset B_2 \supset B_3 \supset \ldots$ such that $R = \bigcup_i A_i \times B_i$. Further, R is called a Ferrers relation **of concepts** of (G, M, I) if and only if there are formal concepts $(A_1, B_1) \leq (A_2, B_2) \leq (A_3, B_3) \leq \ldots$ such that $R = \bigcup_i A_i \times B_i$.

It is well known and easy to see that a relation $R \subseteq G \times M$ is a Ferrers relation if and only if the concept lattice $\underline{\mathfrak{B}}(G, M, R)$ is a chain.

Proposition 2.19 ([GG12])**.** *Any Ferrers relation $R \subseteq I$ is contained in a Ferrers relation of concepts of (G, M, I).*

Proof. If $A_i \times B_i \subseteq I$, then $A_i \times B_i \subseteq A_i'' \times A_i'$ holds. Thus, if $R = \bigcup_i A_i \times B_i \subseteq I$, then $R \subseteq \overline{R} := \bigcup_i A_i'' \times A_i' \subseteq I$, and $\overline{R}$ is a Ferrers relation of concepts. ☐

Definition 2.20. The **width** of a factorising family $\mathcal{F}$ of concepts is the largest number of pairwise incomparable elements of $\mathcal{F}$. The **ordinal factorisation width** of (G, M, I) is the smallest width of a factorising family of concepts.

Theorem 2.21 ([GG12])**.** *The following are equivalent:*

1. *(G, M, I) has ordinal factorisation width $\leq n$.*
2. *(G, M, I) has an ordinal factorisation with $\leq n$ ordinal factors.*
3. *$\underline{\mathfrak{B}}(G, M, G \times M \setminus I)$ has order dimension $\leq n$.*
4. *I can be written as a union of $\leq n$ Ferrers relations.*

Proof. (1) $\Rightarrow$ (2): (G, M, I) has ordinal factorisation width $\leq n$ if and only if there is a factorising family $\mathcal{F}$ of concepts, which as an ordered subset of the concept lattice has width $\leq n$. By a classical theorem of Dilworth this implies that $\mathcal{F}$ can be covered by $\leq n$ chains, i.e., linear ordered families of concepts, each of which induces an ordinal factor. This proves (2).

(2) $\Rightarrow$ (4): The factorising family of an ordinal factor is a chain of concepts, and the incidences occurring in such a chain form a Ferrers relation.

(3) $\Leftrightarrow$ (4) is well known, see for instance [GW96].

(4) $\Rightarrow$ (1): If I can be written as a union of $\leq n$ Ferrers relations, then it can, according to Proposition 2.19, also be written as a union of $\leq n$ Ferrers relations of concepts. These concepts form a factorising family of width $\leq n$. ☐

Example 2.22. Consider once more the formal context from Figure 2.1 (page 40). Its

	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}
g_1	2	2	1		1	2	1		2	2			2	1
g_2	2		1	1	1		1		2				2	1
g_3	2	2		2	1	2	1	2	2	2		2	2	1
g_4	2	2		2	2	2		2	2	2		2	2	2
g_5	2	2		2	1	2	1		2	2			2	1
g_6			1	1	1		1				1	1		1

Figure 2.8.: Two Ferrers relations, the union of which is the incidence relation of the formal context in Figure 2.1.

incidence relation I can indeed be covered by two Ferrers relations, as can be seen from Figure 2.8, where 1 and 2 symbolise the two different relations. Note that the two relations are not disjoint. Thus, to improve readability, we entered a "2" in a context cell, if the corresponding cross is contained in the second Ferrers relation and not in the first.

So the ordinal width of the formal context in Figure 2.1 equals two (a smaller value is obviously impossible). This was to be expected, since an ordinal factorisation with two ordinal factors was given in Figure 2.5. Moreover the order dimension of the lattice in Figure 2.3 (page 41) is apparently equal to two.

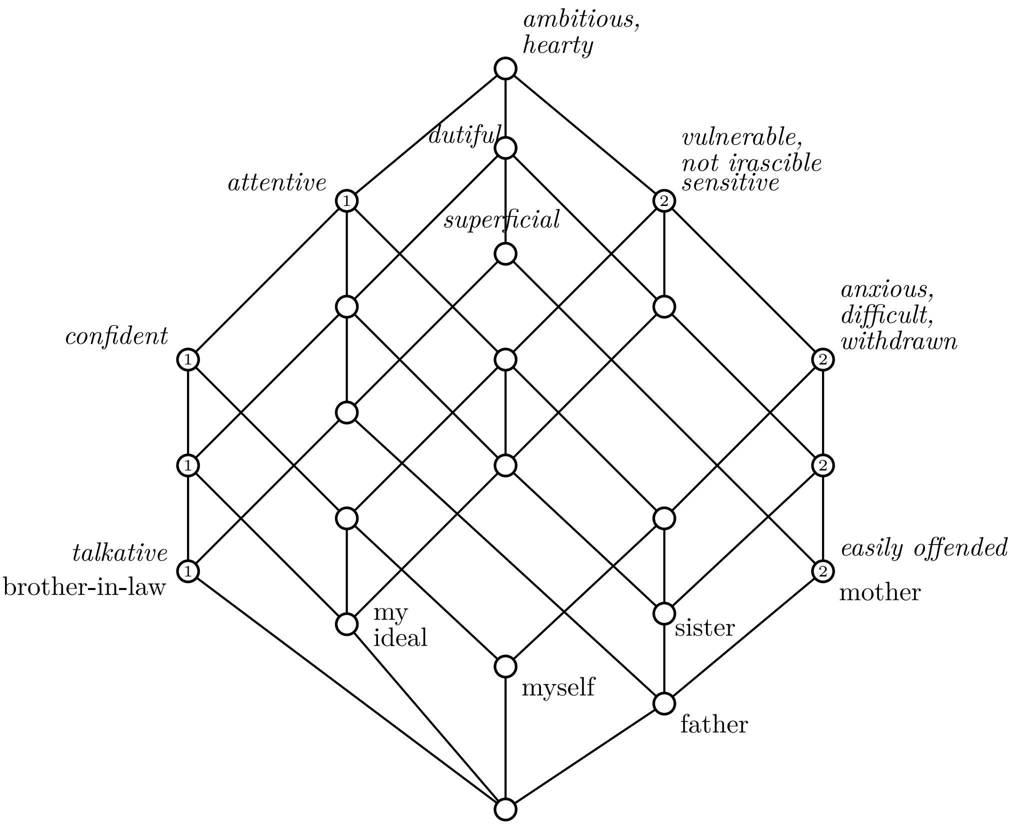

Figure 2.9.: Concept lattice of the anorexia nervosa context. The encircled numbers mark a factorising family of concepts of width two.

The concept lattice of the formal context in Figure 2.1 is shown in Figure 2.9. Two chains are marked in the diagram. These cover all (tight) incidences, i.e., whenever $(g, m) \in I$, then the interval $[\gamma g, \mu m]$ contains some concept from one of these chains. Therefore, these concepts form a factorising family of width two.

An immediate consequence of Theorem 2.21 is that for any $k \geq 3$ the decision problem whether a formal context has factorisation width $\leq k$ is NP-complete. This follows from Yannakakis' result ([Yan82]) that "order dimension $\leq k$" is hard to decide for $k \geq 3$. Another consequence is that one can easily determine the factorisation width of some elementary scales:

Corollary 2.23 ([GG12]). *1. (G, M, I) has ordinal factorisation width 1 if and only if I is Ferrers.*

2. The (one-dimensional) contraordinal scale has ordinal factorisation width 2, independent of its size (> 1).

3. The interordinal scale has ordinal factorisation width 2, independent of its size (> 1).

The corollary gives first clues of how algorithmically difficult interordinal and contraordinal factorisation (yet to be developed) will be. The *nominal* scale with n scale values obviously has ordinal factorisation width n.

The underlying ideas of *Nonmetric Factor Analysis*, developed by Coombs and Kao ([Coo64, CK55]), are similar to those of the ordinal factorisations. As the title suggests, the authors distance themselves as well from a metric handling of data. The theory was further expanded by Doignon, Ducamp and Falmagne in [DDF84].

2.3. Applications

In this section we put the ordinal factorisation to work and test it on three real-world data sets. These are chosen in a way such that they cover different areas and data collecting methods. Our first case study is done on a repertory grid data from [CH08]. It was obtained by asking a man with numerous tattoos to make character judgements about people based on their looks. In the second study we analyse the data set from [VD89] concerning psychiatric symptoms. The data was analysed in [BVG01] with a Latent Class approach, and we compare the results from there with the ones obtained by the ordinal factorisation. In the last case study we use a data set collected for the purpose of performing ordinal factorisations. The data was registered in June 2012 in the otolaryngology clinic at the university hospital "Titu Maiorescu" from Bucharest, Romania by an otolaryngologist.

Usually in Factor Analysis one is interested in an approximate factorisation rather than in an exact one. Moreover, one wishes for few factors that explain a large part of the data and 2-3 factors are preferable.

As we have already discussed in Section 2.1 the approximate factorisations are of two kinds, with positive and with negative discrepancies. The latter are easy to obtain. One does not cover the entire incidence relation of the context. We will use this kind of approximate factorisation. Our algorithm works as follows: First it selects a maximal Ferrers relation that covers most of the incidence relation. Afterwards maximal Ferrers relations are selected that cover most of the yet uncovered incidence relation. The halting condition of the algorithm is the number of factors to be extracted. Note that the so-obtained factors are not necessarily present in an optimal or/and an exact ordinal factorisation.

One may also attempt to perform an approximate factorisation with positive discrepancies. This is however not only trickier from the algorithmic point of view but also with regard to the interpretation. In the previous case we knew that not the entire information of the data is contained in the factorisation, whereas in this case we are altering the data by adding new object-attribute tuples to the incidence relation.

If we are looking for a factorisation with two ordinal factors we may apply an additional trick. Before presenting it, we have to introduce *incomparability graphs*. The **incomparability graph** (V, E) of a formal context (G, M, I) is given by

$$V := (G \times M) \setminus I,$$
$$\{(g,m),(h,n)\} \in E :\Longleftrightarrow (g,m) \notin I, (h,n) \notin I, (g,n) \in I, (h,m) \in I.$$

It was shown in [DDF84] that the *Ferrers-dimension*, the number of optimal ordinal factors, is two if and only if the incomparability graph is bipartite. Thus, one could try to alter the context such that its incomparability graph is bipartite.

Tattoo Data

The first to propose the application of Formal Concept Analysis on *repertory grid* data were Wolff and Spangenberg in [SW91]. We have already seen in the previous section the

	Self	Ideal self	Future self	Self as I'd like others to see me	Self as seen by others	Self before I had a tattoo	Self now no tattoo	Someone without a tattoo	Someone with a tattoo	Someone who conforms	Someone who pushes boundaries	Someone I don't like the look of	
happy	2	2	2	2	2	1	3	2	1	5	1	5	miserable
patient	2	1	2	2	2	3	2	3	4	5	2	5	noxious
caring	5	2	3	1	2	5	1	1	2	5	1	4	neutral
kind	2	2	2	2	1	2	2	1	2	5	2	3	callous
sense of humour	1	1	1	1	1	3	2	2	2	5	2	5	dour
sociable	5	3	5	5	5	4	5	1	1	2	1	4	hermit-like
friendly	2	2	2	2	5	2	2	2	1	4	2	4	aloof
active	1	1	2	1	1	1	1	1	2	1	1	4	dead beats
confident	1	1	1	1	1	4	2	2	2	3	1	5	weak
organised	1	1	1	2	3	3	2	5	4	2	1	4	shambles
hard working	2	2	3	1	2	2	2	1	2	1	1	4	lazy
honest	1	1	1	2	1	1	1	1	1	2	2	4	slimy

Figure 2.10.: Repertory grid taken from [CH08]

results of an ordinal factorisation on such data, i.e., the grid from the anorexia nervosa therapy. In [CH08] the authors have applied Formal Concept Analysis on a different real-world data set. The latter was obtained from a 57 year old man with approximately ten tattoos. He made his first tattoo with the age of 19 and his last one was ready a few weeks before the grid was filled out. Here we reproduce the initial data from [CH08] in Figure 2.10.

Let us first explain briefly how such grids are designed and how they are transformed into formal contexts. First the interviewee has to choose persons from their close environment. Afterwards, he/she is required to compare and contrast successive sets of three persons, *triads*, from those already named. In this process the interviewee has to state some way in which two persons from the triad are alike and some way the third person is different from the other two. These pairs of attributes/characteristics are called *constructs*, whereas the first attribute is called the *left pole* and the second one the *right pole*. Thereafter, the interviewee has to assign to each person exactly one value from a *scale* for each construct. These values yield a bipolar ordering, indicating which pole of the construct characterises the person. Usually the scale values range from 1 to 5 or from 1 to 7. In our data set the values range from 1 to 5. The values 1 and 2 are associated to the left pole, where 1 means that the left pole characterises the person, and 2 is a gradation of 1. The right pole is associated with the values 4 and 5, where 5 denotes the decision that the person was assigned to the right pole, and 4 is a gradation of 5. The value 3 indicates that a person can neither be characterised by the left pole nor by the right pole, or that both attributes apply in the same extent. Let us take a look at the second row of the grid from Figure 2.10. "Ideal self" is characterised as patient, the persons that have a 2 are considered fairly patient, whereas "Someone who conforms" is seen as noxious and "Someone with a tattoo" is regarded as fairly noxious. The persons that have a 3 are neither seen as patient nor as noxious. The so-obtained data can be easily represented by a formal context. The persons build the object set. Both poles of each construct become an attribute. The incidence relation is determined by the values, i.e., 1 and 2 are assigned to the left pole, the values 4 and 5 to the right pole and a score of 3 is indicated by a blank.

The so-obtained formal context can be factorised with four ordinal factors. However, by choosing two ordinal factors we can explain $77,44\%$ of the incidence relation. These two factors were obtained by the approximate algorithm described above. The factors are "plotted" in Figure 2.11. Once again this diagram does not contain metric information, it is based only on ordinal data. One can easily see that the first factor contains the judgements with positive connotations, whereas the second factor contains those with negative ones. However, if we perform an exact ordinal factorisation, then the description of two factors roughly corresponds to the approximate factors. The other two factors represent a mixture between positive and negative judgements. In ordinary Factor Analysis it is desirable that the objects load high on only one factor, i.e., have many attributes of one factor. We can see from the "biplot" that this is also the case for our ordinal factorisation.

The outcome of the approximate factorisation is quite satisfying. On the one hand, it covers a large proportion of the incidence data. On the other hand, the description yields information about the positive and negative judgements of the interviewee.

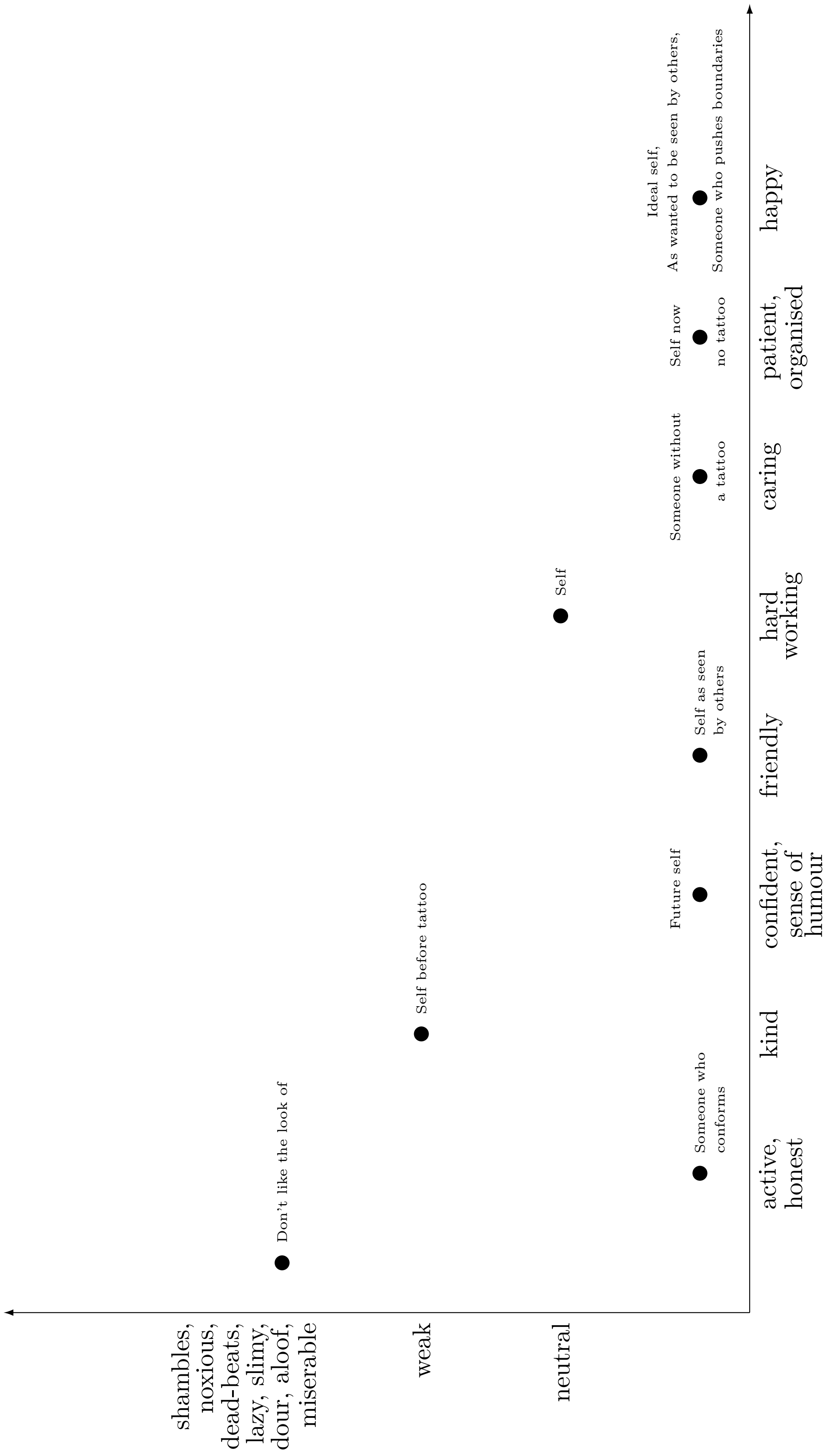

Figure 2.11.: The “biplot” of two ordinal factors of the scaled context to the repertory grid from Figure 2.10

Psychiatric Symptoms Data

Our second case study is on the data set collected by Van Mechelen and De Boeck ([VD89]). It consist of yes-no judgements made by a psychiatrist about the presence of 23 psychiatric symptoms on 30 patients. The data is displayed in Figure 2.12. It was re-analysed in [BVG01] with a variant of *Latent Class Analysis*. In such a data analysis technique one attempts to detect the *latent classes*, in this example the diseases, and associate to them the measured variables, in our case the symptoms. The authors from [BVG01] have settled for a solution with three latent classes for the data in Figure 2.12. These were described as follows: Class 1 is associated with high probabilities on the symptoms agitation, ideas of persecution and hallucinations. These symptoms indicate a psychosis syndrome. Class 2 is associated with depression, anxiety and suicide, and can be interpreted as an affective syndrome. Class 3 is associated primarily with alcohol abuse. Moreover, the authors show for each symptom the probabilities to which it belongs to each of the three classes. From there we find out that there are also other symptoms, besides the ones listed before, that belong to the classes with a high probability. Here we show just those attributes that have at least 20%.

Class 1	Class 2	Class 3
ideas of persecution ≈ 90%	depression ≈ 95%	leisure time impairment ≈ 95%
hallucinations ≈ 90%	leisure time impairment ≈ 85%	daily routine impairment ≈ 90%
inappropriate affect ≈ 80%	daily routine impairment ≈ 85%	alcohol abuse ≈ 70%
agitation ≈ 70%	social isolation ≈ 70%	inappropriate affect ≈ 55%
leisure time impairment ≈ 70%	anxiety ≈ 70%	memory impairment ≈ 40%
daily routine impairment ≈ 70%	suicide ≈ 60%	antisocial acts ≈ 40%
social isolation ≈ 45%	inappropriate affect ≈ 50%	retardation ≈ 40%
anxiety ≈ 40%	somatic concerns ≈ 20%	social isolation ≈ 30%
depression ≈ 35%		negativism ≈ 30%
suicide ≈ 20%		
somatic concerns ≈ 20%		

Of course it is neither adequate nor fair to compare the outcome of the Latent Class analysis with the outcome of the ordinal factorisation as they are based on different philosophies and techniques. However, the results of the ordinal factorisation are substantiated by the results of the Latent Class technique. For the conceptual factorisation we used just a part of the symptoms and left seven attributes out, namely disorientation, obsession, lack of emotion, speech disorganisation, overt anger, grandiosity and drug abuse. This choice is justified by the fact that these attributes have probabilities between little above zero and under 20% of appearing in any of the three classes. The formal context obtained after removing these attributes can be factorised with seven ordinal factors. However, we are looking for three factors and therefore apply the approximate ordinal factorisation. These factors are “tri-plotted” in Figure 2.13 and they cover $82,54\%$ of the incidence relation of the context. Even more, there is a strong correspondence between the ordinal factors and the latent classes. The most evident one is between the third latent class and the third ordinal factor. In both cases “leisure time impairment” and “daily routine impairment” are the most common symptoms and “social isolation” the least common one. Moreover, the one-valued attributes contributing to the formation of the ordinal factor are a subset of the attributes that occur in the latent class with high probability. Though the orderings are a little bit different. The first latent class may be identified with the second ordinal

	disorientation	obsession	memory impairment	lack of emotion	antisocial acts	speech disorganisation	overt anger	grandiosity	drug abuse	alcohol abuse	retardation	negativism	somatic concerns	ideas of persecution	hallucinations	agitation	suicide	anxiety	social isolation	inappropriate affect	depression	leisure time impairment	daily routine impairment
g_1									×										×		×		
g_2																	×	×			×		
g_3													×					×			×	×	×
g_4		×																×		×		×	×
g_5					×				×											×		×	×
g_6							×			×						×						×	×
g_7																	×	×		×	×	×	×
g_8																	×	×	×		×	×	×
g_9																	×	×	×		×	×	×
g_{10}													×					×		×	×	×	×
g_{11}													×					×	×		×	×	×
g_{12}								×									×	×	×		×	×	×
g_{13}												×				×	×		×		×	×	×
g_{14}				×										×	×				×		×	×	×
g_{15}								×						×	×	×				×		×	×
g_{16}							×		×					×	×	×		×		×			
g_{17}													×				×		×	×	×	×	×
g_{18}																	×	×	×	×	×	×	×
g_{19}			×		×					×	×	×										×	×
g_{20}					×					×		×							×	×	×	×	×
g_{21}								×				×						×	×	×	×	×	×
g_{22}											×						×	×	×	×	×	×	×
g_{23}											×						×	×	×	×	×	×	×
g_{24}						×		×						×	×	×				×		×	×
g_{25}														×	×	×		×	×	×		×	×
g_{26}							×			×						×		×		×	×	×	×
g_{27}									×			×	×					×	×	×	×	×	×
g_{28}	×		×			×				×	×								×	×		×	×
g_{29}													×	×	×	×	×		×	×	×		×
g_{30}				×		×					×		×	×	×		×	×	×	×	×	×	×

Figure 2.12.: Psychiatric symptoms data from [VD89]

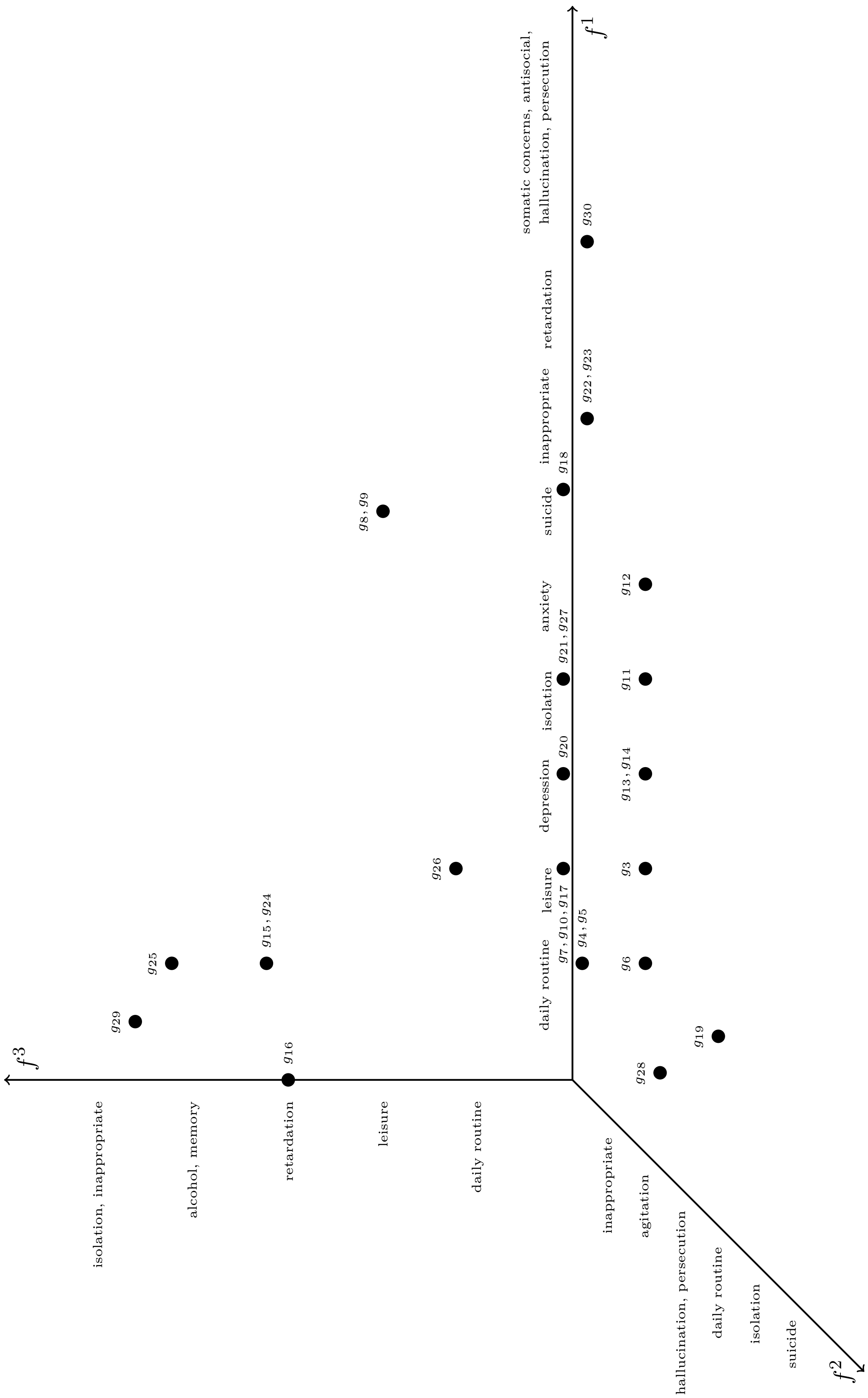

Figure 2.13.: A three-dimensional “biplot” of an ordinal factorisation of the context from Figure 2.12 using three approximate factors

factor. For both “inappropriate affect” and “agitation” are common and “suicide” is the most specific one, i.e., it appears seldom. The second latent class corresponds to the first ordinal factor. In both cases the first seven attributes are the same, the only difference lies in the position of the attribute “depression”. In contrast to the latent class, the ordinal factors contains more “most specific” attributes. The attributes of the latent class are a subset of the ordinal factors’ attributes.

In the analysis of this real-world data set we have seen that there is quite a tight relation between the outcome of the ordinal factorisation and the Latent Class analysis. Thus, the ordinal factors obtained by the conceptual factorisation are verified by the latent classes.

Otolaryngology Data

The data to be analysed in the following was collected by an otolaryngologist in June 2012 at the otolaryngology clinic of the university hospital “Titu Maiorescu” from Bucharest, Romania for the purpose of ordinal factorisation.

The column “Principal symptoms” refers to the most pronounced symptoms of the patients and “Secondary symptoms” contains those symptoms that were more or less unincisive. In “Principal diagnosis” the principal diagnoses of the otolaryngologist are noted, whereas in “Secondary diagnosis” alternative diagnoses are contained. Let us first explain some notions from the data set:

- dysphagia = swallow difficulties;
- odynophagia = pain while swallowing;
- nasal obstruction = blockage of the nasal passages;
- lump in throat = the feeling of having a foreign body in the throat;
- rhinorrhea = “runny nose”;
- epistaxis = nosebleed;
- otalgia = pain in the ear;
- otorrhea = drains to the outside of the ear;
- tinnitus = perception of sound in the absence of it;
- vertigo = dizziness;
- autophony = unusually loud hearing of a person’s own voice.

We have scaled the many-valued context as follows: For every attribute value we have introduced a one-valued attribute. We have made a cross in the corresponding line of the patient and the corresponding column of the attribute if the patient suffered from that symptom, indifferent from being principally or secondary, or if he was diagnosed principally or secondary with that disease. Further, purulent, mucous and serous rhinorrhea imply the attribute ”rhinorrhea”. Moreover, any acute or chronic form of a disease implies the disease. We have ignored the many-valued attributes “age” and “sex”.

	Sex	Age	Principal symptoms	Secondary symptoms	Principal diagnosis	Secondary diagnosis
1	M	75	Dysphagia	Lump in Throat, Cough	Chronic Pharyngitis	Deviated Septum, Chronic Sinusitis
2	M	72	Cervical Lymphadenopathy		Renal Carcinoma	Metastatic Cervical Lymphadenopathy
3	M	25	Purulent Rhinorrhea	Headache, Hyposmia, Chronic Nasal Obstruction	Acute Sinusitis	Deviated Septum
4	M	85	Epistaxis		Epistaxis	Blood Hypertension
5	F	36	Rhinorrhea, Headache, Chronic Nasal Obstruction	Dysphagia, Otalgia	Deviated Septum	Turbinate Hypertrophy, Chronic Sinusitis, Chronic Pharyngitis
6	F	63	Mucous Rhinorrhea	Chronic Nasal Obstruction	Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
7	F	54	Lump in Throat	Dysphagia	Tonsillar Tumour	Chronic Laryngitis, Deviated Septum
8	F	36	Dysphagia, Odynophagia	Headache	Acute Tonsillitis	Acute Pharyngitis, Deviated Septum, Chronic Sinusitis, Chronic Laryngitis
9	F	34	Tinnitus Auris	Hearing Loss, Otalgia, Chronic Nasal Obstruction	Deviated Septum	Chronic Otitis Media, Chronic Sinusitis, Chronic Pharyngitis
10	F	20	Fever, Headache, Odynophagia	Muscle and Joint Pain	Acute Tonsillitis	Acute Pharyngitis, Deviated Septum
11	M	67	Lump in Throat	Dysphagia	Uvula Hypertrophy	Chronic Pharyngitis
12	M	47	Purulent Rhinorrhea	Headache	Acute Sinusitis	Deviated Septum, Chronic Pharyngitis
13	M	32	Headache	Chronic Nasal Obstruction	Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
14	M	31	Hemoptysis	Lump in Throat		Deviated Septum, Chronic Pharyngitis
15	F	56	Headache	Otalgia, Tinnitus Auris	Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
16	M	63	Fever	Myalgia, Headache	Acute Pharyngitis	Pneumopathy, Deviated Septum
17	M	25	Chronic Nasal Obstruction		Deviated Septum	Turbinate Hypertrophy
18	M	65	Dysphagia	Odynophagia, Otalgia	Tonsillar Tumour	
19	M	38	Dysphonia	Odynophagia	Acute Laryngitis	Chronic Pharyngitis
20	F	31	Fever	Trembling Headache, Mucous Rhinorrhea	Chronic Sinusitis	Pneumopathy, Deviated Septum
21	F	54	Lump in Throat	Dysphagia	Chronic Sinusitis	Chronic Pharyngitis
22	F	44	Hearing Loss	Ear Fullness	Turbinate Hypertrophy	Chronic Otitis Media
23	M	7	Chronic Nasal Obstruction	Mucous Rhinorrhea	Adenoidal Hypertrophy	
24	M	74	Purulent Rhinorrhea	Headache, Hyposmia, Chronic Nasal Obstruction	Chronic Sinusitis	Chronic Sinusitis, Chronic Pharyngitis
25	F	58	Hearing Loss	Otorrhea	Chronic Otitis Media	Deviated Septum
26	F	32	Epistaxis		Epistaxis	Deviated Septum, Turbinate Hypertrophy
27	M	23	Purulent Rhinorrhea	Odynophagia, Chronic Nasal Obstruction	Acute Sinusitis	Acute Pharyngitis
28	M	26	Hearing Loss	Fever	Chronic Otitis Media	Deviated Septum, Chronic Pharyngitis
29	M	72	Chronic Nasal Obstruction	Serous Rhinorrhea, Headache	Polypoid Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
30	F	36	Vertigo	Headache,	Vestibular Syndrome	Acute Sinusitis, Deviated Septum

				Mucous Rhinorrhea		
31	F	55	Tinnitus Auris	Vertigo, Headache, Hearing Loss	Tinnitus Auris	Sensorineural Hearing Loss, Vestibular Syndrome, Chronic Pharyngitis
32	F	29	Hearing Loss	Autophony	Chronic Otitis Media	Septum Deviation
33	M	21	Dysphagia	Fever, Trembling, Odynophagia	Acute Tonsillitis	Deviated Septum
34	F	9	Chronic Nasal Obstruction	Hearing Loss	Adenoidal Hypertrophy	Chronic Otitis Media
35	F	25	Nasal Obstruction	Serous Rhinorrhea, Headache	Chronic Sinusitis	Deviated Septum, Turbinate Hypertrophy
36	M	30	Nasal Obstruction	Serous Rhinorrhea, Headache	Chronic Sinusitis	Deviated Septum, Turbinate Hypertrophy
37	M	59	Tinnitus Auris	Vertigo	Tinnitus Auris	Vestibular Syndrome, Deviated Septum, Chronic Sinusitis, Chronic Pharyngitis
38	M	50	Odynophagia	Lump in Throat	Chronic Pharyngitis	Chronic Tonsillitis, Deviated Septum, Chronic Pharyngitis
39	F	40	Ear Fullness	Autophony, Headache, Hearing Loss	Otosclerosis	Chronic Otitis Media, Deviated Septum, Chronic Sinusitis, Chronic Pharyngitis
40	F	75	Lump in Throat	Serous Rhinorrhea, Headache	Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
41	F	12	Odynophagia	Serous Rhinorrhea, Dysphagia	Chronic Pharyngitis	Chronic Tonsillitis
42	F	56	Hearing Loss	Vertigo	Chronic Mastoiditis	Deviated Septum, Chronic Pharyngitis
43	M	75	Odynophagia	Lump in Throat, Chronic Nasal Obstruction, Hipoacusis	Chronic Pharyngitis	Deviated Septum, Presbyacusis
44	M	36	Cervical Lymphadenopathy		Rhinopharynx Carcinoma	
45	M	34	Chronic Nasal Obstruction	Mucous Rhinorrhea	Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
46	F	24	Headache	Fever, Trembling Rhinorrhea	Acute Pharyngitis, Trigeminal Neuralgia	Chronic Sinusitis,
47	F	20	Chronic Nasal Obstruction	Mucous Rhinorrhea, Headache	Chronic Sinusitis	Deviated Septum, Chronic Pharyngitis
48	F	33	Ear Fullness	Hearing Loss, Chronic Nasal Obstruction	Chronic Otitis Media	Deviated Septum, Chronic Pharyngitis
49	F	35	Chronic Nasal Obstruction	Sleep Apnea, Headache	Turbinate Hypertrophy	Chronic Tonsillitis
50	F	43	Headache	Purulent Rhinorrhea	Chronic Sinusitis	Deviated Septum
51	F	75	Tinnitus Auris	Vertigo, Hearing Loss	Tinnitus Auris	Presbyacusis
52	M	45	Otalgia	Hearing Loss	External Otitis	Deviated Septum, Chronic Pharyngitis

With this data set we will perform two case studies. First, we apply the ordinal factorisation on a subcontext of the scaled context containing almost all the symptoms. After finding the ordinal factors, we associate to each of them a disease based on the symptoms they contain. Thereafter we try to validate them. We can do so since we have the true diagnoses of the patients. In the second experiment we consider the symptoms which are typical for throat and nose illnesses. Once again, with the help of the initial data sets, we will try to validate the ordinal factors.

First let us take a subcontext of the scaled context containing each symptom that appeared more than twice in the original data set. The "ignored" attributes can be seen as atypical symptoms. Further, instead of taking all forms of rhinorrhea and nasal obstruction we just consider the attributes "rhinorrhea" and "nasal obstruction". The so-obtained context has 14 attributes, 57 concepts, 93 maximal Ferrers relations and its concept lattice has 18 atoms. The latter is due to the fact that almost each patient has a unique symptom and disease pattern. We find an optimal factorisation of the context with 10 ordinal factors. Although 10 factors are quite many, we are not surprised of the outcome since in the initial data set there are 35 different diagnoses 13 of which appear just once. Listed below are the diseases, written in italics, that correspond to the description of the factors and the symptoms are ordered from most common to most specific. A separation of attributes by a semicolon means that they are contained in different intents, whereas a comma symbolises that they belong to the same intent. Note that we only used the attributes of the factors in order to associate them the corresponding disease. The ordinal factors are:

- *otitis media*: hearing loss; ear fullness; headache, autophony
- *otitis media*: hearing loss; autophony; ear fullness, headache
- *acute otitis media*: hearing loss; otorrhea
- *deviated septum*: otalgia; hearing loss; nasal obstruction, tinnitus
- *tonsillitis*: headache; rhinorrhea; nasal obstruction; dysphagia, otalgia
- *tonsillitis*: odynophagia; dysphagia; rhinorrhea
- *pharyngitis*: lump in throat; dysphagia
- *sinusitis*: nasal obstruction; rhinorrhea; odynophagia
- *acute sinusitis*: fever; headache; odynophagia
- *vestibular syndrome*: vertigo; tinnitus; hearing loss; headache.

Thus, there are two manifestation types of otitis media having the same symptoms but in a different ordering. A similar remark holds for tonsillitis. Note that not each diagnosis from the initial data set is present as an ordinal factor. If one examines the patients that "load high" on the factors, i.e., those that have many of their attributes, one finds out that these patients, with a few exceptions, were indeed diagnosed to have that disease.

Further, one may be interested in 2-3 factors that explain most of the data. Finding these would give us some information about the diseases which are most common. We have

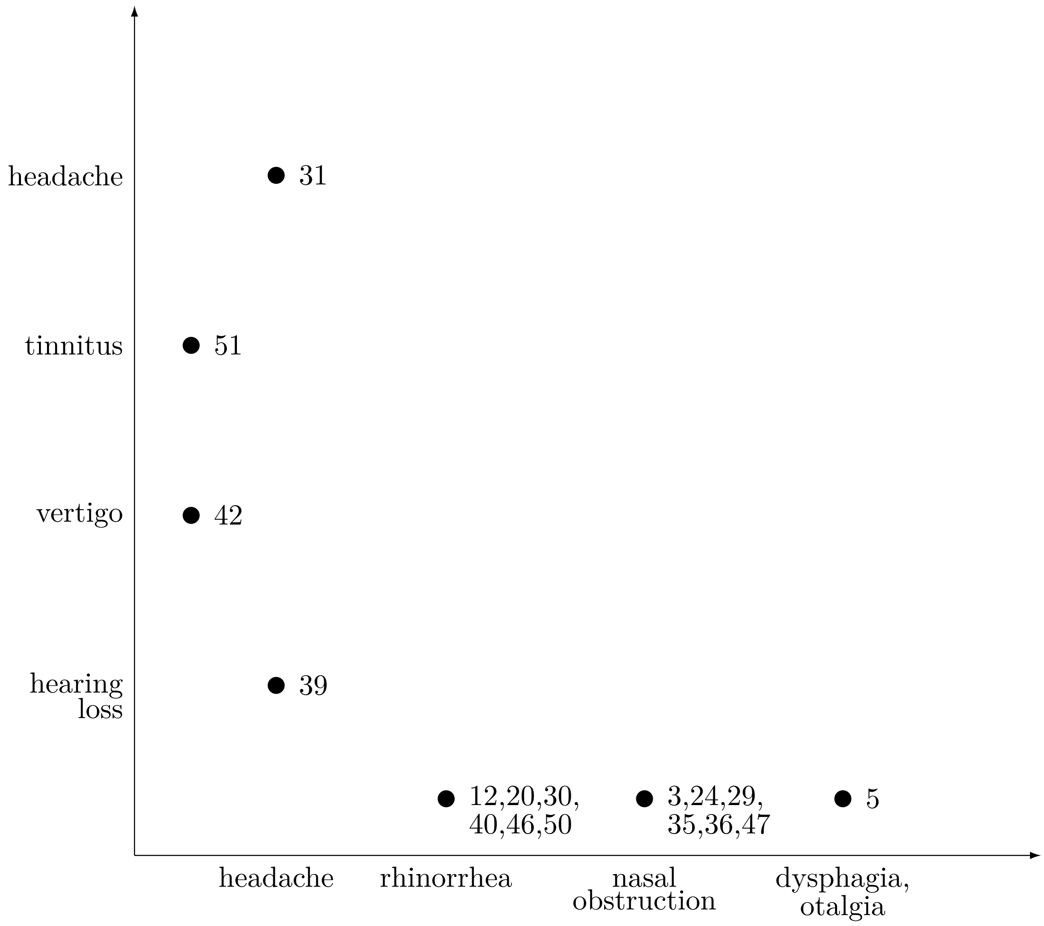

Figure 2.14.: A "biplot" of an ordinal factorisation of the otolaryngology data based on two approximate ordinal factors

chosen 2 factors which cover 50% of the incidences. They are "biplotted" in Figure 2.14. We did not include the patients which suffer only from the most common symptom of the two diseases, i.e., from either nasal obstruction for the first factor or headache for the second. The symptoms of the first ordinal factor belong to deviated septum, which was diagnosed in 35 cases, whereas the second one corresponds to vestibular syndrome. The latter was diagnosed only 3 times, however none of its symptoms, besides the last one, were covered by the first factor. Further, by comparing the "biplot" with the original data set it turns out that the patients that "load high" on the factors indeed suffer from the corresponding disease.

Let us take now a different subcontext containing the symptoms that are typical for throat and nose diseases and that were diagnosed at least twice in the original data set. This time we also include the different forms of rhinorrhea and nasal obstruction. The so-obtained context has 12 attributes, 51 concepts, 23 maximal Ferrers relations and its concept lattice has 15 atoms. The optimal factorisation consists of 7 ordinal factors. As we stated before the high number of factors is not surprising since in the initial data set there are numerous diagnoses. Further, there are distinct manifestations of the same disease, either through different symptoms or through various orderings of the attributes within the factor. The ordinal factors are listed below. The diseases, written in italics,

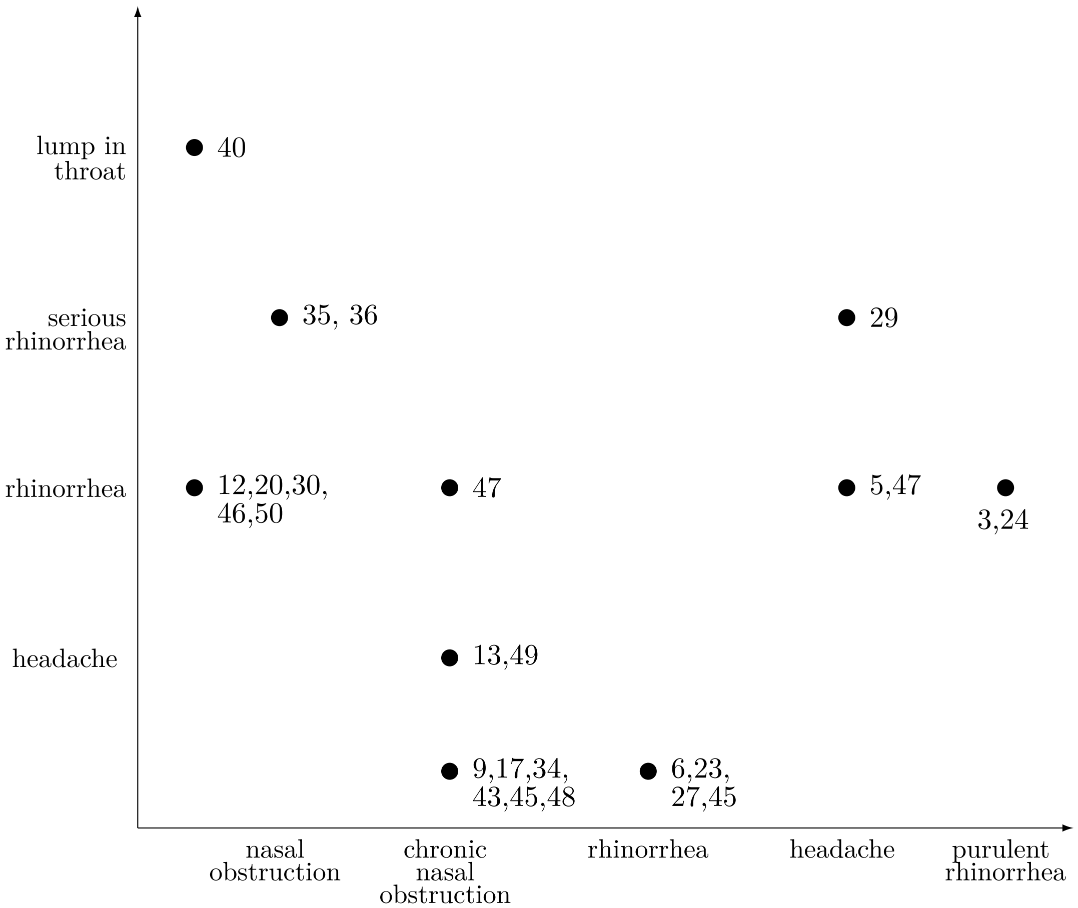

Figure 2.15.: A "biplot" of an ordinal factorisation of the otolaryngology data based on two approximate ordinal factors for symptoms that are typical for nose and throat diseases

correspond to the description of the factors and the symptoms are ordered from most common to most specific. Note once again that the semicolon separates the attributes from different intents, whereas the comma separates the attributes of the same intent. The ordinal factors are:

- *chronic sinusitis*: rhinorrhea; mucous rhinorrhea; headache; trembling headache
- *chronic sinusitis*: headache; rhinorrhea; serous rhinorrhea; nasal obstruction; chronic

nasal obstruction

- *acute sinusitis*: rhinorrhea; purulent rhinorrhea; nasal obstruction; chronic nasal obstruction; odynophagia
- *pharyngitis*: lump in throat; dysphagia
- *deviated septum*: nasal obstruction; chronic nasal obstruction; headache; rhinorrhea; dysphagia
- *deviated septum*: fever; headache; trembling headache
- *acute tonsillitis*: odynophagia; dysphagia; rhinorrhea; serous rhinorrhea.

Notice once again that we only used the attributes contained in the factors to determine their meanings. The most common disease seems to be sinusitis, which in the initial data set was diagnosed 26 times. Although deviated septum was diagnosed in 35 cases, it appears here only twice. One could conclude that the latter has less manifestation forms than sinusitis.

We are interested in two factors that explain most part of the data. These are “biplotted” in Figure 2.15 and explain $63,71\%$ of the incidences. Both factors describe sinusitis. This is not surprising in view of the fact that the optimal factorisation contains 3 factors out of 7 that describe sinusitis. Once more the factors are validated by the initial data set, as the patients that “load high” on the two factors indeed suffer from sinusitis. In Figure 2.15 we did not include the patients that have only the most common symptoms from either of the two factors.

In the previous case study the data was obtained by a checklist inspecting whether or not the patients suffer from the symptoms of the list. The data used in this study contains the records about the symptoms and the diseases of the patient without any restrictions. Our analyses have shown that even on such a freely-obtained data set the ordinal factors perform well.

2.4. Conceptual Factorisation of L-Contexts

The factorisation of **L**-contexts was introduced in [Běl08] and further studied in [BV09a]. In accordance with the rest of this thesis we deviate from the authors' notations.

Definition 2.24. A **factorisation** of an **L**-context (G, M, I) consists of two **L**-contexts (G, F, I_{GF}) and (F, M, I_{FM}) such that

$$I(g,m) = l \iff I_{GF}(g,f) \otimes I_{FM}(f,m) = l \text{ for some } f \in F.$$

The set F is called the **factor set**, its elements the **(L-)factors**, and (G, F, I_{GF}) and (F, M, I_{FM}) are said to be the **factorisation contexts**. We write

$$(G, M, I) = (G, F, I_{GF}) \circ (F, M, I_{FM})$$

to indicate a factorisation.

As in the crisp setting, we may associate to each factorisation a **factorising family** $\{(A_f, B_f) \mid f \in F\}$ given by the **L**-sets $A_f \in \mathbf{L}^G$ and $B_f \in \mathbf{L}^M$ defined as $A_f(g) := I_{GF}(g, f)$ and $B_f(m) := I_{FM}(f, m)$ for all $g \in G$ and $m \in M$. A family $\{(A_f, B_f) \mid f \in F\}$ is a factorising family of (G, M, I) if and only if

$$I = \bigcup_{f \in F} A_f \bigotimes B_f \tag{2.5}$$

where $A_f \otimes B_f$ is the **L**-set $A_f \otimes B_f : G \times M \to L$ given by

$$(A_f \bigotimes B_f)(g, m) := A_f(g) \otimes B_f(m) \tag{2.6}$$

for all $g \in G$ and $m \in M$. Expressed differently, $\{(A_f, B_f) \mid f \in F\}$ is a factorising family of (G, M, I) if and only if

$$I = \bigcup_{f \in F} f^{I_{GF}} \bigotimes f^{I_{FM}}.$$

Once more these factorising families correspond precisely to those families of **L**-preconcepts of (G, M, I) that cover the **L**-relation I. Similarly to the crisp case a tuple (A, B) is called an **L-preconcept** if $A \in \mathbf{L}^G$ and $B \in \mathbf{L}^M$ such that $A^{\uparrow} \subseteq B$ and $B^{\downarrow} \subseteq A$. By enlarging these preconcepts we obtain a factorising family of **L**-concepts. In the following we will call such factorisations **L-conceptual**. Note however that this enlargement is not unique. The advantage is thus that we are searching in a smaller set for a covering of the **L**-relation without increasing the number of factors.

From the definition of the factorisation contexts it is straightforward to see that the relationship between objects and attributes from (G, M, I) is explained by the factors of F. Indeed, object g has attribute m if and only if there is a factor f which applies to g and for which m is one of its manifestations. As we are dealing with **L**-sets the notions "applies to" and "is a manifestation of" have truth values. Thus, for a factor f there is a degree $A_f(g)$ to which f applies to g and a degree $B_f(m)$ to which m is a manifestation of f. To obtain the degree to which "f applies to g and m is a manifestation of f", we have to compute $A_f(g) \otimes B_f(m)$.

It was shown in [Běl08] that using **L**-concepts in the factorisation of **L**-contexts yields the smallest possible number of factors. It follows trivially from the crisp case that finding an optimal factorisation is NP-hard. In the light of this fact [Běl08] and [BV09a] provide us with greedy approximation algorithms. Further, in the crisp case we were able to characterise mandatory factors, i.e., factors that have to be present in every factorisation. However, this is not possible in the fuzzy case. We will see in Section 4.4 that different **L**-concepts may yield the same maximal rectangle. Whether this occurs depends highly on the choice of the residuated lattice and the **L**-context.

2.5. Conclusion

In this chapter we presented Factor Analysis of qualitative data in a somewhat new light. First we focused on factorisations with Boolean factors. These factors have reduced expressiveness due to their unary nature. This drawback leads us to the many-valued factorisations, particularly ordinal factorisations, that group the Boolean factors into conceptual scales. The so-obtained factors provide us with a compact representation of the data that is easily interpretable.

Usually in Factor Analysis one plots the factors in order to have a better overview and a graphical representation that allows the user to quickly grasp the outcome of the analysis. Therefore, we have developed such diagrams for our setting. However, unlike ordinary biplots, the ones presented in this chapter do not encode metric data.

The ordinal factorisations and their formal concept analytical interpretation are newly developed. Thus, the method has to be applied in practice to see whether there is more behind it than an abstract mathematical theory. To address this matter, we have run some analyses in Section 2.3. Their outcome shows that ordinal factors seem to be serious competitors of latent attributes used in ordinary Factor Analysis and related techniques, having a similar expressiveness.

3 Triadic Factor Analysis

The importance of triadic (three-way) data in psychology lies in the modelling of human perception. Typical examples are people associating characteristics to other members of their group, evaluating situations, measurements undertaken several times and so on. Factor Analysis and related data analysis techniques were generalised to the triadic case, for example Three-mode Factor Analysis ([Tuc66]), Three-mode Principal Components Analysis ([Kro83]) and Three-way Hierarchical Classes Analysis ([LVDR99], see Section 4.2 for the connection between this research field and the one presented in this chapter). Other applications of three-way factorisation include image analysis, experimental design, spectroscopy, chromatography, see [SBG04]. Due to the wide applicability of Three-way Factor Analysis, it arose as a natural wish to also generalise the concept analytical approach to the triadic case.

This chapter is based on [BGV12, Glo10].[1] Here we present these results from a more general point of view, in accordance with Chapter 2. Further, we also show new results originating from the generalisation of findings from the dyadic case.

The works [BV10b, Glo10] were not the first to study the decomposition of triadic data with Triadic Concept Analysis. The first connection was established in [KSOG94], however, through a different approach. We will comment on this aspect later, after introducing our framework.

As we have already pointed out in Section 2.1, applying decomposition methods to binary data that were designed for real-valued data distorts the meaning of the data and of the results (see for instance [MMG+08, TMGM06]). On the other hand, decomposition methods based on the Boolean matrix product are interpreted in a straightforward way and are therefore preferable ([MMG+08]). As it will turn out, the decomposition involving the Boolean matrix product admits a natural generalisation for the case of triadic binary

[1] During the revision period of [Glo10] it turned out there is yet unpublished work of Bělohlávek and Vychodil dealing with the same subject [BV10b]. Hence, [BGV12] resulted as a joint paper covering the content of the aforementioned works of the authors as well as some new aspects.

data.

In order for the reader to have a better overview of what is happening in this chapter we present its structure: In Section 3.1 we develop the mathematical foundation of our framework proving that triconcepts yield optimal factorisations of triadic data and provide further mathematical insight into the triadic factorisation problem. In Section 3.2 we introduce mappings which transform a description of a given object in terms of attributes and conditions into a description of the same object in terms of factors. The proper algorithms for the factorisation of triadic data are developed in Section 3.3. There are situations in which the users are just roughly interested in the information contained in the data. Then, approximate factorisations are adequate. These are the topic of Section 3.4. In Section 3.5 we generalise some factor analytical tools to our setting. The last section contains an overall conclusion of the results presented in this chapter.

Until explicitly said otherwise, $\mathbb{K}$ will denote a tricontext for the remainder of this chapter.

3.1. Triconceptual Factorisations

We start our work by introducing the so-called *dyadic cuts* of a tricontext, which will prove themselves useful for our framework.

Definition 3.1. For a tricontext $\mathbb{K} = (K_1, K_2, K_3, Y)$ a **dyadic-cut** (shortly **d-cut**) is defined by

$$c^i_\alpha := (K_j, K_k, Y^{jk}_\alpha),$$

where $\{i, j, k\} = \{1, 2, 3\}$ and $\alpha \in K_i$.

Obviously, d-cuts are a special case of $\mathbb{K}^{ij}_{X_k} = (K_i, K_j, Y^{ij}_{X_k})$ for $X_k \subseteq K_k$ and $|X_k| = 1$. Thus, each d-cut is itself a dyadic context. For every tricontext there are three families of d-cuts:

$$c^1 := \{c^1_g := (K_2, K_3, Y^{23}_g)\}_{g \in K_1}, \tag{3.1}$$
$$c^2 := \{c^2_m := (K_1, K_3, Y^{13}_m)\}_{m \in K_2}, \tag{3.2}$$
$$c^3 := \{c^3_b := (K_1, K_2, Y^{12}_b)\}_{b \in K_3}. \tag{3.3}$$

Hence, (3.1) represents cuts in $\mathbb{K}$ for each object $g \in K_1$. The family $\{c^1_g\}_{g \in K_1}$ of d-cuts contains (at most) $|K_1|$ d-cuts. Such a d-cut is itself a dyadic context, namely (K_2, K_3, Y^{23}_g) with $g \in K_1$. For a fixed $g \in K_1$ the d-cut c^1_g contains the incidence relation between the attribute and condition sets of $\mathbb{K}$ under the object g. Equation (3.2) represents cuts in $\mathbb{K}$ for every attribute $m \in K_2$. Such a d-cut contains the relationships between all the objects and all the conditions for the attribute which generated the d-cut. Accordingly, equation (3.3) represents cuts in $\mathbb{K}$ for each condition $b \in K_3$. Such a d-cut contains the relationships between all the objects and all the attributes for the condition which generated the d-cut.

Obviously, one can reconstruct the tricontext $\mathbb{K}$ from the d-cuts by "gluing" them together. For the d-cut families of a tricontext the following equations hold:

$$Y = \bigcup_{g \in K_1} \{g\} \times Y^{23}_g = \bigcup_{m \in K_2} \{m\} \times Y^{13}_m = \bigcup_{b \in K_3} \{b\} \times Y^{12}_b.$$

For different d-cuts of a d-cut family we may have identical dyadic contexts. This happens whenever the incidence relation between K_i and K_j is the same for some elements from K_k with $\{i,j,k\} = \{1,2,3\}$.

We are ready now to start the investigation of conceptual factorisations in a triadic setting.

Definition 3.2. A **factorisation** of a tricontext (K_1, K_2, K_3, Y) consists of formal contexts $(K_1, F, I_1), (K_2, F, I_2)$ and (K_3, F, I_3) such that

$$(a_1, a_2, a_3) \in Y \iff (a_i, f) \in I_i \text{ for some } f \in F \text{ and for } i = 1, 2, 3.$$

The set F is called the **factor set**, its elements the **(triadic Boolean) factors**, and (K_i, F, I_i) $(i = 1, 2, 3)$ are said to be the **factorisation contexts**. We write

$$(K_1, K_2, K_3, Y) = \circ\,((K_1, F, I_1), (K_2, F, I_2), (K_3, F, I_3))$$

to indicate a factorisation.

From the definition of the factorisation contexts it is straightforward to see that they represent relationships between objects and factors, attributes and factors, conditions and factors, respectively. Therefore, $(a_1, f) \in I_1$ means that the object a_1 can be described through the factor f. In the same way, $(a_2, f) \in I_2$ means that the attribute a_2 is a particular manifestation of the factor f, and $(a_3, f) \in I_3$ stands for the fact that the factor f exists under the condition a_3.

As in the dyadic case, we may associate to each factorisation a **factorising family**

$$\{(A_f^1, A_f^2, A_f^3) \mid f \in F\},[2]$$

given by

$$A_f^i := \{a_i \in K_i \mid a_i\; I_i\; f\} \text{ for } i = 1, 2, 3.$$

Such families are easy to characterise: A family $\{(A_f^1, A_f^2, A_f^3) \mid f \in F\}$ is a factorising family of (K_1, K_2, K_3, Y) if and only if

$$Y = \bigcup_{f \in F} A_f^1 \times A_f^2 \times A_f^3,$$

or, expressed differently, if and only if

$$Y = \bigcup_{f \in F} f^{I_1} \times f^{I_2} \times f^{I_3}.$$

Once again the factorising families are precisely the families of triadic preconcepts[3] of (K_1, K_2, K_3, Y) covering all incidences. By enlarging these preconcepts we obtain a factorising family of triconcepts. In the following we will call such factorisations **triconceptual**.

[2] There should be no confusion with this notation and the one introduced for d-cuts. The latter is used in tricontexts whereas this one is defined for the dyadic factorisation contexts.

[3] A triple (A_1, A_2, A_3) is called a triadic preconcept of (K_1, K_2, K_3, Y) if $A_i \subseteq K_i$ $(i = 1, 2, 3)$ such that $A_i^{A_j} \subseteq A_k$ for all $\{i, j, k\} = \{1, 2, 3\}$.

Similarly to the dyadic case this enlargement is not unique. The advantage is thus that we are searching in a smaller set for a covering of the ternary incidence relation of the tricontext without increasing the number of triadic Boolean factors.

We may uniquely determine a factorisation context through the other two. Indeed, we conclude from $a_i\ I_i\ f$ and $a_j\ I_j\ f$ that $a_k \in A^k_f = (A^i_f)^{A^j_f} = (f^{I_i})^{f^{I_j}}$. Note that the derivation operators are applied in $\mathbb{K}$.

There are however different representations among the factorisation contexts. For instance, we may be interested in the relationship between objects and attributes for each factor independent from the conditions. Then, we define the tricontext $\mathbb{K}^J_{12} \coloneqq (K_1, K_2, F, J)$, where

$$(a_1, a_2, f) \in J :\Longleftrightarrow a_1\ I_1\ f \text{ and } a_2\ I_2\ f.$$

Each factor d-cut in $\mathbb{K}^J_{12}$ represents the relationship between the objects and attributes for that factor. However, one may also be interested in the relationship between attributes and conditions independent from the objects. Then, in a similar manner as before, one can build $\mathbb{K}^J_{23}$. To put it more generally, we have tricontexts $\mathbb{K}^J_{ij} \coloneqq (K_i, K_j, F, J)$, where

$$(a_i, a_j, f) \in J :\Longleftrightarrow a_i\ I_i\ f \text{ and } a_j\ I_j\ f.$$

It is easy to reconstruct the factorisation from these contexts. Indeed, we have

$$(a_i, a_j, a_k) \in Y \Longleftrightarrow (a_i, a_j, f) \in J \text{ and } a_k\ I_k\ f \text{ for some } f \in F. \tag{3.4}$$

As we will shortly see, these alternative representations of the factorisation contexts are not just useful for the interpretation.

That a triconceptual factorisation indeed covers the incidence relation of a tricontext is clear. It is however unclear if such factorisations serve the main purpose of a factorisation. The major aim of a factorisation, independent from the nature of the data, is to find an *optimal factorisation*, i.e., the smallest possible number of factors covering the incidence relation. In the dyadic case, we could show that the formal concepts yield an optimal factorisation using the 2-dimension. However, in the triadic case, we are lacking the notion of dimension. Therefore, we will take a slightly different approach to show that the triconcepts indeed yield an optimal factorisation. As this result is the most important of all, it appears in various articles from the literature.

Theorem 3.3 ([BGV12, BV10b, Glo10])**.** *Triconceptual factorisations yield optimal factorisations, i.e., the smallest possible number of factors.*

Proof. Let $\mathbb{K} = (K_1, K_2, K_3, Y)$ be a tricontext and let (K_i, H, J_i) be formal contexts for $i = 1, 2, 3$ such that

$$Y = \bigcup_{h \in H} h^{J_1} \times h^{J_2} \times h^{J_3}.$$

Obviously, for every $h \in H$ we have $h^{J_1} \times h^{J_2} \times h^{J_3} \subseteq Y$. Therefore, $h^{J_1} \times h^{J_2} \times h^{J_3}$ must be contained in some maximal rectangular box full of crosses. We know that these correspond to triconcepts. Thus, there is $(A^h_1, A^h_2, A^h_3) \in \mathfrak{T}(\mathbb{K})$ such that

$$h^{J_1} \times h^{J_2} \times h^{J_3} \subseteq A^h_1 \times A^h_2 \times A^h_3$$

for every $h \in H$. Now, we build factorisation contexts (K_i, F, I_i) $(i = 1,2,3)$ such that for every $h \in H$, $f^{I_i} = A_i^h$ holds for some $f \in F$ and $i = 1,2,3$. Thus, $h^{J_i} \subseteq f^{I_i}$ for all $h \in H$ and the corresponding $f \in F$. We have the following:

$$Y = \bigcup_{h \in H} h^{J_1} \times h^{J_2} \times h^{J_3} \subseteq \bigcup_{f \in F} f^{I_1} \times f^{I_2} \times f^{I_3} \subseteq Y.$$

Since $|F| \leq |H|$, we are done. □

Note that a tricontext can be factorised using different sets of triconcepts and that the cardinalities of these sets may differ from one another. However, the above theorem states that among all these triconceptual factorisations there is at least one with the smallest possible number of factors.

We know now that triconcepts provide optimal factorisations. But how large can such a triconceptual factorisation be? We give an upper bound for this matter:

Theorem 3.4 ([BGV12]). *Let $\mathbb{K} = (K_1, K_2, K_3, Y)$ be a tricontext with $|K_1| = p, |K_2| = q$ and $|K_3| = r$ and let F be the factor set of an optimal triconceptual factorisation of $\mathbb{K}$. Then,*

$$|F| \leq \min\{pq, pr, qr\}.$$

Proof. We will use the ik-bounds $\mathfrak{b}_{13}(a_1, a_3)$ (see Equation (1.15) on page 21) to obtain a triconceptual factorisation with factor set F, i.e., let

$$f^{I_i} := \{B_i \mid B_i \text{ } i\text{-th component of } \mathfrak{b}_{13}(a_1, a_3), a_1 \in K_1, a_3 \in K_3\}$$

for $i \in \{1,2,3\}$ and $f \in F$. Obviously, we have $n := |F| \leq |K_1| \cdot |K_3| = pr$. As $\mathfrak{b}_{13}(a_1, a_3)$ is a triconcept we have that

$$\bigcup_{l=1}^{n} f_l^{I_1} \times f_l^{I_2} \times f_l^{I_3} \subseteq Y.$$

On the other hand, if $(a_1, a_2, a_3) \in Y$, then $(a_1, a_2, a_3) \in \mathfrak{b}_{13}(a_1, a_3)$ holds and therefore $(a_1, a_2, a_3) \in \bigcup_{l=1}^{n} f_l^{I_1} \times f_l^{I_2} \times f_l^{I_3}$. Thus, we have shown $|F| \leq pr$. Similarly, one can prove $|F| \leq pq$ and $|F| \leq qr$, finishing the proof. □

Now we describe mandatory factors of $\mathbb{K}$, i.e., triconcepts that have to be present in every factorisation of $\mathbb{K}$. In the dyadic case, mandatory factors of a dyadic context (G, M, I) are exactly those dyadic concepts of (G, M, I) that are both object concepts and attribute concepts ([BV10a]). In the triadic case, the concepts of (K_1, K_2, K_3, Y) that can be seen as analogous to object and attribute concepts are $\mathfrak{b}_{ij}(a_i, a_j)$ for $\{i, j\} = \{1, 2\}$, $\mathfrak{b}_{jk}(b_j, b_k)$ for $\{j, k\} = \{2, 3\}$, and $\mathfrak{b}_{ik}(c_i, c_k)$ for $\{i, k\} = \{1, 3\}$ where $a_1, c_1 \in K_1$, $a_2, b_2 \in K_2$, $b_3, c_3 \in K_3$.

Lemma 3.5 ([BGV12]). *Let (K_1, K_2, K_3, Y) be a tricontext, $a_1, c_1 \in K_1$, $a_2, b_2 \in K_2$, and $b_3, c_3 \in K_3$. If there exist $\{i_a, j_a\} = \{1, 2\}$, $\{i_b, j_b\} = \{2, 3\}$ and $\{i_c, j_c\} = \{1, 3\}$ such that*

$$\mathfrak{b}_{i_a j_a}(a_{i_a}, a_{j_a}) = \mathfrak{b}_{i_b j_b}(b_{i_b}, b_{j_b}) = \mathfrak{b}_{i_c j_c}(c_{i_c}, c_{j_c}),$$

then the triconcept (D_1, D_2, D_3) described equivalently by any of the formulas $\mathfrak{b}_{i_x j_x}(x_{i_x}, x_{j_x})$ is the only triconcept for which

$$\{a_1, c_1\} \times \{a_2, b_2\} \times \{b_3, c_3\} \subseteq D_1 \times D_2 \times D_3. \tag{3.5}$$

Proof. (3.5) follows from the fact that the p-th and q-th component of $\mathfrak{b}_{pq}(x_p, x_q)$ contain x_p and x_q, respectively. Let $(C_1, C_2, C_3) \in \mathfrak{T}(K_1, K_2, K_3, Y)$ be a triconcept that satisfies $\{a_1, c_1\} \times \{a_2, b_2\} \times \{b_3, c_3\} \subseteq C_1 \times C_2 \times C_3$. Then, $\{a_{i_a}\} \subseteq C_{i_a}$ and $\{a_{j_a}\} \subseteq C_{j_a}$. Since D_3 is the third component of $\mathfrak{b}_{i_a j_a}(a_{i_a}, a_{j_a})$, we have

$$D_3 = \{a_{i_a}\}^{\{a_{j_a}\}} \supseteq C_{i_a}^{C_{j_a}} = C_3.$$

In a similar way, one proves $D_1 \supseteq C_1$ and $D_2 \supseteq C_2$. Since $(C_1, C_2, C_3) \in \mathfrak{T}(K_1, K_2, K_3, Y)$, the maximality of triconcepts implies $(C_1, C_2, C_3) = (D_1, D_2, D_3)$. □

Definition 3.6. A triconcept $(D_1, D_2, D_3) \in \mathfrak{T}(K_1, K_2, K_3, Y)$ is called **mandatory** if in every triconceptual factorisation of (K_1, K_2, K_3, Y) with factor set F there is $f \in F$ such that $f^{I_i} = D_i$, $i = 1, 2, 3$.

The following lemma is a crucial observation when it comes to describing mandatory triconcepts.

Lemma 3.7 ([BGV12]). *For a triconcept $\mathfrak{d} \in \mathfrak{T}(K_1, K_2, K_3, Y)$ and $a_1 \in K_1$, $a_2 \in K_2$ and $a_3 \in K_3$, the following conditions are equivalent:*

(i) *$\mathfrak{d}$ is the only concept of (K_1, K_2, K_3, Y) that covers (a_1, a_2, a_3).*

(ii) *There exists $\{i, j, k\} = \{1, 2, 3\}$ such that*

$$\mathfrak{d} = \mathfrak{b}_{ij}(a_i, a_j) = \mathfrak{b}_{jk}(a_j, a_k) = \mathfrak{b}_{ik}(a_i, a_k). \tag{3.6}$$

(iii) *For every $\{i, j, k\} = \{1, 2, 3\}$ we have*

$$\mathfrak{d} = \mathfrak{b}_{ij}(a_i, a_j) = \mathfrak{b}_{jk}(a_j, a_k) = \mathfrak{b}_{ik}(a_i, a_k).$$

Proof. "(i)⇒(ii)": If $\mathfrak{d}$ covers (a_1, a_2, a_3), then a_1, a_2, and a_3 are related. Therefore, $\mathfrak{b}_{ij}(a_i, a_j)$, $\mathfrak{b}_{jk}(a_j, a_k)$, and $\mathfrak{b}_{ik}(a_i, a_k)$ all cover (a_1, a_2, a_3). Since $\mathfrak{d}$ is the only triconcept covering (a_1, a_2, a_3), (3.6) follows.

"(ii)⇒(i)": Follows from Lemma 3.5.

"(ii)⇒(iii)": Assume that (3.6) holds for some $\{i, j, k\} = \{1, 2, 3\}$. Then, $\mathfrak{b}_{ji}(a_j, a_i)$ has the same k-th component as $\mathfrak{b}_{ij}(a_i, a_j)$ and therefore contains a_k. Since the i-th and j-th components of $\mathfrak{b}_{ji}(a_j, a_i)$ contain a_i and a_j, $\mathfrak{b}_{ji}(a_j, a_i)$ covers (a_1, a_2, a_3). Because (ii) implies (i), $\mathfrak{d}$ is the only triconcept that covers (a_1, a_2, a_3), whence $\mathfrak{d} = \mathfrak{b}_{ji}(a_j, a_i)$. In the same way one proves that $\mathfrak{d} = \mathfrak{b}_{kj}(a_k, a_j)$ and $\mathfrak{d} = \mathfrak{b}_{ki}(a_k, a_i)$, yielding (iii).

"(iii)⇒(ii)": Trivial. □

Theorem 3.8 ([BGV12]). *A concept $\mathfrak{d} \in \mathfrak{T}(K_1, K_2, K_3, I)$ is mandatory if and only if there exist $a_1 \in K_1$, $a_2 \in K_2$, and $a_3 \in K_3$ that satisfy* (ii)*, or, equivalently* (iii) *of Lemma 3.7.*

Proof. Clearly, $\mathfrak{d}$ is mandatory if and only if there exist $a_1 \in K_1$, $a_2 \in K_2$, and $a_3 \in K_3$ such that $\mathfrak{d}$ is the only triconcept that covers a_1, a_2, and a_3. The claim thus follows from Lemma 3.7. □

Remark 3.9. A claim such as (iii) in the scenario from Lemma 3.7 does not hold in Lemma 3.5. In fact, consider a tricontext $\mathbb{K} = (\{a_1, c_1\}, \{a_2, b_2, x_2\}, \{b_3, c_3\}, Y)$ with

$$Y = \{a_1, c_1\} \times \{a_2, b_2\} \times \{b_3, c_3\} \cup \{a_1\} \times \{x_2\} \times \{b_3, c_3\}$$

such that x_2 is distinct from both a_2 and b_2. Then, $\mathfrak{d} := (\{a_1, c_1\}, \{a_2, b_2\}, \{b_3, c_3\})$ is the only triconcept covering $\{a_1, c_1\} \times \{a_2, b_2\} \times \{b_3, c_3\}$. On the one hand we have the equalities $\mathfrak{d} = \mathfrak{b}_{12}(a_1, a_2) = \mathfrak{b}_{23}(b_2, b_3) = \mathfrak{b}_{13}(c_1, c_3)$. However, on the other hand we have that $\mathfrak{b}_{21}(a_2, a_1) = (\{a_1\}, \{a_1, b_2, x_2\}, \{b_3, c_3\}) \neq \mathfrak{d}$.

Example 3.10. Let us consider the example about hostels from Figure 1.3 (page 20). The factorisation contexts of the triconceptual factorisation are displayed in Figure 3.1 and 3.2. The triconcepts used in the factorisation are drawn larger in the trilattice from Figure 1.4 (page 23).

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
Nuevo S.	×				×			
Samay	×		×	×	×	×		
Oasis B.	×	×	×	×	×		×	×
One		×		×	×	×	×	×
Ole B.		×		×	×	×	×	×
Garden B.	×	×	×		×	×		

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
character		×				×		×
safety		×		×		×	×	×
location	×	×	×		×	×	×	
staff		×	×	×	×	×	×	×
fun		×					×	
cleanliness		×	×	×		×	×	×

Figure 3.1.: Factorisation contexts for objects and attributes

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
hostelworld	×		×	×				×
hostels	×	×	×	×	×		×	×
hostelbookers	×		×	×	×	×	×	

Figure 3.2.: Factorisation context for conditions

The factors also have a verbal description which can be read from the factorisation contexts. In our case the factors are the overall ratings of the users. The context on the left in Figure 3.1 associates to each object the factors which describe it (hostels are described by their ratings), the context on the right associates to each attribute the factors which contain it (which services are taken into account for each rating) and the context in Figure 3.2 shows which factor exists under which condition (which users contributed to the formation of each rating). Let us once again number the elements from K_i consequently, i.e., $K_i := \{0, \dots, |K_i|\}$ with $i \in \{1, 2, 3\}$. For example, f_1 stands for *best location* because the users from all platforms ($f_1^{I_3} = K_3$) have rated the hostels from $f_1^{I_1}$ as having the best location. Factor f_4 can be interpreted as hostels with *best facilities* because all users agreed that the hostels $1, 2, 3, 4$ are excellent concerning safety, staff and cleanliness. Factor f_5 shows that the users from platform 1 and 2 consider that every hostel is excellent regarding location and staff. The *best deals*, according to the users from the second platform, are represented by f_2. On the other hand, the users from the third platform consider the hostels from $f_6^{I_1}$ the *best deals*.

The triconceptual factorisation offers the possibility to describe the hostels through 8 factors while in the tricontext they are described through 6 attributes under 3 conditions, i.e., 18 items. Thus, the factors yield a more parsimonious way of information representation.

Let us see how some results from Section 2.1 can be translated into the triadic setting.

Proposition 3.11. *For any triconceptual factorisation with factor set F the intersection of the attribute orders of (K_i, F, I_i) and (K_j, F, I_j) is contained in the dual attribute order of (K_k, F, I_k) for $\{i, j, k\} = \{1, 2, 3\}$.*

Proof. Suppose we have

$$f^{I_i} \subseteq h^{I_i} \text{ and } f^{I_j} \subseteq h^{I_j}$$

for some $f, h \in F$. Then,

$$A_f^i = f^{I_i} \subseteq h^{I_i} = A_h^i \text{ and } A_f^j = f^{I_j} \subseteq h^{I_j} = A_h^j$$

both hold. As two components of a triconcept uniquely determine the third one, we immediately obtain

$$A_f^k = (A_f^i \times A_f^j)^{(k)} \supseteq (A_h^i \times A_h^j)^{(k)} = A_h^k,$$

finishing the proof. □

Now we may characterise the factorisation contexts. The conditions are the triadic analogons of the dyadic ones. As we will see, the alternative definition of factorisation contexts given by condition (3.4) (page 70) turns out to be very useful for this task.

Theorem 3.12. *$(K_1, F, I_1), (K_2, F, I_2)$ and (K_3, F, I_3) are factorisation contexts of a triconceptual factorisation if and only if every intent of (K_i, F, I_i) is a modus of the tricontext $(K_j, K_k, F, (K_j \times K_k \times F \setminus J))$, for all $\{i, j, k\} = \{1, 2, 3\}$ and $j < k$.*

Proof. First let us note the following equivalences:

$$\begin{aligned}(a_1, a_2, a_3) \in Y &\iff (a_i, f) \in I_i \text{ for some } f \in F \text{ and } i = 1, 2, 3\\ &\iff (a_i, a_j, f) \in J \text{ and } a_k\ I_k\ f \text{ for some } f \in F \text{ and } \{i, j, k\} = \{1, 2, 3\}\\ &\iff \bigcap_{i=1}^{3} a_i^{I_i} \neq \varnothing.\end{aligned}$$

Hence, we also have

$$(a_1, a_2, a_3) \notin Y \iff a_k^{I_k} \subseteq F \setminus \{a_i^{I_i} \cap a_j^{I_j}\}.$$

Let (K_i, F, I_i) ($i = 1, 2, 3$) be factorisation contexts. We have to show that for any $a_i \in K_i$ the object intent $a_i^{I_i}$ is the intersection of modi of

$$(K_j, K_k, F, J^c) := (K_j, K_k, F, (K_j \times K_k \times F \setminus J)).$$

Assume this is not true. Then, there must be some $f \in F$ which is contained in all modi of (K_j, K_k, F, J^c) that contain $a_i^{I_i}$ but which does not belong to $a_i^{I_i}$, i.e., $f \notin a_i^{I_i}$. This, on the other hand, is equivalent to $a_i \notin f^{I_i} = A^i_f$, and therefore $a_i \notin (A^j_f \times A^k_f)^{(i)} = (f^{I_j} \times f^{I_k})^{(i)}$, where $(-)^{(i)}$ is applied in (K_1, K_2, K_3, Y). Thus, there must be elements $a_j \in f^{I_j}$ and $a_k \in f^{I_k}$ such that $(a_i, a_j, a_k) \notin Y$. As we have noted at the beginning, the latter is equivalent to $a_i^{I_i} \subseteq F \setminus (a_j^{I_j} \cap a_k^{I_k})$. Since $F \setminus (a_j^{I_j} \cap a_k^{I_k})$ is a modus of (K_j, K_k, F, J^c) containing $a_i^{I_i}$ but not f, we have reached a contradiction.

For the converse direction suppose the three conditions are fulfilled. The derivation operators of (K_j, K_k, F, J) are denoted by $(-)^{(j)_J}, (-)^{(k)_J}$ and $(-)^{(l)_J}$. For (K_j, K_k, F, J^c) we proceed similarly with the subscription J^c. We have to show that the factorisation is triconceptual. In order to do so we have to prove that $(f^{I_j})^{f^{I_k}} \subseteq f^{I_i}$ or equivalently that $(f^{I_j} \times f^{I_k})^{(i)} \subseteq f^{I_i}$ for any $f \in F$ and for all $\{i, j, k\} = \{1, 2, 3\}$. We will show the second statement. To this end, suppose $a_i \notin f^{I_i}$ which is equivalent to $f \notin a_i^{I_i}$. Since $f \in F$ there must exist elements $b_i \in K_i, a_j \in K_j$ and $a_k \in K_k$ such that $(b_i, a_j, a_k) \in Y$, i.e., b_i, a_j, a_k are related with f in the factorisation contexts. Further, since $a_i^{I_i}$ is an intent of (K_i, F, I_i), there must be, according to the condition, a modus $(a_j \times a_k)^{(l)_{J^c}}$ of (K_j, K_k, F, J^c) containing $a_i^{I_i}$ but not f (as discussed above, $(a_j \times a_k)^{(l)_J} = \varnothing$ or equivalently $(a_j \times a_k)^{(l)_{J^c}} = F$ cannot happen). Expressed differently, there are $a_j \in f^{I_j}$ and $a_k \in f^{I_k}$ such that $a_i^{I_i} \cap (a_j \times a_k)^{(l)_J} = \varnothing$. Consequently, for any $h \in (a_j \times a_k)^{(l)_J}$ we have $(a_i, h) \notin I_i$. Thus, $(a_i, a_j, a_k) \notin Y$ and consequently $a_i \notin (f^{I_j} \times f^{I_k})^{(i)}$. □

Trivially, the triconceptual factorisation is “invariant” under the clarification and the interchange of the roles of objects, attributes and conditions in tricontexts in the sense detailed for the dyadic case in Proposition 2.6 and Proposition 2.7 (page 42).

For computational simplicity we will present also the matrix notation for triconceptual factorisations. This allows a better manipulation of the factorisation contexts when designing and implementing software. By replacing in the cross table the crosses by 1’s and the blanks by 0’s we obtained in the dyadic case the corresponding Boolean matrix. If we proceed alike in the triadic case we obtain from the three-dimensional cross table a **Boolean** $3d$**-matrix** (shortly $3d$**-matrix**) that is a rectangular box $\mathbb{B}_{p \times q \times r}$ such that $b_{ijk} \in \{0, 1\}$ for all $i \in \{1, \dots, p\}, j \in \{1, \dots, q\}, k \in \{1, \dots, r\}$. For a $3d$-matrix $\mathbb{B}$ we write $\mathbb{B} = \mathbb{B}_1 \mid \dots \mid \mathbb{B}_r$, where the $\mathbb{B}_1, \dots, B_r$ are $p \times q$ binary matrices called **layers**.

In order to define the *Boolean* $3d$*-matrix product* for a triconceptual factorisation with factor set F, we first have to build matrices A_F, B_F, C_F which basically are the Boolean matrix representation of the factorisation contexts, i.e.,

$$(A_F)_{il} := \begin{cases} 1, & i \in f_l^{I_1}, \\ 0, & i \notin f_l^{I_1}, \end{cases} \quad (B_F)_{jl} := \begin{cases} 1, & j \in f_l^{I_2}, \\ 0, & j \notin f_l^{I_2}, \end{cases} \quad (C_F)_{kl} := \begin{cases} 1, & k \in f_l^{I_3}, \\ 0, & k \notin f_l^{I_3}, \end{cases}$$

for all $l \in \{1, \dots, |F|\}$.

Definition 3.13. For $p \times n, q \times n$ and $r \times n$ binary matrices P, Q and R, the **Boolean** $3d$**-matrix product** (shortly $3d$**-product**) is defined as the ternary operation

$$\circ (P, Q, R)_{ijk} := \bigvee_{l=1}^{n} P_{il} \cdot Q_{jl} \cdot R_{kl}, \tag{3.7}$$

with $i \in \{1, \dots, p\}, j \in \{1, \dots, q\}$ and $k \in \{1, \dots, r\}$. As in the dyadic case, $\bigvee$ denotes the maximum and $\cdot$ is the usual product.

As we have already stated at the beginning, the first link between factorisation of triadic data and Triadic Concept Analysis was established in [KSOG94]. The authors were concerned with a covering of Y by triconcepts, however in a way different from the one proposed here. Indeed, they search for a triconcept that covers a large part of Y. The subrelation of Y corresponding to such a triconcept is then removed from Y. Such a removal continues until Y contains no entries. It is easy to find an example showing that the method from [KSOG94] may produce decompositions with a larger than optimal number of factors.

In [Mie11] the factorisation of triadic data was also studied but without any connection to Triadic Concept Analysis and without the factor analytical interpretation presented here. The paper contains various algorithms for the factorisation problem. There is a small overlap in the theoretical part, namely it also contains the assertion from our Theorem 3.4.

3.2. Transformation Between the Attribute and Condition Space and the Factor Space

The results of this section were already presented in [BGV12]. In contrast to that work, however, we present the problem in the notation of Triadic Concept Analysis and elaborate further on the topics.

In Section 2.1 we have already got acquainted with transformations between the space of attributes and the space of factors. These were developed for the dyadic case in [BV10a] and were utilised in [Out10] for improving classification of binary data. In this section we aim to generalise these mappings to the triadic case. That is, given a tricontext $\mathbb{K} = (K_1, K_2, K_3, Y)$ (or equivalently a $p \times q \times r$ $3d$-binary matrix Y) with its factorisation contexts (K_i, F, I_i) $(i = 1, 2, 3)$, one wishes for mappings that transform a description of a given object in terms of attributes and conditions into a description of the same object in terms of factors. That is, one asks for transformations between the attribute $\times$ condition space and the factor space.

In the attribute $\times$ condition space, the object $g \in K_1$ is represented by the object d-cut of g, i.e., by (K_2, K_3, Y_g^{23}). In the factor space, g is represented by a row of (K_1, F, I_1), namely by g^{I_1}. As presented in [BGV12], these transformations can be modelled as follows: Let $n \coloneqq |F|$. Define mappings $\varphi : \{0,1\}^{q\times r} \to \{0,1\}^n$ and $\psi : \{0,1\}^n \to \{0,1\}^{q\times r}$ for $P \in \{0,1\}^{q\times r}$ and $S \in \{0,1\}^n$ by

$$(\varphi(P))_l \coloneqq \bigwedge_{j=1}^{q} \bigwedge_{k=1}^{r} \left((B_{jl} \cdot C_{kl}) \to P_{jk}\right), \tag{3.8}$$

$$(\psi(S))_{jk} \coloneqq \bigvee_{l=1}^{n} \left(S_l \cdot B_{jl} \cdot C_{kl}\right), \tag{3.9}$$

where $B_{j\_}$ and $C_{k\_}$ are the j-th and k-th rows of the factorisation matrices corresponding to (K_2, F, I_2) and (K_3, F, I_3), respectively, for $l \in \{1, \ldots, n\}, j \in \{1, \ldots, q\}$ and $k \in \{1, \ldots, r\}$. Here, $\to$, $\cdot$, $\bigwedge$, and $\bigvee$ denote the truth function of classical implication, the usual product, minimum, and maximum, respectively. If P represents a description of an object i in terms of attributes and conditions and S a description in terms of factors, then given the interpretation of the factor matrices, φ and ψ have the following meaning: (3.8) says that factor l applies to i if and only if object i has every attribute j under every condition k such that j is a manifestation of l and k is one of the conditions under which l appears; (3.9) says that object i has attribute j under condition k if there exists a factor l such that l applies to i, j is a manifestation of l, and k is one of the conditions under which l appears.

We may redefine these mappings in the language of Triadic Concept Analysis as follows: $\varphi : \mathfrak{P}(K_2 \times K_3) \to \mathfrak{P}(F)$ and $\psi : \mathfrak{P}(F) \to \mathfrak{P}(K_2 \times K_3)$ given by

$$\varphi(P) := \{f \in F \mid f^{I_2} \times f^{I_3} \subseteq P\}, \tag{3.10}$$

$$\psi(S) := \{(k_2, k_3) \in f^{I_2} \times f^{I_3} \mid f \in S\}, \tag{3.11}$$

for $P \in \mathfrak{P}(K_2 \times K_3)$ and $S \in \mathfrak{P}(F)$.

The next theorem shows that φ and ψ can be considered as appropriate transformations between the attribute $\times$ condition space and the factor space.

Theorem 3.14 ([BGV12])**.** *For any $g \in K_1$ we have*

$$\varphi(g^{(1)}) = g^{I_1} \quad \text{and} \quad \psi(g^{I_1}) = g^{(1)}.$$

That is, φ maps the object d-cuts of (K_1, K_2, K_3, Y) to the rows of (K_1, F, I_1), and ψ maps the rows of (K_1, F, I_1) to the object d-cuts of (K_1, K_2, K_3, Y).

Proof. The assertions follow directly from (3.10) and (3.11). We have

$$\varphi(g^{(1)}) = \{f \in F \mid f^{I_2} \times f^{I_3} \subseteq g^{(1)}\},$$
$$\psi(g^{I_1}) = \{(k_2, k_3) \in f^{I_2} \times f^{I_3} \mid f \in g^{I_1}\}$$

for any $g \in K_1$. Indeed, $\psi(g^{I_1}) = g^{(1)}$ results from the definition of triconceptual factorisations (Definition 3.2). For the first item let $g \in K_1$ such that $g \in f^{I_1}$ for some $f \in F$. Thus, for any $(k_2, k_3) \in f^{I_2} \times f^{I_3}$ we have $(g, k_2, k_3) \in Y$ by Definition 3.2. Hence, $(k_2, k_3) \in g^{(1)}$ and thereby $f^{I_2} \times f^{I_3} \subseteq g^{(1)}$. The converse is similar. □

The next lemma, which follows easily from the definition of the two mappings, shows that φ and ψ form an isotone Galois connection.

Lemma 3.15 ([BGV12])**.** *For $P, Q \in \mathfrak{P}(K_2 \times K_3)$ and $S, T \in \mathfrak{P}(F)$, we have*

$$P \subseteq Q \Longrightarrow \varphi(P) \subseteq \varphi(Q), \tag{3.12}$$

$$S \subseteq T \Longrightarrow \psi(S) \subseteq \psi(T), \tag{3.13}$$

$$\psi(\varphi(P)) \subseteq P, \tag{3.14}$$

$$S \subseteq \varphi(\psi(S)). \tag{3.15}$$

Keeping in mind the results of Theorem 3.14 the conditions (3.12)–(3.15) represent natural requirements. Indeed, (3.12) says that the more attributes under more conditions an object has, the more factors apply, while (3.13) asserts the converse relationship. Consequently, for an object, having “more attributes” and “more conditions” is positively correlated with having “more factors”. (3.14) means that common attributes in common modi associated to all the factors which apply to a given object are contained in the collection of all attributes and modi possessed by that object. The meaning of (3.15) is analogous.

Corollary 3.16. *For $P, P_j \in \mathfrak{P}(K_2 \times K_3)$ and $S, S_j \in \mathfrak{P}(F)$ with $j \in J$, we have*

$$\begin{aligned}
\varphi(P) &= \varphi\psi\varphi(P), \\
\psi(S) &= \psi\varphi\psi(S), \\
\varphi\Big(\bigcap_{j\in J} P_j\Big) &= \bigcap_{j\in J} \varphi\left(P_j\right), \\
\psi\Big(\bigcup_{j\in J} S_j\Big) &= \bigcup_{j\in J} \psi\left(S_j\right).
\end{aligned}$$

A geometry behind the transformations is described by the following assertion. For $P \in \mathfrak{P}(K_2 \times K_3)$ and $S \in \mathfrak{P}(F)$, we put

$$\begin{aligned}
\varphi^{-1}(S) &:= \{P \in \mathfrak{P}(K_2 \times K_3) \mid \varphi(P) = S\}, \\
\psi^{-1}(P) &:= \{S \in \mathfrak{P}(F) \mid \psi(S) = P\}.
\end{aligned}$$

Recall that for a universe U, a subset $C \subseteq \mathfrak{P}(U)$ is called *convex* if $Y \in C$ whenever $X \subseteq Y \subseteq Z$ for some $X, Z \in C$.

Theorem 3.17 ([BGV12]). (1) *$\varphi^{-1}(S)$ is a convex partially ordered subspace of the attribute and condition space, and $\psi(S)$ is the least element of $\varphi^{-1}(S)$.*
(2) *$\psi^{-1}(P)$ is a convex partially ordered subspace of the factor space, and $\varphi(P)$ is the largest element of $\psi^{-1}(P)$.*

Proof. By standard application of the properties of isotone Galois connections. □

According to Theorem 3.17, the space $K_2 \times K_3$ of attributes and conditions and the space F of factors are partitioned into an equal number of convex subsets. The subsets of the space of attributes and conditions have least elements and the subsets of the space of factors have greatest elements. Hence, φ maps every element of any convex subset of the space of attributes and conditions to the greatest element of the corresponding subset of the factor space, whereas ψ maps every element of some convex subset of the space of factors to the least element of the corresponding convex subset of the space of attributes and conditions.

3.3. Algorithms

It follows trivially from the dyadic case that finding an optimal triconceptual factorisation is an instance of the set covering problem. Thus, the problem is NP-hard and the corresponding decision problem is NP-complete. Recall that there exists a greedy approximation algorithm for the set covering optimisation problem which achieves an approximation ratio $\leq \ln(|U|) + 1$, see [CLRS01]. In our setting the universe U to be covered corresponds to the incidence relation of the tricontext. The family $\mathcal{S}$ of subsets of the universe U that is used for finding a cover corresponds to the set of all triconcepts $\mathfrak{T}(K_1, K_2, K_3, Y)$. In this setting, we are looking for $\mathcal{C} \subseteq \mathcal{S}$ as small as possible such that $\bigcup \mathcal{C} = U$.

In [BGV12] we presented two algorithms for the triconceptual factorisation problem. As this part of the work goes back to the other authors, we will present here just a brief overview of the results obtained for the algorithms. The first algorithm, Algorithm 1, computes all triconcepts of the tricontexts. Afterwards, it searches in a greedy manner for those triconcepts which cover most part of the incidence relation. This algorithm can be significantly improved to get a better performance in terms of computation time. Indeed, the main drawback of it is that it first computes the set $\mathcal{S}$ of all triconcepts of (K_1, K_2, K_3, Y) and then selects (usually) a small subset of it for the triconceptual factorisation by iteratively going through $\mathcal{S}$. It is well-known that the number of elements in $\mathcal{S}$ is usually large and may be larger than exponential with respect to $\min(|K_1|, |K_2|, |K_3|)$ in the worst case. As a result, since the first algorithm iterates through $\mathcal{S}$ every time it computes a new factor, it has an exponential time delay complexity. Therefore, a second algorithm, Algorithm 2, was designed to overcome this problem. It finds the factorisation directly without the need to compute all triconcepts. The way the second algorithm works leads to a polynomial time delay complexity. The algorithm is based on the idea of incremental modification of a promising triconcept by extending its intent and modus such that the concept covers as much of the remaining values in U as possible.

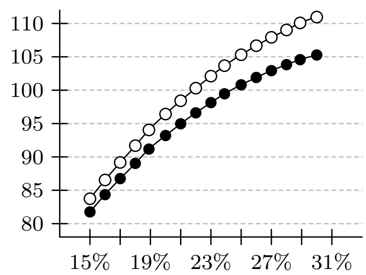

Figure 3.3.: Empirical comparison of the numbers of factors computed by Algorithm 1 and Algorithm 2 (x-axis: percentage of 1s in randomly generated three-dimensional matrices; y-axis: numbers of factors; black nodes: average numbers of factors computed by Algorithm 1; white nodes: average numbers of factors computed by Algorithm 2).

Remark 3.18. In practice, Algorithm 1 produces better results than Algorithm 2 in terms of the number of computed factors. This is expected since Algorithm 1 uses the whole set of triconcepts to search for the factors. On the other hand, Algorithm 2 is faster than Algorithm 1 by an order of magnitude. This is because the expensive operation of computing all triconcepts is omitted in Algorithm 2. Even if Algorithm 2 delivers worse results (on average), our empirical experiments have shown that the average difference of results obtained by both the algorithms is negligible if we compute the triconceptual factorisation for tricontexts with a relatively low number of elements contained in Y.

The results of the experiments related to this issue are depicted in Figure 3.3. The graph in Figure 3.3 shows two curves corresponding to the average numbers of factors computed by Algorithm 1 (curve with black nodes) and Algorithm 2 (curve with white nodes) using a sample of $300,000$ randomly generated tricontexts of various sizes, the incidence relation of each context containing about ten thousand entries. One can see that with growing density of crosses in the context, the difference between the average numbers of factors grows and Algorithm 2 computes more factors than Algorithm 1. With smaller percentages of crosses, the difference is negligible. Since most large real-world data sets represented by tricontexts are typically sparse (very low percentages of crosses), Algorithm 2 is a preferred choice since it delivers almost as good results as Algorithm 1 in considerably less time.

3.4. Approximate Factorisations

For a large tricontext numerous factors are needed, even for an optimal factorisation. In cases where the users would like to know just roughly about the information contained in the data, approximate factorisation is adequate. Then the requirement on the model is weaker in the sense that it has to fit the data set only partially. These approximations can be of two kinds, namely with *negative discrepancies* and with *positive discrepancies*. In the first type there are crosses in the tricontext which remain uncovered, i.e., not the entire ternary relation of the tricontext is explained, and the second type covers some blank entries with crosses, i.e., one explains more than there is. Trivially, one can combine the two kinds.

Negative Discrepancies

For a given tricontext (K_1, K_2, K_3, Y) we will be searching for approximate factorisation contexts (K_i, F, I_i) $(i = 1, 2, 3)$ such that

$$\bigcup_{f \in F} f^{I_1} \times f^{I_2} \times f^{I_3} \subseteq Y.$$

These inexact factorisation contexts approximate Y from below. By adding further elements to F, we obtain a more precise approximation. Such approximate factorisations for the dyadic case were presented in [BV10a].

For the tricontext in Figure 1.3 (page 20) we consider the approximate factorisations given by the factor sets $F_1 = \{f_4\}, F_2 = \{f_4, f_5\}, F_3 = \{f_4, f_5, f_2\}$ and $F_4 = \{f_4, f_5, f_2, f_6\}$ (see figures 3.1 and 3.2 on page 73). Notice that just using an optimal factor we are able to cover 45% of the incidence relation of the tricontext. The set F_2, containing just a quarter of the total number of factors, covers 65%. The factor sets F_3 and F_4 also tighten the approximation but not as drastically as the first two factors. Although the first four factors cover 85% of the incidence relation, the last 4 factors are needed to cover the remaining 15%. While the exact optimal factorisation contains eight factors we can explain 65% of the data by using just two factors through the approximate factorisation.

Usually one is interested in the degree of approximation, i.e., a ratio to which the approximate factorisation contexts explain the tricontext. This ratio is given by

$$\frac{|\bigcup_{f \in F} f^{I_1} \times f^{I_2} \times f^{I_3}|}{|Y|}. \tag{3.16}$$

Obviously, the ratio equals 1 (or 100%) if and only if F is an exact factorisation. From the point of view of approximation, one is interested in finding, given a ratio r, a factor set F such that the degree of approximation given by (3.16) is at least r. In other words, we are interested in factors which explain at least $100 \cdot r\%$ of the input data. Notice that the algorithms presented in Subsection 3.3 ([BGV12]) can be easily modified to compute approximate factorisations by adding an additional parameter r and a new halting condition which stops looking for further factors whenever the threshold value r has been reached.

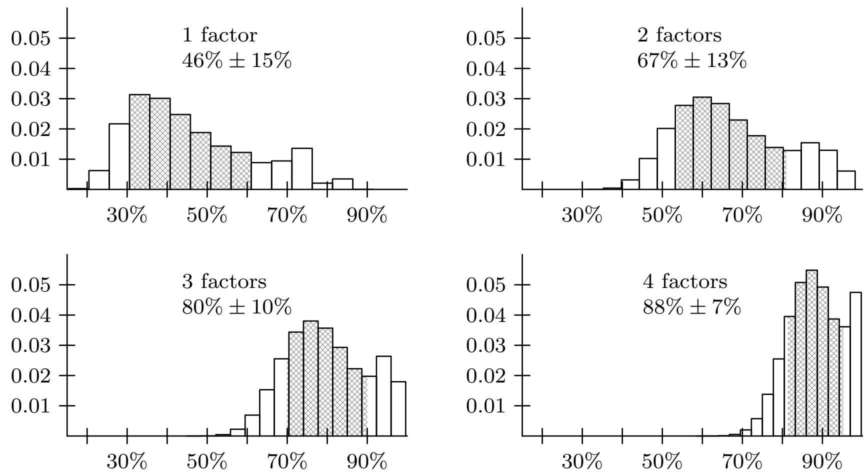

Figure 3.4.: Relative frequency histograms of degrees of approximation obtained by using one up to four factors of randomly-generated 7-factorisable three-dimensional matrices.

Remark 3.19. In [BGV12] a series of experiments were performed to explore the behaviour of approximate decompositions, in particular to observe the (average) numbers of factors that are needed to achieve a high approximation degree (for instance, 80 % and higher). The experiments show that with the first few factors computed, either of the algorithms presented in Section 3.3 (usually) achieve relatively high degrees of approximation. This observation is based on experiments with approximate factorisations of randomly generated tricontexts with various densities.

Figure 3.4 depicts relative frequency histograms of degrees of approximation obtained by using the first four factors in a sample of $850,000$ randomly generated 7-factorisable three-dimensional matrices of various sizes, each matrix containing about one million entries. The top-left histogram shows the degree of approximation obtained by using the first factor. The top-right histogram shows the degree of approximation obtained by using the first two factors. Analogously, the bottom-left and bottom-right histograms refer to the first three and first four factors. The hatched areas of the histograms are delimited by the intervals of the mean degrees of approximation ± the standard deviations. The mean values are also presented in the diagrams. As one can see, the first two factors cover nearly 70 % of the input data (on average), and the first four factors cover nearly 90 % of the input data. Thus, even if the generated matrices are 7-factorisable, i.e., 7 factors are needed to achieve 100 % degree of factorisation (the exact factorisation), only the first four factors are sufficient to achieve 90 % degree of approximation which can be quite surprising. Hence, the approximate decomposition can help to reveal important factors covering most part of the data. In this particular case of 7-factorisable matrices, we can say that (typically) the first four factors are the most important ones.

Concluding, while exact factorisation may require a large number of factors, a considerably smaller number of factors may cover a large portion of the data.

Positive Discrepancies

If we are looking for an approximate factorisation with *positive discrepancies*, i.e.,

$$Y \subseteq \bigcup_{f \in F} f^{I_1} \times f^{I_2} \times f^{I_3},$$

we have to compute dense rectangular boxes (dense boxes) instead of full rectangular boxes. *Dense boxes* correspond to triples (A_1, A_2, A_3) with $A_i \subseteq K_i$ $(i = 1, 2, 3)$ allowing blanks (0 entries). A dense box (A_1, A_2, A_3) has the interpretation "most objects from A_1 have most attributes from A_2 under most conditions from A_3". In many cases the factorisation through dense boxes is justified by the errors and noisiness contained in the data.

For the dyadic case the computation of dense rectangles of a given formal context was studied in [GV93, BV06b], however the factorisation problem was not considered.

As in [GV93], for a dense box $d = (A_1, A_2, A_3)$, we denote the number of zeros contained in $A_1 \times A_2 \times A_3$ by z_d and its density by $\rho_d := z_d/(|A_1| \cdot |A_2| \cdot |A_3|)$. Note that triconcepts have density 0.

Definition 3.20. The factorisation of a tricontext (K_1, K_2, K_3, Y) with dense boxes is a set of dense boxes $\mathcal{D} = \{(A_1, B_1, C_1), \ldots, (A_n, B_n, C_n)\}$ such that

$$Y \subseteq \bigcup_{(A,B,C) \in \mathcal{D}} A \times B \times C.$$

The most straightforward approach to obtain factorisations with positive discrepancies, i.e., with dense boxes, is to consider formal concepts of d-cuts and compute the corresponding triconcepts with a given density α.

For example, in the tricontext from Figure 1.3 (page 20) let us consider the concept $(A, B) := (\{0, 1, 2, 3, 4, 5\}, \{2, 3\})$ of the d-cut c_2^1. Now $(A \times B)^{(3)} = \{1, 2\}$, but $A \times B$ is also partially contained in the d-cut of condition 0. If we take the dense box $d_1 := (\{0, 1, 2, 3, 4, 5\}, \{2, 3\}, \{0, 1, 2\})$, then we get $z_{d_1} = 3$ and $\rho_{d_1} = 1/12$. By using this dense box we eliminate a factor from the optimal triconceptual factorisation (see Figure 3.1 and 3.2), namely $(\{0, 1, 2, 5\}, 2, \{0, 1, 2\})$, because the incidence is already covered by the dense box. Proceeding alike, we obtain the factorisation by dense boxes $\mathcal{D} = \{d_1, d_2, d_3\}$, where $d_2 := (\{2, 3, 4, 5\}, \{0, 1, 2, 3, 4, 5\}, \{0, 1, 2\})$ with $\rho_{d_2} = 1/8$ and $d_3 := (\{1, 3, 4, 5\}, \{0, 1, 2, 3, 5\}, \{0, 1, 2\})$ with $\rho_{d_3} = 1/10$. Of course we have to specify at the beginning which is the maximal allowed density. Through this method we obtained a factorisation containing 3 dense boxes instead of 8 triconcepts.

The lower the density is that we allow for the dense boxes, the more precise the factorisation becomes.

In the algorithms presented in Section 3.3 we may replace the conditions for triconcepts by conditions for α-dense boxes. In [GV93] more approaches for the computation of dense blocks (dense dyadic concepts) were presented. We are particularly interested in the one which allows blocks with a given density α. The algorithm can be easily generalised to the triadic case. The difference to the approach presented before is that any two components of a dense box do not necessarily form a dyadic concept.

3.5. Factor Analytical Tools

In this section we generalise some tools from ordinary Factor Analysis to our framework. It turns out that such methods can be easily applied in our setting.

3.5.1. Confirmatory Factor Analysis

So far we have presented Exploratory Factor Analysis, however Factor Analysis can also be used in a confirmatory way. In this approach the researcher has to make some assumptions about the number and/or the structure of the factors. Such hypothesis must be done under consideration of the attributes, how they interact with another, and which are likely to be contained in the same factor. If the hypothesis is appropriate for the data, then the validity of the factor analytical model tends to be more accepted.

When the researcher has to make assumptions just on the number of factors, there is no difference between Confirmatory and Exploratory Factor Analysis as the structure of factors is not influenced. If the researcher makes assumptions on the structure of factors, he/she has to specify the attributes belonging to each factor. Such hypothesised factorisation could yield an exact factorisation, however due to possible errors in the data or wrong assumptions of the user it is very likely that it would be an approximate factorisation. Further it could happen that such factorisations are not necessarily optimal. After applying Confirmatory Factor Analysis one could check whether the entire incidence relation of the context was covered and perform Exploratory Factor Analysis on it. The comparison between the results of the confirmatory and exploratory methods proved to be useful, see for instance [Lon83].

Algorithm 1 performs Confirmatory Factor Analysis on a triadic data set. The user must enter the tricontext $\mathbb{K} = (K_1, K_2, K_3, Y)$ and two factorisation contexts, namely (K_2, F, I_2) and (K_3, F, I_3), of the hypothetical triconceptual factorisation. The algorithm first computes the components of the triconcepts for each $f \in F$, lines $2-5$, and removes the rectangular boxes, not necessarily maximal, generated by them from U in line 6. The factorisation context (K_1, F, I_1) is built in line 7. In order to know which part of the incidence relation remains uncovered, we store the negative discrepancies induced by the hypothetical factorisation in line 9.

Algorithm 1: ConfirmatoryFactorAnalysis($\mathbb{K}, (K_2, F, I_2), (K_3, F, I_3)$)

1 **set** U **to** Y;
2 **foreach** $f \in F$ **do**
3   **set** A_2 **to** f^{I_2};
4   **set** A_3 **to** f^{I_3};
5   **set** A_1 **to** $(A_2 \times A_3)^{(1)}$;
6   **set** U **to** $U \setminus (A_1 \times A_2 \times A_3)$;
7   **set** f^{I_1} **to** A_1;
8 **end**
9 **set** $negDis$ **to** $\{(i,j,k) \mid (i,j,k) \in U\}$;
10 **return** (K_1, F, I_1), $negDis$

The algorithm can be modified in line 5 in order to accept dense boxes with some density instead of triconcepts or rectangular boxes.

Another possible way to use Confirmatory Factor Analysis is to check whether the factorisation of a data set is applicable on another data set. Suppose we have performed a study on more groups of people and we have computed an optimal factorisation of the first data set. The question arises whether this factorisation also applies to the persons from another study group or/and whether it applies to all persons participating in the study.

Let us now put the confirmatory triconceptual factorisation to work.

Example 3.21. We want to find out whether the triconceptual factorisation of the hostels from Seville works as well if we apply it to hostels from other cities. Recall that the hostels from Figure 1.3 (page 20) were chosen such that they were the best rated hostels in Seville that were present in all three hostel booking websites. Their triconceptual factorisation is given in Figure 3.1 and 3.2 (page 73). Let us now choose hostels from Malaga under the exact same conditions, and let us extract the data analogously. We obtain what is displayed in Figure 3.5. Again, we made a cross in the corresponding cell of the hostel, service and hostel booking website if the service of the hostel was rated as excellent by the users of the website.

	0: hostelworld						1: hostels						2: hostelbookers					
	0: character	1: safety	2: location	3: staff	4: fun	5: cleanliness	0: character	1: safety	2: location	3: staff	4: fun	5: cleanliness	0: character	1: safety	2: location	3: staff	4: fun	5: cleanliness
0: Oasis B. Malaga			×	×		×		×	×	×		×		×	×	×		×
1: Picasso's Corner			×	×					×						×	×		
2: Melting Pot	×		×	×	×	×	×		×	×	×	×	×	×	×	×	×	×
3: Residencia Univ.			×						×					×	×	×		×
4: Pink House			×	×				×	×	×		×			×	×		

Figure 3.5.: Hostels from Malaga

We will use Algorithm 1 with the tricontext $\mathbb{K}$ from Figure 3.5 and the attribute and condition factorisation contexts from the Seville example given in Figure 3.1 and 3.2. Whenever we write $(-)^{I_2}$ and $(-)^{I_3}$ we refer to the factorisation contexts for the hostels from Seville. In turn, we use the derivation operators $(-)^{(1)}$ and $(-)^{I_1}$ to refer to the tricontext about the hostels from Malaga and its factorisation context for objects, respectively. Recall that factor f_1 stands for best location since all the users agreed on this matter. For f_1 we have $(f_1^{I_2} \times f_1^{I_3})^{(1)} = K_1$, so this factor is also present in a (not necessarily optimal) factorisation of the hostels from Malaga. Even more, $(f_1^{I_1}, f_1^{I_2}, f_1^{I_3})$ is a triconcept of $\mathbb{K}$. Further, we have $f_3^{I_1} = \{0, 2\}$, $f_5^{I_1} = \{0, 2, 4\}$ and $f_6^{I_1} = \{2\}$. All these factors can be found in the new data set but not all of them induce triconcepts of $\mathbb{K}$. The other factors cannot be found in the data since we have $f_2^{I_1} = f_4^{I_1} = f_7^{I_1} = f_8^{I_1} = \varnothing$. Thus, our assumption that both data sets have the same underlying structure of factors turned out to be wrong. However, there is a significant overlap between the two structures. Indeed, half of the factors from the triconceptual factorisation for the hostels from Seville apply to the hostels from Malaga. Using the factors that are present in the data, $\{f_1, f_3, f_5, f_6\}$, we are able to cover $67,39\%$ of the incidence relation, where the larger crosses in Figure 3.5 are those that have not been covered by the confirmatory triconceptual factorisation.

	f_1	f_2	f_3	f_4	f_5
Oasis Malaga	×		×	×	×
Picasso's Corner	×				
Melting Pot	×	×	×		×
Residencia Univ.	×				×
Pink House	×		×	×	

	f_1	f_2	f_3	f_4	f_5
character		×			
safety				×	×
location	×	×	×	×	×
staff		×	×	×	×
fun		×			
cleanliness		×		×	×

Figure 3.6.: Factorisation contexts for objects and attributes

	f_1	f_2	f_3	f_4	f_5
hostelworld	×	×	×		
hostels	×	×		×	
hostelbookers	×	×	×		×

Figure 3.7.: Factorisation context for conditions

Note however that factor f_1 and f_5 from the Seville hostels appear also in the optimal factorisation of the hostels from Malaga. The entire conceptual factorisation is displayed in Figure 3.6 and Figure 3.7. For this data set we have only five factors. This is due to the fact that the data is sparser and that the users from the three hostel booking website tend to agree more in their opinions.

3.5.2. Projection of External Elements

Suppose that we have already computed the triconceptual factorisation of a tricontext $\mathbb{K} = (K_1, K_2, K_3, Y)$ but want to introduce new objects, attributes or conditions without recomputing the factorisation. For these new elements we want to check how good they fit the factor model. We will study the case for new objects. The projection of new attributes or conditions is done analogously. For the dyadic case such projections were studied among others in [DRV93].

Let us denote the factor set of $\mathbb{K}$ by F and the new object by g. This new object induces a tricontext (g, K_2, K_3, J). How good g fits a factor $f \in F$ can be determined by the following formula:

$$\text{proj}_f(g) := \frac{|\{g\}^{(1)} \cap (f^{I_2} \times f^{I_3})|}{|f^{I_2} \times f^{I_3}|}.$$

Obviously, $0 \leq \text{proj}_f(g) \leq 1$ for all $f \in F$. The higher the value of $\text{proj}_f(g)$ is, the better factor f explains the object g. If $\text{proj}_f(g) = 1$, we say "object g can be fully projected on factor f" and we can extend the extent f^{I_1} by g. Now, if

$$g^{(1)} \subseteq \bigcup \{ f^{I_2} \times f^{I_3} \mid f \in F, \mathrm{proj}_f(g) = 1 \}, \tag{3.17}$$

then the whole incidence relation of (g, K_2, K_3, J) is explained by some factors of F, i.e., object g fits the factor model and we can extend the first factorisation context. If the above equation is not fulfilled, then g cannot be explained by F and we would need additional factors for an exact factorisation of $(K_1 \cup \{g\}, K_2, K_3, Y \cup J)$.

Example 3.22. Consider once again our running example from Figure 1.3 (page 20) with its triconceptual factorisation given in Figure 3.1 and 3.2 (page 73). We want to add

	0						1						2					
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
6: Triana B.	×	×		×		×	×	×	×	×		×	×	×	×	×		×
7: Living Roof	×	×		×			×	×	×	×			×	×	×	×		

Figure 3.8.: Projection of new objects

new hostels to the tricontext in Figure 1.3 which are displayed in Figure 3.8. We have $\mathrm{proj}_{f_5}(6) = \mathrm{proj}_{f_6}(6) = \mathrm{proj}_{f_7}(6) = 1$ and $6^{(1)} \subseteq \bigcup \{ f^{I_2} \times f^{I_3} \mid f \in \{f_5, f_6, f_8\}\}$ satisfying Equation 3.17. Thus, the hostel Triana Backpackers fits the triconceptual factorisation. On the other hand, the hostel Living Roof cannot be explained by the factorisation. For this hostel, the factor f_5 is the only one on which the hostel can be fully projected. However this factor does not cover the entire incidence relation of $(\{7\}, K_2, K_3, J)$. Hence, this hostel does not satisfy Equation 3.17 and it is therefore not explained fully by the factorisation.

3.6. Conclusion

We proposed a new approach to Factor Analysis of three-way binary data. The approach utilises triconcepts as factors. Such a choice is justified by a theorem showing that triadic concepts provide us with optimal factorisations of triadic binary data. In addition, as illustrated by an example, triadic concepts are easily interpretable, resulting in a simply to understand output of the Factor Analysis. Further, we gave an upper bound for the number of factors, described those triconcepts which are present in any factorisation of a given tricontext, and characterised the factorisation contexts. We also introduced mappings which transform a description of a given object in terms of attributes and conditions into a description of the same object in terms of factors. We proposed a greedy algorithm for computing suboptimal triconceptual factorisations and provided its experimental evaluation. Moreover, we studied the triconceptual factorisations with positive and negative discrepancies. In order to obtain a better fitting of our approach to Factor Analysis we also generalised some factor analytical tools to our framework.

Due to the wide applicability of Factor Analysis in a triadic setting it arose as a natural wish to generalise also the conceptual factorisation to the triadic setting. This generalisation is also of major importance to show that the conceptual factorisation might be as powerful as Factor Analysis.

4
Hierarchical Classes Analysis

Hierarchical Classes Analysis was developed by De Boeck and Rosenberg at the end of the 80's ([DR88]). It is a discrete, categorical data analysis model for binary data. Its development was driven by applications in personality organisation and implicit belief systems. Let us mention beforehand that there is a great similarity between the conceptual factorisation and Hierarchical Classes Analysis.

Although the development of Hierarchical Classes Analysis started almost thirty years ago, there is still recent work going on in this field, especially concerning generalisations of the model to non-binary and non-dyadic data. We will also study these variants.

This chapter does not aim to develop some complex mathematical theory. As we will see this is also not necessary. Our main goal is to connect Hierarchical Classes Analysis with the conceptual factorisations and to show how they can benefit from each other. We will establish this connection by translating the notions of Hierarchical Classes Analysis into the language of Formal Concept Analysis and conceptual factorisation. Our aim is to present the whole picture. Therefore, we will investigate the motivations, theoretical frameworks and algorithmic approaches of the methods.

The results regarding the connection between conceptual factorisation and Hierarchical Classes Analysis for binary data presented in Section 4.1 are partially based on the work developed in [Glo11b]. Those findings motivated us to investigate the link between the two methods also for non-dyadic and numerical data. A brief survey of this field was given in [Glo11a].

Before we start our work let us present the body of this chapter. In Section 4.1 we give a brief introduction to Hierarchical Classes Analysis for binary dyadic data and study its connection to conceptual factorisation. As there exists a tight link between the two methods, we investigate the similarities of them also for other kinds of data. In Section 4.2 we study the triadic versions of Hierarchical Classes Analysis and establish a link between these methods and triconceptual factorisation. Section 4.3 deals with Hierarchical Classes Analysis for numerical data. Completely new land is reached in Section 4.4, where we develop the theory for Hierarchical Classes Analysis in a fuzzy setting. An overall conclusion of this chapter is given in Section 4.5.

4.1. Hiclas

We start our investigation by giving a brief introduction to Hierarchical Classes Analysis and translate the notions into the language of Formal Concept Analysis. The interested reader may find a more detailed introduction to Hierarchical Classes Analysis in [DR88, DRV93].

In Hierarchical Classes Analysis a $p \times q$ binary matrix W is decomposed into the Boolean matrix product $X \circ Y^T$ of a $p \times n$ binary matrix X and a $q \times n$ binary matrix Y with n being the Schein rank of W, i.e., the smallest possible value such that $W = X \circ Y^T$ where Y^T is the transpose matrix of Y. The matrices X and Y have as columns the characteristic vectors of the object and attribute bundles (see below), respectively.

In the following we just present the definitions and notions for objects, the ones for attributes can be done analogously by interchanging objects with attributes.

The *object by attribute data* from Hierarchical Classes Analysis is a binary matrix which corresponds to the Boolean matrix representation of a formal context. In Hierarchical Classes Analysis two objects are called *equivalent* if and only if the same attributes apply to them. An *object class* is the set of all objects that are equivalent to one another in a given object by attribute data. If the objects $g_1, \dots, g_n$ form an object class, we write $[g_1, \dots, g_n]$. The class of objects to which none attributes apply is called the *undefined class*.

Each object class is characterised by the set of attributes that apply to all objects of the class. Therefore, the object classes can be ordered by the super-/subset relation of their attribute sets. This order is a partial one, called the *hierarchy of object classes*.

We may translate the above presented notions from Hierarchical Classes Analysis into the language of Formal Concept Analysis as follows:

Definition 4.1. In a formal context (G, M, I) two objects $g_1, g_2 \in G$ are called **equivalent** if and only if $g_1' = g_2'$. The set

$$[g_1] := \{g \in G \mid g' = g_1'\}$$

is called the **object class** of the object g_1. We define a partial order relation between the classes of G, the so-called **hierarchy of object classes**, as follows:

$$[g_1] \leq [g_2] :\Longleftrightarrow g_1' \subseteq g_2',$$

for any objects $g_1, g_2 \in G$.

Example 4.2. Consider the context given in Figure 4.1 with its object set $G = \{1, \dots, 10\}$ and attribute set $M = \{a, \dots, h\}$. Objects 1 and 2 are equivalent and therefore form an object class. There are all in all six object classes, namely $[1,2], [3,4], [5], [6,7]$, $[8]$ and $[9,10]$, where the last one is an undefined class, because no attribute applies to its objects. The attribute classes are $[a,b], [c,d], [e,f]$ and $[g,h]$. This example does not contain an undefined attribute class.

	a	b	c	d	e	f	g	h
1	×	×	×	×	×	×	×	×
2	×	×	×	×	×	×	×	×
3	×	×	×	×			×	×
4	×	×	×	×			×	×
5			×	×	×	×	×	×
6	×	×					×	×
7	×	×					×	×
8					×	×	×	×
9								
10								

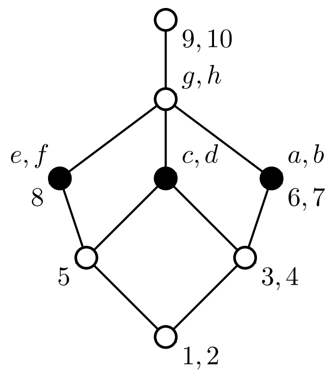

Figure 4.1.: A formal context and its concept lattice. The concepts corresponding to the black coloured circles are used in the conceptual factorisation.

The *association relation* between the hierarchy of object and of attribute classes allows a simultaneous graphical representation of the two hierarchies. An object class is *associated* with all attribute classes that apply to the objects of that object class. By the association of an object class with an attribute class, the first is also associated with all the superordinate classes of the second one and vice versa. Therefore, it is sufficient to associate the bottom classes of the two hierarchies. Graphically, one hierarchy is represented upside-down and the association relation by zigzag lines. The graphical representation of the context from Figure 4.1 is displayed in Figure 4.2 and the concept lattice in Figure 4.1, where the black coloured circles correspond to concepts used in the conceptual factorisation.

The object (attribute) classes are ordered through the super-/subset relation of their attribute (object) sets. However, the undefined object (attribute) class is not included into the hierarchy of objects (attributes), because it interferes with the graphical representation. Including the undefined classes into the hierarchies would make them the bottom classes and, through the association relation of the bottom classes, every object class would be associated with every attribute class.

The graphical representation of the Hierarchical Classes Analysis model contains all the object and all the attribute concepts. There are, however, cases where these concepts are not enough for the correct representation of the association relation or/and for the optimal factorisation. Therefore, the diagram may contain also other concepts, whereas, the concept lattice contains all the concepts of a given context.

The object set which corresponds to an attribute class can be decomposed into object classes such that their size is maximal and their number minimal. These objects are then called *object bundles*. An object bundle is a set of objects which is associated with the same bottom class of attributes. The three relations of the model can be reconstructed from the bundles.

In the language of Formal Concept Analysis we give the following definition of bundles:

Definition 4.3. An **object (attribute) bundle** is the extent (intent) of some formal concept.

Example 4.4. The graphical representation sometimes requires empty classes for the correct illustration of the association relation. For instance, in Figure 4.2 the hierarchy of object classes contains an empty bottom class, because the classes $[3,4]$ and $[5]$ both apply to the attribute class $[c,d]$. A direct zigzag line from $[3,4]$ and $[5]$ to $[c,d]$ is not permitted because the two object classes are not bottom classes.

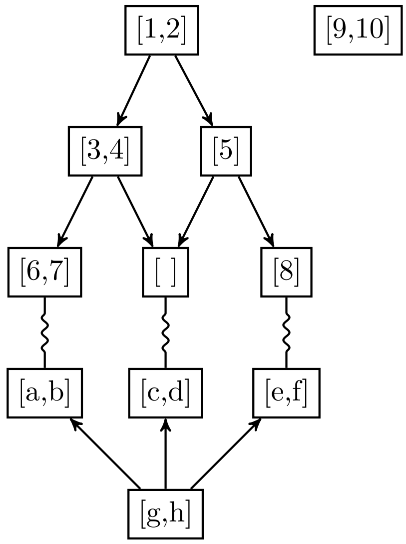

Figure 4.2.: Hiclas representation

The object bundles are $\{1,2,3,4,6,7\}$, $\{1,2,3,4,5\}$ and $\{1,2,5,8\}$ and the attribute bundles are $\{a,b,g,h\}$,$\{c,d,g,h\}$ and $\{e,f,g,h\}$. The bundle decomposition is given in Figure 4.3.

Note that, for the optimal factorisation, the Hierarchical Classes Analysis model yields the same set of formal concepts as the conceptual factorisation. This is not surprising, because we have already seen that both methods actually use formal concepts for the factorisation/decomposition. Let us fix this result in the following proposition:

Proposition 4.5. *For an optimal decomposition there is a one-to-one correspondence between the factorisation contexts of a conceptual factorisation and the object-attribute bundles from Hierarchical Classes Analysis.*

However, when it comes to algorithms, the situation slightly changes. This is caused by the fact that the factorisation problem is NP-hard and the two approaches tackle the problem differently.

The Hiclas algorithm was presented in [DR88, LV01] and it computes for a binary matrix W the best fitting Hierarchical Classes Analysis model for a given solution rank n. The algorithm assumes that $W = Z+E$, where W, Z and E are $p\times q$ matrices, $Z = X \circ Y^{\mathcal{T}}$ with X and Y being $p\times n$ and $q\times n$ binary matrices, respectively, and E is the discrepancies matrix. Then, X and Y are estimated by a least square approach. The user must specify the number of bundles, the rank, of the solution. The algorithm starts an iterative procedure based on either an initial set of attribute or object bundles. The initial bundle set can be determined by: 1) a rational heuristic; 2) a random generation procedure; or 3) the user's input; where the first two are built into the algorithm. Hiclas can also be used in a confirmatory way through method 3). The algorithm stops either when the pre-entered

	Object Bundles				Attribute Bundles		
Obj.	$OB1$	$OB2$	$OB3$	Attr.	$AB1$	$AB2$	$AB3$
1	1	1	1	a	1	0	0
2	1	1	1	b	1	0	0
3	1	1	0	c	0	1	0
4	1	1	0	d	0	1	0
5	0	1	1	e	0	0	1
6	1	0	0	f	0	0	1
7	1	0	0	g	1	1	1
8	0	0	1	h	1	1	1
9	0	0	0				
10	0	0	0				

Figure 4.3.: Bundle decomposition of the context from Figure 4.1

rank is reached or when the number of discrepancies does not decrease in any further iteration. The *optimal number of bundles* is considered to be the number beyond which the discrepancies decrease slightly.

We have already seen the theoretical framework, algorithmic approach of Hierarchical Classes Analysis and its connection to conceptual factorisation. As its development was motivated by applications, we are interested in the analyses where Hierarchical Classes Analysis was used. Indeed, the method proved itself successful in many applications. For instance in sociometrics ([DR88]), in the analysis of data involving patients with medically unexplained somatic symptoms ([GSE+98]), in studies in social personality ([Ros88]), in splitted personality analysis ([RG]), and so on. Based on our results we may conclude that conceptual factorisation can be also applied for such studies.

Comparison of the Models

As we have seen, Hierarchical Classes Analysis as well as conceptual factorisation seek to decompose a formal context/binary matrix into two factorisation contexts/binary matrices such that the number of factors is as small as possible. Both methods actually use formal concepts for this purpose, i.e., there is a one-to-one correspondence between concepts and object-attribute bundles. The major differences lie in the graphical representation and the algorithms. Concerning the latter, the algorithms of both methods search for the smallest possible subset of formal concepts which covers the incidence relation of a context, but their approaches are different. Due to the fact that the factorisation problem is NP-hard, a greedy approximation algorithm was considered in [BV10a]. This algorithm is efficient but can possibly yield suboptimal solutions, as discussed in Section 2.1. Commonly, data used in Factor Analysis is very large and this algorithm is applicable. A brute-force set covering algorithm is applicable on moderate size data. Further, we may use an approximate factorisation with positive or/and negative discrepancies, see Section 2.1. On the other hand, the Hiclas algorithm is based on a branch-and-bound approach and always yields an optimal factorisation. However, a branch-and-bound approach can be applied only on smaller data sets due to its high complexity. Therefore, Hiclas was implemented to compute factorisations with up to 15 bundles.

Note that Formal Concept Analysis and Hierarchical Classes Analysis were first linked in [CY05]. However, back then the factorisation problem was not considered. Hence, this work does not represent a competitor for our approach presented here.

Hierarchical Classes Analysis was developed to find the latent variables hidden in data about patients suffering from different disorders. We have already mentioned above that this technique was successfully applied in various clinical evaluations. The 2-dimension originates from a completely different research field. However, the two problems meet through the conceptual factorisation. For the 2-dimension the factorisation is a significant application. But the other side also profits from this connection. Indeed, the conceptual factorisation/Hierarchical Classes Analysis is based on a solid mathematical framework.

Hierarchical Classes Analysis was generalised in [LVD01] to the treatment of ordinal data. The roots of this generalisation lie in the Nonmetric Factor Analysis developed by Coombs and Kao [CK55] that we have already mentioned in Chapter 2. The three standard relations of Hierarchical Classes Analysis are adapted to the new data type. The main concern of [LVD01] lies in finding an appropriate algorithm for the optimisation of the nonmetric models. The outcome uses a least square and least absolute deviation function. However, since there are no connections between [LVD01] and our ordinal factorisation, we will stop the investigation of that work here.

4.2. Triadic Hiclas

Based on the results from the previous section, we also expect similarities between the conceptual factorisation and Hierarchical Classes Analysis in the triadic setting. As we will see in the following, this is indeed the case. There are two different approaches to Hierarchical Classes Analysis in the triadic setting, the Indclas model and the Tucker3 model, which we will study in the upcoming subsections.

Both methods have a common core. They are applicable on three dimensional binary data matrices. The three dimensions, the so-called *modes*, are considered to be *objects, attributes* and *sources*. In Triadic Concept Analysis we identify this data matrix with a tricontext by replacing the zeros with blanks and the ones with crosses. Further, the three modes correspond to objects, attributes and conditions, respectively.

4.2.1. Indclas

Indclas, a triadic version of Hierarchical Classes Analysis, was presented in [LVDR99]. It decomposes a binary $p \times q \times r$ $3d$-matrix W into $p \times n, q \times n$ and $r \times n$ binary matrices P, Q and R with n being as small as possible. Then, the matrices P, Q and R are called *object bundle, attribute bundle* and *source bundle* matrices, respectively, and n is said to be the *rank of the model*. The decomposition is done through the Boolean $3d$-matrix product given by

$$W_{ijk} = \bigvee_{l=1}^{n} P_{il} \cdot Q_{jl} \cdot R_{kl},$$

for all $i \in \{1, \dots, p\}, j \in \{1, \dots, q\}$ and $k \in \{1, \dots, r\}$. This matrix product is identical with the one used in the triconceptual factorisation, see Definition 3.13 (page 75).

For an element x of either of the three modes, $W(x)$ denotes the set of tuples of elements from the other two modes that are associated with x in W. This operator corresponds to the $(-)^{(i)}$-derivation operators in Triadic Concept Analysis, see Definition 1.28 (page 19).

The *association relation* is a ternary relation between the three modes given by the Boolean $3d$-matrix product.

Two elements x and y of some mode are *equivalent* if $W(x) = W(y)$. Equivalent objects (attributes, sources) constitute an *object* (*attribute, source*) *class*. Consequently, equivalent objects (attributes, sources) have identical bundle patterns in the Indclas model. The *object* and *attribute bundles* are defined as in the dyadic case and the *source bundles* analogously.

An element x is *hierarchically below* an element y if $W(x) \subseteq W(y)$. This order is a partial one and forms the *hierarchical order* in each mode.

To sum up, the notions presented above translate into the language of Triadic Concept Analysis in the following way:

Definition 4.6. Let (K_1, K_2, K_3, Y) be a tricontext. Two elements $x, y \in K_i$ with $i \in \{1, 2, 3\}$ are **equivalent** if and only if $x^{(i)} = y^{(i)}$. For an object (attribute, condition) $x \in K_i$ with $i \in \{1, 2, 3\}$ its **object** (**attribute, condition**) **class** is defined by

$$[x] := \{y \in K_i \mid y^{(i)} = x^{(i)}\}.$$

An **object** (**attribute, condition**) **bundle** is the extent (intent, modus) of some triconcept.

Consequently, the i-th rows of the three bundle matrices correspond to the components of some triconcept. Similarly to the dyadic case we have the following result:

Proposition 4.7. *For an optimal factorisation there is a one-to-one correspondence between the factorisation contexts of a triconceptual factorisation and the three bundles of Indclas.*

Example 4.8. Consider the tricontext (K_1, K_2, K_3, Y) from Figure 4.4 with its object set $K_1 = \{1, \dots 7\}$, attribute set $K_2 = \{a, \dots e\}$ and condition set $K_3 = \{A, B, C\}$. For object

	A					B					C				
	a	b	c	d	e	a	b	c	d	e	a	b	c	d	e
1	×			×	×		×	×	×		×	×	×	×	×
2	×			×	×		×	×	×		×	×	×	×	×
3						×			×	×	×			×	×
4		×	×	×								×	×	×	
5	×			×	×	×	×	×	×	×	×	×	×	×	×
6	×	×	×	×	×		×	×	×		×	×	×	×	×
7						×			×	×	×			×	×

Figure 4.4.: Example of a tricontext

7 we obtain $W(7) = \{(a, B), (d, B), (e, B), (a, C), (d, C), (e, C)\}$ and hence $W(7) = 7^{(1)}$.

	Object Bundles					Attribute Bundles					Source Bundles			
Obj.	$OB1$	$OB2$	$OB3$	$OB4$	Attr.	$AB1$	$AB2$	$AB3$	$AB4$	Sour.	$SB1$	$SB2$	$SB3$	$SB4$
1	0	1	1	0	a	0	0	1	1	A	1	0	1	0
2	0	1	1	0	b	1	1	0	0	B	0	1	0	1
3	0	0	0	1	c	1	1	0	0	C	1	1	1	1
4	1	0	0	0	d	1	1	1	1					
5	0	1	1	1	e	0	0	1	1					
6	1	1	1	0										
7	0	0	0	1										

Figure 4.5.: Indclas bundle decomposition of the tricontext from Figure 4.4

The objects 1 and 2 share the same attributes under each mode (source) and thus, are equivalent. Consequently, they form an object class.

The decomposition of the tricontext is given by the object, attribute and source bundles in Figure 4.5. These bundles correspond to the following four triconcepts:

$$
\begin{aligned}
&(\{4,6\},\{b,c,d\},\{A,C\}),\\
&(\{1,2,5,6\},\{b,c,d\},\{B,C\}),\\
&(\{1,2,5,6\},\{a,d,e\},\{A,C\}),\\
&(\{3,5,7\},\{a,d,e\},\{B,C\}),
\end{aligned}
$$

which are exactly the triconcepts used in the triconceptual factorisation. The circles corresponding to these triconcepts are drawn larger in the trilattice from Figure 4.7.

Now let us turn our attention to the graphical representation. The Indclas graphical representation of the tricontext from Figure 4.4 is given in Figure 4.6. In the upper part of the diagram the hierarchical order of object classes is displayed, in the lower part the upside-down hierarchy of attribute classes, and in the middle the circles containing the corresponding sources. The zigzag lines represent the association relation. Compared to the graphical representation of the dyadic case, the zigzag lines in the triadic setting include circles which contain the sources that belong to the corresponding bundles. We can read the association relation from the diagram as follows: Object x is associated with attribute y by source z if and only if object x is connected with attribute y by a path that goes through source z. For instance, in Figure 4.6 a path goes from the object class $[4]$ to the attribute classes $[b,c]$ and $[d]$ through the circle containing the sources A and C.

However, through this graphical representation the hierarchical order of the sources is lost. One has to draw a separate diagram for the hierarchy of the source classes. To also include the hierarchical order of the source classes into the diagram, the trilattice representation is adequate. The trilattice of the tricontext from Figure 4.4 is displayed in Figure 4.7 (see the text next to Figure 1.4 on page 23 for details on how to read such diagrams). The triconcepts belonging to the triconceptual factorisation are drawn larger in the trilattice.

The Indclas algorithm is an alternating elementary binary discrete least squares procedure ([LVDR99]), which starts from an initial configuration of bundles. The algorithm is a generalisation of the one from the dyadic case.

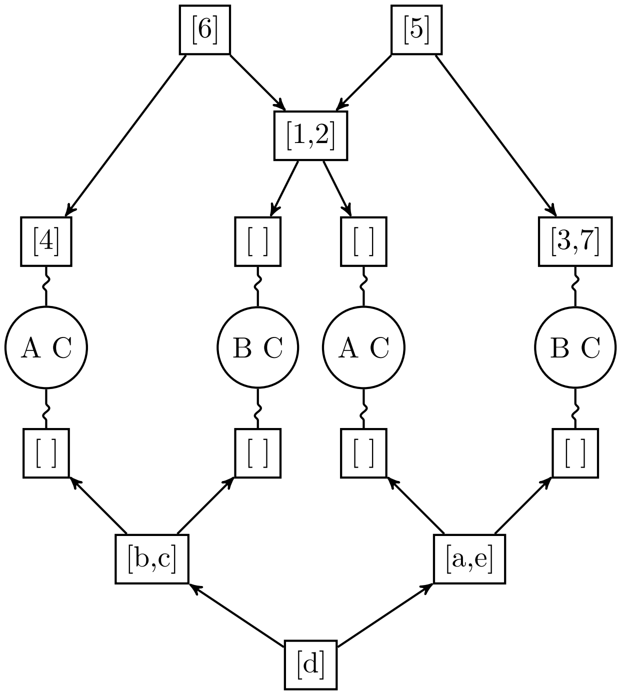

Figure 4.6.: Indclas representation of Figure 4.5

The triadic version of Hierarchical Classes Analysis was successfully applied in different analyses of patients suffering from mental disorders. For instance in [LVDR99] thirty case descriptions of psychiatric inpatients (objects) were evaluated based on twenty-three symptoms (attributes) by fifteen clinicians (sources/conditions). Once again we may conclude that triconceptual factorisation would be applicable in the analysis of such clinical data.

4.2.2. Tucker3 Hierarchical Classes Analysis

The Tucker3 Hierarchical Classes Analysis was developed in [CVL03] as a generalisation of Indclas. It has the same underlying structure as Indclas but the matrices P, Q and R are constructed differently. The Tucker3 model does not restrict these matrices to have the same rank, i.e., the number of bundles in a decomposition may differ from mode to mode. Additionally, the Tucker3 model allows a more complex linking structures among the hierarchies.

The Tucker3 model implies the decomposition of a $p \times q \times r$ $3d$-binary matrix W into a $p \times n_1$ binary *object bundle matrix* P, a $q \times n_2$ binary *attribute bundle matrix* Q, a $r \times n_3$ binary *source bundle matrix* R and an $n_1 \times n_2 \times n_3$ binary $3d$-matrix T, where (n_1, n_2, n_3) is the *rank of the model.* Then, T is a ternary association relation between the three bundle matrices and is called the *core array.*

The *equivalence* and *hierarchical* relations of the Tucker3 model are defined in a completely identical manner as in the Indclas model. However, the *association relation* of the

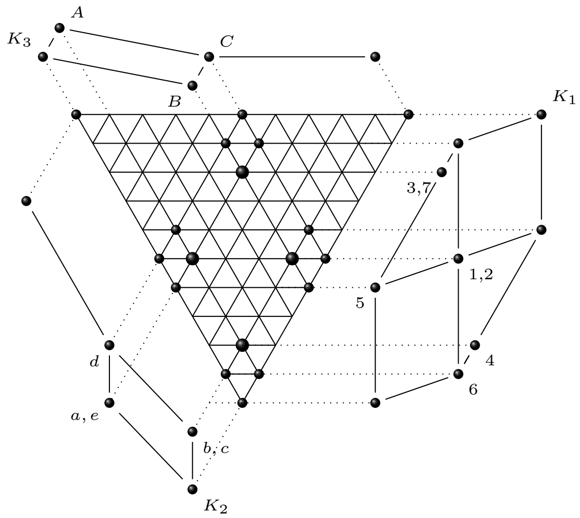

Figure 4.7.: Trilattice of the tricontext from Figure 4.4. The larger circles correspond to the triconcepts used in the triconceptual factorisation.

Obj.	Obj. Bundles $O1$	$O2$	$O3$	Attr.	Attr. Bundles $A1$	$A2$	Sour.	Source Bundles $S1$	$S2$	Obj. Bund.	Attr. Bund.	Source Bundles $S1$	$S2$
1	0	1	0	a	1	0	A	0	1	$O1$	$A1$	1	0
2	0	1	0	b	0	1	B	1	0	$O1$	$A2$	0	0
3	1	0	0	c	0	1	C	1	1	$O2$	$A1$	0	1
4	0	0	1	d	1	1				$O2$	$A2$	1	0
5	1	1	0	e	1	0				$O3$	$A1$	0	0
6	0	1	1							$O3$	$A2$	0	1
7	1	0	0										

Figure 4.8.: Tucker3 bundle decomposition of the tricontext from Figure 4.4

Tucker3 model is defined differently, namely by

$$W_{ijk} = \bigvee_{i_1=1}^{n_1} \bigvee_{i_2=1}^{n_2} \bigvee_{i_3=1}^{n_3} P_{ii_1} \cdot Q_{ji_2} \cdot R_{ki_3} \cdot T_{i_1 i_2 i_3}$$

for all $i \in \{1, \ldots, p\}, j \in \{1, \ldots, q\}$ and $k \in \{1, \ldots, r\}$. Then, $W_{ijk} = 1$ if and only if there exist object, attribute and source bundles to which object i, attribute j and source k, respectively, belong to and which are associated in the core array T. In Figure 4.8 the Tucker3 bundle decomposition of the tricontext from Figure 4.4 is given. From it we can read for example that object 3 is associated with attribute a through source B, because the bundles $O1$, $A1$ and $S1$ to which the elements belong to, respectively, are associated in the core array T.

The graphical representation of the Tucker3 model is similar to the Indclas model. However, the Tucker3 model provides a different linking between the three modes in the diagram due to its association relation. In the graphical representation of the Tucker3 model a path is drawn between the bottom classes of the object and attribute hierarchies if they are associated in the core array T. The path contains a circle with the sources of the source bundle for which the association holds. Every association relation contained in the core array is allowed to be represented.

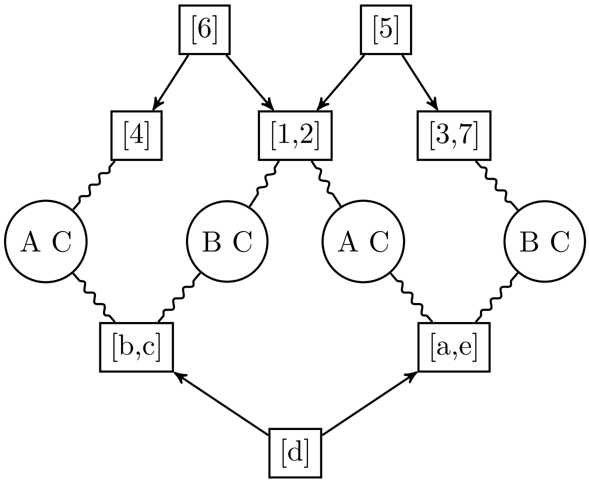

Figure 4.9.: Tucker3 representation

In the Indclas model the i-th rows of P, Q and R represented the extent, intent and modus, respectively, of some triconcept. The matrices P, Q and R in the Tucker3 model still have as rows extents, intents and modi, respectively, however the i-th rows of them do not necessarily belong to the same triconcept. As a triconcept is uniquely determined by two of its components, the same extent, intent or modus can appear in different triconcepts. The Tucker3 model takes advantage of these equivalence relations ($\sim_i$ for $i \in \{1, 2, 3\}$, see (1.14) on page 21) and contains in its bundles just one representative for each equivalence class. The core array on the other hand connects the three bundles with each other, i.e., it "puts" the components of the triconcepts together.

It is arguable whether the Tucker3 model is better than the Indclas one. The price one has to pay for a possibly more compact representation of the bundles is quite high, because there is a supplementary matrix needed to link the bundles with each other.

4.2.3. Comparison of the Models

As we have seen in the previous two subsections, all approaches to Triadic Hierarchical Classes Analysis use formal triconcepts for the bundle decomposition. Once again, the mathematical formalisations are different. The Indclas model decomposes a $3d$-matrix into object, attribute and source bundles, which basically correspond to the factorisation contexts of the triconceptual factorisation. The Tucker3 Hiclas uses the equivalences on the components of the triconcepts, but the factorisation is also done with triconcepts. The Tucker3 Hiclas presents the three bundle matrices in a possibly more compact way, however it requires a fourth matrix for the association of the bundle matrices. Thus, with the knowledge gained from Triadic Concept Analysis we may clarify the difference between the Indclas and the Tucker3 models.

The factorisation problem is reducible to the set covering problem and hence is NP-hard. Therefore, greedy approximation algorithms were considered in [BGV12, BV10b, Glo10] for the triconceptual factorisation. Once again we may use approximate factorisations with positive and negative discrepancies, see Section 3.4.

In general, the Indclas algorithm performs better than the greedy approximation algorithms since it is a branch-and-bound algorithm. However it is not applicable to large data sets due to complexity issues.

Indclas and Triadic Concept Analysis were first linked in [Hwa07], however, once again the factorisation problem was not considered.

Concluding, there is once more a one-to-one correspondence between triconceptual factorisation and the triadic versions of Hierarchical Classes Analysis. Further, we have seen how the methods may benefit from each other, i.e., structural explanations, graphical representations, algorithms and applications.

4.3. Disjunctive Hiclas-R and RV-Hiclas

The disjunctive Hiclas-R model was presented in [VLC07]. It decomposes of a $p \times q$ matrix W with integer entries from $V = \{1, \ldots, v\}$ into a binary $p \times n_1$ *object bundle matrix* P, a binary $q \times n_2$ *attribute bundles matrix* Q and an $n_1 \times n_2$ *core matrix* T which takes n_3 different non-zero values, where $n_3 \leq v$. The *rank* of the model is (n_1, n_2, n_3).

Similarly to the Tucker3 model, the Hiclas-R model also allows that the numbers of object and attribute bundles are different from each other.

The *equivalence* and *hierarchical relations* are defined analogously to the binary Hiclas model, but in this case we have integer values instead of binary values as matrix entries. Two objects x, y are *equivalent* if they have identical rows in the data table W, i.e., $W(x) = W(y)$. The equivalence of attributes is defined on the columns, i.e., two attributes are equivalent if they have identical columns in the data table. An object x is *hierarchically below* an object y, written $x \leq y$, if $W(x) \leq W(y)$ with respect to the component-wise order.

The *association relation* is given by

$$W_{ij} = \bigvee_{h=1}^{n_1} \bigvee_{k=1}^{n_2} P_{ih} \cdot Q_{jk} \cdot T_{hk},$$

for all $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, q\}$. Object i is associated with attribute j at the maximum value of association indicated by the core matrix T for the pair of bundles which contain object i and attribute j.

The RV-Hiclas model is a generalisation of the Hiclas-R model, the main difference between the two are the matrix entries. The RV-Hiclas model contains real-valued data. Therefore, in the following we work with the RV-Hiclas model.

Example 4.9. Consider the data table with real values given in Figure 4.10. Objects b and c have identical rows, thus they are equivalent. We have, for instance, $W(c) \leq W(d)$ with respect to the component-wise order, and hence object c is hierarchically below object d. Figure 4.11 contains the RV-Hiclas bundle decomposition of the data from Figure 4.10. According to it, object d is associated with attribute β in value 1, because the object bundle $OB3$ (to which object d belongs) and the attribute bundle $AB3$ (to which attribute

	α	β	γ
a	0.25	0	0.25
b	0	0.5	0.25
c	0	0.5	0.25
d	0	1	0.25

Figure 4.10.: Data table with real values

β belongs) are associated in the core matrix with value 1. Note that object d belongs to the object bundle $OB2$ as well. However, the value to which $OB2$ and $AB3$ are associated in the core matrix is 0.5 and this does not represent the maximum of association between d and β.

	Obj. Bundles				Attr. Bundles				Core matrix T		
Obj.	$OB1$	$OB2$	$OB3$	Attr.	$AB1$	$AB2$	$AB3$		$AB1$	$AB2$	$AB3$
a	1	0	0	α	1	0	0	$OB1$	0.25	0	0
b	0	1	0	β	0	0	1	$OB2$	0	0.25	0.5
c	0	1	0	γ	1	1	0	$OB3$	0	0	1
d	0	1	1								

Figure 4.11.: RV-Hiclas bundle decomposition of the data from Figure 4.10

The graphical representation of the RV-Hiclas model is similar to the Tucker3 model but the circles contain the maximum values of the association instead of the sources. The graphical representation of the data from Figure 4.10 is displayed in Figure 4.12. From the diagram we can also read the association relation. For example, a path goes from d to β which includes a circle containing the value 1.

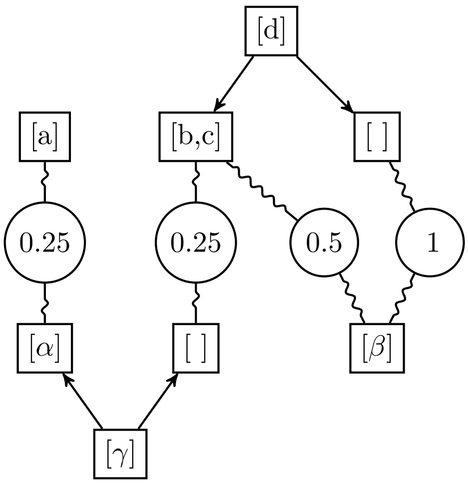

Figure 4.12.: RV-Hiclas representation of Figure 4.11

The Hiclas-R and RV-Hiclas models can also be formulated in the language of Formal Concept Analysis. We can scale the integer-valued data (many-valued context) into a dyadic formal context and perform a conceptual factorisation. For the data from Figure 4.10 an ordinal or nominal scaling could be applied. In the previous models there was a one-to-one correspondence between the object (attribute) bundles and extents (intents). However, this is not the case for the Hiclas-R and RV-Hiclas models. If we consider the corresponding nominal and ordinal scaled contexts, then for example the object bundle $\{b, c, d\}$ is an extent of the ordinal scaled context and $\{d\}$ is an extent of the nominally scaled context. Hence, there is no connection between the methods regarding numerical data.

4.4. Fuzzy Hierarchical Classes Analysis

Up until now we have taken notions from Hierarchical Classes Analysis and translated them into the language of Formal Concept Analysis. In this section we will do the opposite step, namely use the conceptual factorisation from the fuzzy setting and reinterpret them, mutatis mutandis, from the point of view of Hierarchical Classes Analysis. The so obtained result will then be the fuzzy approach to Hierarchical Classes Analysis. Why do we do that if we already know how to factorise contexts with fuzzy attributes? Well, first of all we would like to spread these results to other communities as well, and second we would like to have an even tighter connection between the two techniques. Further, to the best of our knowledge, this is the first attempt of bringing Hierarchical Classes Analysis and fuzzy theory together.

First of all, let us make some remarks about concepts in a fuzzy setting. These will motivate our approach to Fuzzy Hierarchical Classes Analysis.

Remark 4.10. In the crisp case each maximal rectangle full of crosses corresponds uniquely to a formal concept. The only possible exceptions are the concepts (G, G') and (M', M) if and only if $G' = M' = \varnothing$. However, this remark does not hold for the fuzzy setting. First of all, different **L**-concepts of an **L**-context may yield the same maximal rectangle. Further the rectangles do not have to be maximal with respect to their values. For an **L**-concept (A, B) its corresponding maximal rectangle $A \bigotimes B$ is an **L**-set given by (2.6) (page 65). It has the value $A(g) \otimes B(m)$ for the tuple (g, m). Now, if $(A_1, B_1), (A_2, B_2)$ are different **L**-concepts, we may have $A_1 \bigotimes B_1 = A_2 \bigotimes B_2$ or $A_1 \bigotimes B_1 \leq A_2 \bigotimes B_2$ with respect to the point-wise order or, of course, none of the mentioned situations. Where does this non-uniqueness of maximal rectangles come from? This can be traced back to the properties of the different fuzzy logics. Suppose we have a residuated lattice **L** with L being its support set and $a, b \in L$. Applying the Gödel logic, we assume without loss of generality that $a \otimes b = \min(a, b) = a$. However, there can be some $c \in L$ such that $c \neq b$ and $a \leq c$. Thus, $\min(a, b) = \min(a, c) = a$. Through the Łukasiewicz logic we obtain $a \otimes b = \max(a + b - 1, 0)$. It is possible to have $c, d \in L$ such that $a \otimes b = c \otimes b$. Finally, for the Goguen t-norm we obtain similar remarks. Indeed, for different tuples $(a, b), (c, d) \in L^2$, $a \otimes b = c \otimes d$ may hold.

We already know from Section 2.4 that **L**-concepts provide us with optimal factorisations of **L**-contexts and that the factorising families correspond to the families of preconcepts covering the **L**-relation of the context. It will turn out that **L**-preconcepts are more suitable than **L**-concepts in the setting of Fuzzy Hierarchical Classes Analysis.

Definition 4.11. Let (G, M, I) be an $\mathbf{L}$-context and L the support set of the residuated lattice. Two objects $g_1, g_2 \in G$ are **equivalent for values** $l_1, l_2 \in L$ if $\{^{l_1}/g_1\}^{\uparrow} = \{^{l_2}/g_2\}^{\uparrow}$. Equivalent objects form an **object class**. For two objects $g_1, g_2 \in G$ we call g_1 **hierarchically below** g_2 for values l_1, l_2, written $\{^{l_1}/g_1\} \leq \{^{l_2}/g_2\}$, if $\{^{l_1}/g_1\}^{\uparrow} \subseteq \{^{l_2}/g_2\}^{\uparrow}$.

Note that an object can be hierarchically below itself for different values. For instance, $g_1 \in G$ may yield $\{^{l_1}/g_1\} \leq \{^{l_2}/g_1\}$ for some $l_1, l_2 \in L$. Obviously, if two objects $g_1, g_2 \in G$ have identical rows in the context, then they are equivalent for any value $l \in L$, i.e., $\{^{l}/g_1\}^{\uparrow} = \{^{l}/g_2\}^{\uparrow}$ for all $l \in L$.

Similarly to the other models of Hierarchical Classes Analysis, we build object and attribute bundle matrices which are defined as follows:

Definition 4.12. Let (G, M, I) be an $\mathbf{L}$-context with $|G| = n$ and $|M| = m$. Further let $\mathcal{F} := \{(A_1, B_1), \ldots, (A_k, B_k)\}$ be a subset of $\mathbf{L}$-preconcepts of (G, M, I) covering I, i.e.,

$$I = \bigcup_{l=1}^{k} A_l \bigotimes B_l.$$

The **object bundle matrix** P has as rows the characteristic vectors of the extents from $\mathcal{F}$ and the **attribute bundle matrix** Q has as rows the characteristic vectors of the intents from $\mathcal{F}$. An object bundle is **associated with** an attribute bundle if and only if they form an $\mathbf{L}$-preconcept. The **fuzzy matrix product** is given by

$$(P \circ Q)_{ij} := \bigvee_{l=1}^{n} P_{il} \otimes Q_{lj}$$

for $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$.

In other words, we compute the t-norm multiplication between each element of the l-th column of P with each element of the l-th row of Q, where l runs over $\{1, \ldots, n\}$, and take the maximum of these products.

Example 4.13. Instead of interpreting the values in the data table from Figure 4.10 as real, we will consider them as fuzzy ones. We automatically see that objects b and c are equivalent and attribute α is hierarchically below attribute γ. The optimal factorisation

Obj.	OB1	OB2	Attr.	AB1	AB2
a	0.25	0	α	0.25	0
b	0	0.5	β	0	1
c	0	0.5	γ	0.25	0.25
d	0	1			

Obj.	OB1	OB2	OB3	Attr.	AB1	AB2	AB3
a	1	0	0	α	0.25	0	0
b	0.5	0.5	1	β	0	1	0.5
c	0.5	0.5	1	γ	0.25	0.25	0.25
d	0	1	1				

Figure 4.13.: Optimal bundles with the Gödel (left) and the Łukasiewicz logic (right)

of it with the Gödel logic contains two $\mathbf{L}$-preconcepts and with Łukasiewicz logic three $\mathbf{L}$-preconcepts. They are displayed in Figure 4.13. The graphical representations are given in Figure 4.14. For comparison, the optimal $\mathbf{L}$-conceptual factorisations are given by the following $\mathbf{L}$-concepts: With the Gödel logic we have

$$(\{a\}, \{^{0.25}/\alpha, {}^{0.25}/\gamma\}), (\{^{0.5}/b^{0.5}/c, d\}, \{\beta, {}^{0.25}/\gamma\}),$$

and with the Łukasiewicz logic we have

$$\begin{aligned}
&(\{a, {}^{0.75}/b, {}^{0.75}/c, {}^{0.75}/d, \}, \{{}^{0.25}/\alpha, {}^{0.25}/\gamma\}),\\
&(\{{}^{0.5}/b, {}^{0.5}/c, d\}, \{\beta, {}^{0.25}/\gamma\}),\\
&(\{{}^{0.5}/a, b, c, d\}, \{{}^{0.5}/\beta, {}^{0.25}/\gamma\}).
\end{aligned}$$

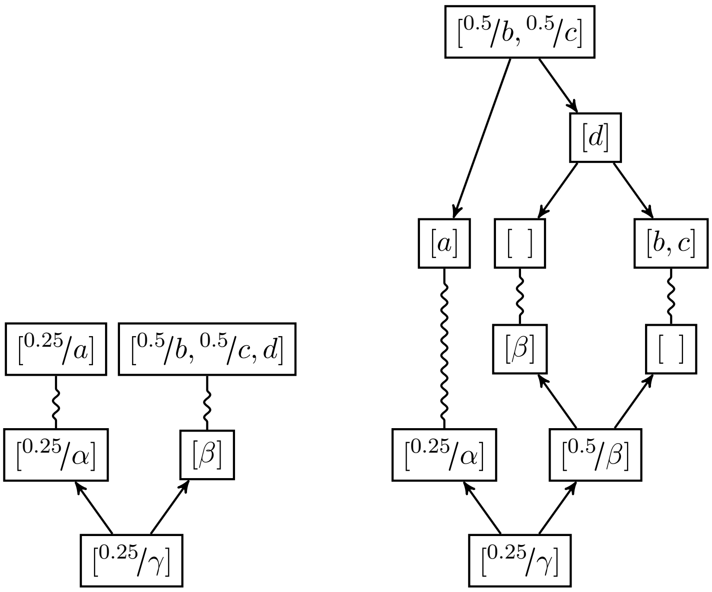

Figure 4.14.: Graphical representation of the bundles from Figure 4.13

The only thing we are still missing is the graphical representation of the Hierarchical Classes Analysis model in the fuzzy setting. The graphical representation of the bundle decomposition from Figure 4.13 is displayed in Figure 4.14. Once again we will draw the object classes and connect them according to the hierarchical ordering. Note that we use a crisp-like subsethood relation and not a fuzzy one, see Definition 4.11. Afterwards, we do the same for the attribute bundles, but we draw them upside-down and below the object hierarchy. Finally, we draw the zigzag lines between the two hierarchies as given by the association relation. In analogy to other Hierarchical Classes models there are cases when empty classes are needed for the correct representation of the association relation. The graphical representation is similar to the RV-Hiclas model but instead of using circles with association values, the object and the attribute classes contain their fuzzy membership values. Further, we associate only the components of the same **L**-preconcept with another. For comparison, the **L**-concept lattices are displayed in Figure 4.16 and 4.17 for the Gödel and Łukasiewicz logic, respectively.

Using the fuzzy Hiclas graphical representation for the optimal conceptual factorisation with the Łukasiewicz logic, one obtains the diagram from Figure 4.15. As it can easily be seen, changing the preconcepts into concepts overcomplicates the diagram. Hence, preconcepts are sufficient for an optimal factorisation and also yield easier to read diagrams.

Comparing our fuzzy approach to Hierarchical Classes Analysis and the RV-Hiclas model we see that the first is a more parsimonious method. First of all, because it does not require

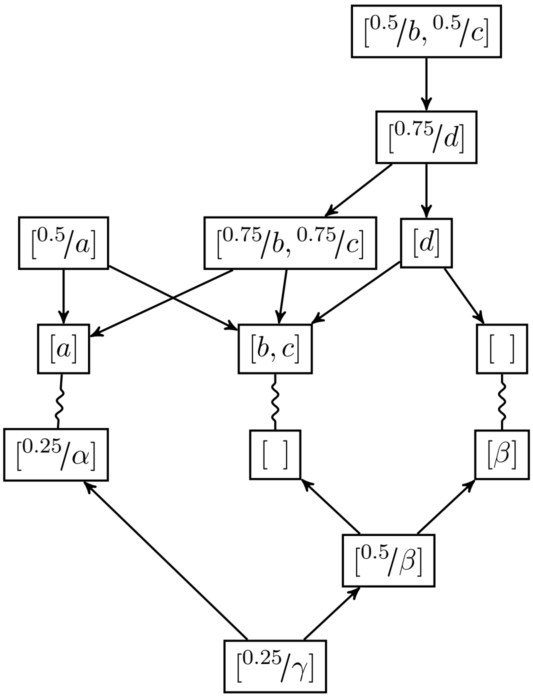

Figure 4.15.: Fuzzy Hiclas graphical representation of the fuzzy conceptual factorisation with the Łukasiewicz logic

a third matrix, namely the core matrix. However, there is an association just between extents and intents of the same concept, whereas the bundle decomposition provides a linking between different object and attribute bundles. Second, the fuzzy approach yields in general a smaller number of factors than the bundle decomposition, due to the properties of the t-norms. A drawback in the fuzzy approach is that it is applicable just for values of the unit interval. Therefore, a data matrix containing different values has to be transformed in such a form. However, the results can then be transformed back to their initial value domains.

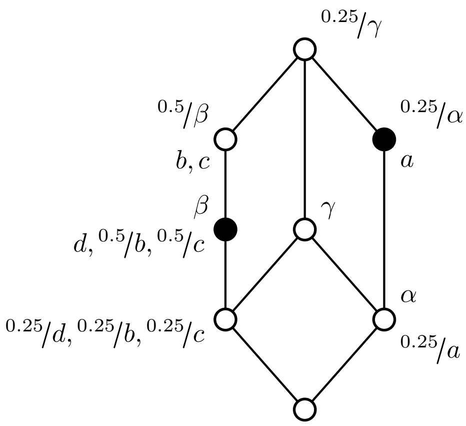

Figure 4.16.: **L**-concept lattice of the **L**-context from Figure 4.10 with the Gödel logic. The black circles correspond to the **L**-concepts used in the **L**-conceptual factorisation.

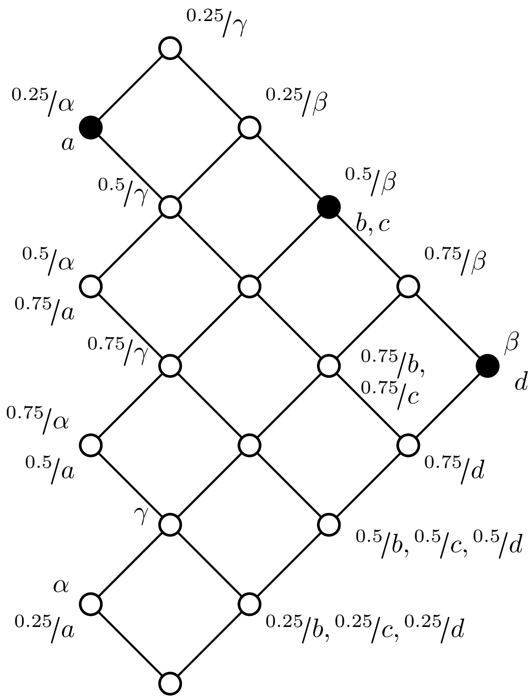

Figure 4.17.: **L**-concept lattice of the **L**-context from Figure 4.10 with the Łukasiewicz logic. The black circles correspond to the **L**-concepts used in the **L**-conceptual factorisation.

4.5. Conclusion

We presented a comparison between conceptual factorisation and Hierarchical Classes Analysis regarding discrete, triadic and real data. It turned out that both methods use formal concepts as factors in the factorisation of binary and triadic data. For the case concerning real data there is however no one-to-one correspondence between **L**-concepts and the bundles.

Further, we presented a generalisations of Hierarchical Classes Analysis to the fuzzy setting. Such a generalisation is useful for patient data, because sometimes it is difficult to decide to which extent a patient suffers from a symptom. We have translated the usual notions from Hierarchical Classes Analysis into the fuzzy setting. Thereby we used **L**-preconcepts for the bundle matrices since they yield a smoother graphical representation than **L**-concepts.

Hence, the formal concept analytical approach to Factor Analysis and the application driven data reduction necessity meet in the common point of Boolean and triadic Boolean factorisations. We showed how the two methods may benefit from each other. On the one hand we have seen in what kind of analyses conceptual factorisations may be used. On the other hand we have presented the results from Hierarchical Classes Analyses in a new light. Further, we have more options for the graphical representation and the algorithmic approach.

Part II.

Fuzzy Data

Introduction to Part II

Fuzzy data has found numerous applications in both theoretical and real-world research fields. We have already pointed out some of them in Section 1.2. However, the list is much longer, starting with approximate reasoning, pattern recognition including clustering and image processing, fuzzy systems like neural networks, automate and dynamic systems, and by far not ending with data bases and decision making. An extensive overview of these and further applications can be found in [KY95].

In the forthcoming chapters we aim to enlarge the list by adding further methods and frameworks for using fuzzy data within Formal Concept Analysis.

In Chapter 5 we generalise a very powerful tool, attribute exploration, into the fuzzy setting. Attribute exploration is used for knowledge discovery and found numerous applications in both theoretical and practical research fields. Thus, bringing fuzzy data and attribute exploration together, the outcome can have nothing but numerous applications!

The method presented in Chapter 6 allows the users to model their preferences over the attribute set of a formal fuzzy context. Based on these preferences the users only obtain the formal fuzzy concepts that are relevant to them. Trivially, the method belongs to the family of data reduction tools, however it can easily be performed and understood by the users as the only input they have to provide is a sort of order on the attributes.

In Chapter 7 we combine fuzzy data with a notion we already got acquainted with in the previous part, namely "triadic data" and come to a framework which we call *Fuzzy-valued Triadic Concept Analysis*. We first present what builds the fundamentals of Formal Concept Analysis in this setting. Afterwards, we study implications in such data sets. There are mainly two kinds, but the symmetry of triadic contexts allows us to investigate nine different implication families. In the last section we bring yet another notion familiar from the previous part into our new setting by developing the fuzzy-valued triconceptual factorisations.

5

Attribute Exploration in a Fuzzy Setting

Attribute exploration is a very powerful tool. We have already seen the result that builds its theoretical basis, namely Proposition 1.25 (page 18). This is what represents its key to success. Thus, the crucial step is to generalise this proposition to the fuzzy setting. We have seen in Subsection 1.5.2 that the choice of the hedge in the residuated lattice influences the existence, uniqueness and minimality of the stem base. Therefore, we will split the development of attribute exploration in a fuzzy setting into two cases, namely the one where we use the globalisation as the hedge, Section 5.1, and the one where we use general hedges, Section 5.2. In Section 5.3 we add background knowledge to the exploration process. That is, we add implications between attributes about which we know in advance that they hold. This has the advantage of shortening the exploration process since less questions have to be answered and fewer counterexamples have to be provided.

The first two sections are based on [Glo12b], whereas the results from Section 5.3 are presented here for the first time.

5.1. Exploration with Globalisation

We start by developing the theoretical ingredients for a successful attribute exploration in a fuzzy setting with the globalisation. Afterwards, we turn our attention to its practical parts. First, we develop an appropriate algorithm for this technique, and thereafter illustrate the method by an example.

Recall that we obtain a unique and minimal set of pseudo-intents for a given **L**-context whenever we use the globalisation as the hedge, see Theorem 1.45 (page 30).

Let us turn our attention to the fuzzy version of Proposition 1.25. Although its proof would be straightforward by the isomorphism between the concepts lattices of an **L**-context (G, M, I) with globalisation and its double-scaled context $(G, M \times L, I^{\Box})$, see Subsection 1.5.2, we give here the proof based on fuzzy logic.

Proposition 5.1. *Let* $\mathbf{L}$ *be a finite residuated lattice with globalisation. Further, let* $\mathcal{P}$ *be the unique system of pseudo-intents of a finite* $\mathbf{L}$*-context* $\mathbb{K}$ *with* $P_1, \dots, P_n \in \mathcal{P}$ *being the first n pseudo-intents in* $\mathcal{P}$ *with respect to the lectic order. If* $\mathbb{K}$ *is extended by an object* g*, the object intent* $g^{\uparrow}$ *of which respects the implications* $P_i \Rightarrow P_i^{\downarrow\uparrow}$ *for all* $i \in \{1, \dots, n\}$*, then* $P_1, \dots, P_n$ *remain the lectically first n pseudo-intents of the extended context.*

Proof. Let $\mathbb{K} = (H, M, J)$ be the initial context and let (G, M, I) be the extended context, namely $G = H \cup \{g\}$ and $J = I \cap (H \times M)$.

First we need to clarify the notion "$g^{\uparrow}$ respects all implications". It means that $\{{}^{l}/g\}^{\uparrow}$ is a model of all implications for any truth value $l \in L$. The hedge $(-)^{*}$ we are using is the globalisation, and therefore $\{{}^{l}/g\}^{\uparrow} = \{{}^{0}/g\}^{\uparrow} = M$ respects all implications for any $l \in L \setminus \{1\}$. Hence, it suffices to check that $\{{}^{1}/g\}^{\uparrow}$ is a model of all implications.

Now we are prepared for the proof which will be done by induction over the number of pseudo-intents.

Let $n = 1$ and $P \in \mathcal{P}$ be the lectically smallest pseudo-intent of (H, M, J). We extend (H, M, J) by an object g, the object intent g^{I} of which is a model of $P \Rightarrow P^{JJ}$. Thus, $P^{JJ} = P^{II}$. Now suppose that there exists a pseudo-intent Q of (G, M, I) that is lectically smaller than P. Since $Q^{J} \subseteq Q^{I}$, we have $Q^{II} \subseteq Q^{JI} = Q^{JJ}$. Further, since Q is a pseudo-intent of (G, M, I) it follows that $Q \neq Q^{JJ}$. As P is the lectically smallest pseudo-intent of (H, M, J), there is no pseudo-intent R of (H, M, J) with $R \subseteq Q < P$, where $<$ is meant with respect to the lectic order. By the definition of pseudo-intents we obtain that Q is a pseudo-intent of (H, M, J), contradicting that P is the smallest pseudo-intent of (H, M, J). Hence, there does not exist a pseudo-intent Q of (G, M, I) such that Q is lectically smaller than P. Particularly, we have $Q \subseteq P$, implying that P is indeed a pseudo-intent of (G, M, I), namely the smallest one.

Now we show the induction step from $n - 1$ to n. To this end let $P_1, \dots, P_{n-1}$ be the lectically first $(n - 1)$ pseudo-intents of (G, M, I) and let g^{I} be a model of $P_i \Rightarrow P_i^{JJ}$ for every $i \in \{1, \dots, n\}$. We have to show that P_n is the lectically next pseudo-intent of (G, M, I) after P_{n-1}. For contradiction, suppose it is not, i.e., assume that there exists a pseudo-intent Q of (G, M, I) such that $P_{n-1} < Q < P_n$. Obviously, from $P_i \subseteq Q$ we obtain $P_i^{JJ} = P_i^{II} = Q$ for all $i \in \{1, \dots, n-1\}$. Further, from $Q^{JJ} \subseteq Q^{II} \neq Q$ we obtain that Q is a pseudo-intent of (H, M, J). Thus, since $P_{n-1} < Q$ it follows that $Q = P_n$, contradicting $Q < P_n$. Therefore, there is no pseudo-intent Q of (G, M, I) such that $P_{n-1} < Q < P_n$. Since $P_1, \dots, P_{n-1}$ are lectically smaller pseudo-intents of (G, M, I) than P_n, we have shown that P_n is a pseudo-intent of (G, M, I), namely the lectically next one. □

We have now the key to a successful attribute exploration in the fuzzy setting, at least when we use the globalisation. With this result we are able to generalise the attribute exploration algorithm to the fuzzy setting as presented by Algorithm 2. Its input is the $\mathbf{L}$-context $\mathbb{K}$ and the residuated lattice $\mathbf{L}$. The first intent or pseudo-intent is the empty set. If it is an intent, add it to the set of intents of the context. Otherwise, ask the expert whether the implication is true in general. If so, add this implication to the stem base, otherwise ask for a counterexample and add it to the context (line 2 – 11). Until A is different from the whole attribute set, repeat the following steps: Search for the largest attribute i in M with its largest value l such that $A(i) < l$. For this attribute compute its closure with respect to the cl_{T}-closure operator given by (1.24) (page 30) and check

Algorithm 2: FuzzyExploration($\mathbb{K}, \mathbf{L}$)

```
ℒ := ∅; A := ∅;
if A = A^{↓↑} then
    add A to Int(𝕂)
else
    Ask expert whether A ⇒ A^{↓↑} is valid;
    if yes then
        add A ⇒ A^{↓↑} to ℒ
    else
        Ask for counterexample g and add it to 𝕂
    end
end
while A ≠ M do
    for i = n, ..., 1 and l = max L, ..., min L with A(i) < l do
        B := cl_T(A);
        if A ↘ i = B ↘ i and A(i) < B(i) then
            A := B;
            if A = A^{↓↑} then
                add A to Int(𝕂)
            else
                Ask expert whether A ⇒ A^{↓↑} is valid;
                if yes then
                    add A ⇒ A^{↓↑} to ℒ
                else
                    Ask for counterexample g and add it to 𝕂
                end
            end
        end
    end
end
```

whether the result is the lectically next intent or pseudo-intent (line 12 – 16). Thereby, $A \searrow i := A \cap \{1, \dots, i-1\}$. If the result is an intent, add it to the set of intents (line 17–18), otherwise ask the user whether the implication provided by the pseudo-intent holds. If the implication holds, add it to the stem base otherwise ask the user for a counterexample (line 20 – 24).

The algorithm generates interactively the stem base of the **L**-context. As in the crisp case we enumerate the intents and pseudo-intents in the lectic order. Due to Proposition 5.1 we are allowed to extend the context by objects whose object intents respect the already confirmed implications. This way, the pseudo-intents already contained in the stem base do not change. Hence, the algorithm is sound and correct.

Example 5.2. We want to explore the size and distance of the planets. We include some of them into the object set and obtain the context displayed in Figure 5.1. In this example we use the Łukasiewicz logic with the globalisation as hedge.

	small (s)	large (l)	far (f)	near (n)
Earth	1	0	0	1
Mars	1	0	0.5	1
Pluto	1	0	1	0

Figure 5.1.: Initial context

We start the attribute exploration. The first pseudo-intent is $\{\varnothing\}$ and we are asked

Do all objects have the attribute small to degree 1*?*

This is of course not true and we provide a counterexample:

	small (s)	large (l)	far (f)	near (n)
Jupiter	0	1	1	0.5

The next pseudo-intent is $\{n\}$ and we are asked

Do objects having attribute near to degree 1 *also have attribute small to degree* 1*?*

This is true and we confirm the implication. The next pseudo-intent is $\{f, {}^{0.5}/n\}$ which yields the following question:

Do objects having attributes far and near to degree 1 *and* 0.5*, respectively, also have attribute large to degree* 1*?*

We also confirm this implication since it is true. The algorithm proceeds with

Do objects having attribute large to degree 0.5 *also have the attributes large, far and near to degree* $1, 1$ *and* 0.5*, respectively?*

This implication is not true for our planet system and we give a counterexample:

	small (s)	large (l)	far (f)	near (n)
Uranus	0.5	0.5	1	0

The following four implications are true, so we will confirm them:

$$\begin{aligned} {}^{0.5}/l &\Rightarrow f, \\ l, f &\Rightarrow {}^{0.5}/n, \\ {}^{0.5}/s, {}^{0.5}/n &\Rightarrow s, n, \\ s, {}^{0.5}/l, f &\Rightarrow l, n. \end{aligned}$$

And the attribute exploration has stopped. Now we have an extended **L**-context, namely the one containing Jupiter and Uranus besides the objects given in Figure 5.1. Note that we did not have to include all the planets into the object set, just a representative part of them. The other planets with their attributes are displayed in Figure 5.2. These objects contain just redundant information and the knowledge provided by them is already incorporated into the stem base of the extended context.

	small (s)	large (l)	far (f)	near (n)
Mercury	1	0	0	1
Venus	1	0	0	1
Saturn	0	1	1	0.5
Neptune	0.5	0.5	1	0

Figure 5.2.: Superfluous planets

Hence, the result of the exploration is the set of implications of the stem base and an extended **L**-context containing the representative objects for the whole theory.

Concluding, we are able to perform a successful attribute exploration in a fuzzy setting provided we choose the globalisation as the hedge. In the forthcoming section we will study the same problem but use a general hedge in the residuated lattice.

5.2. Exploration with General Hedges

In this section it turns out that there are several obstacles that make a straightforward generalisation of attribute exploration to a fuzzy setting using general hedges impossible. At the end of the section we will discuss which approaches may lead to a successful exploration. However, it is also an open question whether an exploration in such a framework is desirable.

Recall that when using a general hedge in the residuated lattice, neither the existence nor the uniqueness of stem base is ensured. We have already presented these results from [BV05b] in Subsection 1.5.2 where we also explained how the computation of stem bases in this general framework can be done (see Theorem 1.48 and Example 1.49 on page 31).

Let us start with an example.

Example 5.3. Consider once again the **L**-context from Example 1.49 (page 32), namely $(\{g\}, \{a, b\}, I)$ with $I(g, a) = 0.5$ and $I(g, b) = 0$ and **L** being the three-element Łukasiewicz chain with the identity hedge. We have already computed its systems of pseudo-intents and the corresponding stem bases in Example 1.49 and display them once again here in Figure 5.3.

$\mathcal{T}_2$	$\mathcal{T}_4$
(1) $\ {}^{0.5}/b \Rightarrow a$	(3) $\ {}^{0.5}/a, {}^{0.5}/b \Rightarrow a$
(2) $\ a \Rightarrow {}^{0.5}/b$	(4) $\ a \Rightarrow {}^{0.5}/b$

Figure 5.3.: Stem bases

Now we could start an attribute exploration, for instance in $\mathcal{T}_4$. The algorithm would ask us:

Do objects having attribute a and b both to degree 0.5 *also have attribute a to degree* 1*?*

Let us answer this question affirmatively. The next question is:

Do objects having attribute a to degree 1 *also have attribute b to degree* 0.5*?*

We deny this implication and provide a counterexample, the object h with $I(h,a) = 1$ and $I(h,b) = 0$. This counterexample obviously respects the already confirmed implication so the context is extended by the new object h. For this extended context we want to find its stem bases and therefore compute the sets V and E as given by (1.25) and (1.26) (page 31). From the graph $\mathbf{G} = (V, E \cup E^{-1})$ we obtain four maximal independent sets, three of which form systems of pseudo-intents since they fulfil the condition from Theorem 1.48. The stem bases which they induce are displayed in Figure 5.4. At the beginning we have

$\mathcal{T}_4'$	$\mathcal{T}_4''$	$\mathcal{T}_4'''$
(5) $\ {}^{0.5}/b \Rightarrow a$	(6) $\ \{\} \Rightarrow {}^{0.5}/a$	(8) $\ b \Rightarrow a$
	(7) $\ {}^{0.5}/a, b \Rightarrow b$	

Figure 5.4.: Stem bases of the extended context

confirmed implication (3) from Figure 5.3. However, this implication is now not present in any stem base of the extended context.

By extending the context with objects that respect the already confirmed implications, the latter may disappear from the stem base of the extended context. Hence, we do not have an analogon of Proposition 1.25 for general hedges.

The attribute exploration with general hedges raises a lot of questions and open problems. We have more than one stem base for a context. These bases are equally powerful with respect to their expressiveness. The major problem however is how to perform an attribute exploration successfully. It is an open problem how to enumerate the pseudo-intents obtained by general hedges such that the already confirmed implications still remain in the stem base of the extended context. One could for instance make some constraints on the counterexamples. However, such an approach is not in the spirit of attribute exploration.

As we have already discussed in Subsection 1.5.2, the problem originates from the issues regarding the stem bases of **L**-contexts with general hedges. There, it is yet unclear which conditions have to be satisfied in order to have a system of pseudo-intents. Further open questions regard the uniqueness and minimality of such systems.

5.3. Exploration with Background Knowledge

The user may know some implications between attributes in advance. We will call such kind of implications *background implications*. In the rest of this section we will focus on finding a minimal list of implications, which together with the background implications will describe the structure of the concept lattice.

The theory about background knowledge for the crisp case was developed in [Stu96] and a more general form of it in [Gan99]. In the latter, the implications (Horn clauses) are replaced by so-called *cumulated clauses* which make the task of finding a minimal base cumbersome. The results presented in [Stu96] for implication bases with background knowledge follow by some slight modifications of the results about implication bases without background knowledge presented in [GW96]. The same applies for the fuzzy variant of this method. Hence, if we choose the empty set as the background knowledge, we obtain the results presented in [BV06a, BCV04].

Particularly appealing is the usage of background knowledge in the exploration process. This proves itself to be very useful and time saving for the user. He/she will have to answer less questions, as the algorithm does not start from scratch. Hence, the user will have to enter also less counterexamples.

We start by investigating the stem bases of **L**-contexts relative to a set of background implications. Afterwards we show how some notions and results for fuzzy implications and their stem bases change for our new setting. Note that we will present these results for general hedges. However, when it comes to the exploration we do not have any other choice than to consider the globalisation. In order to arrive at the exploration with background knowledge we will present the lectic order, the "key proposition" and an appropriate algorithm for attribute exploration in this setting. At the end we explore the characteristics of the planets using background knowledge and compare the results with the exploration without background knowledge.

Definition 5.4. Let $\mathbb{K}$ be a finite **L**-context and let $\mathcal{L}$ be a set of background implications. A set $\mathcal{B}$ of fuzzy implications of $\mathbb{K}$ is called **$\mathcal{L}$-complete** if every implication of $\mathbb{K}$ is entailed by $\mathcal{L} \cup \mathcal{B}$. We call $\mathcal{B}$, **$\mathcal{L}$-non-redundant** if no implication $A \Rightarrow B$ from $\mathcal{B}$ is entailed by $(\mathcal{B} \setminus \{A \Rightarrow B\}) \cup \mathcal{L}$. If $\mathcal{B}$ is both $\mathcal{L}$-complete and $\mathcal{L}$-non-redundant, it is called an **$\mathcal{L}$-base**.

Note that if we have $\mathcal{L} = \varnothing$ in the above definition, then all $\mathcal{L}$-notions are actually the notions introduced for sets of fuzzy implications. This remark holds also for the other notions introduced in this section.

Until explicitly said otherwise, the attribute sets of **L**-contexts and the residuated lattices **L** will be considered finite.

For any set $\mathcal{L}$ of background implications and any **L**-set $X \in \mathbf{L}^M$ we define an **L**-set

$X^{\mathcal{L}} \in \mathbf{L}^M$ by

$$X^{\mathcal{L}} := X \cup \bigcup \{B \otimes \mathrm{S}(A, X)^* \mid A \Rightarrow B \in \mathcal{L}\}. \tag{5.1}$$

Further, we define an $\mathbf{L}$-set $X^{\mathcal{L}_n} \in \mathbf{L}^M$ for each non-negative integer n as follows:

$$X^{\mathcal{L}_n} := \begin{cases} X, & n = 0, \\ (X^{\mathcal{L}_{n-1}})^{\mathcal{L}}, & n \geq 1. \end{cases} \tag{5.2}$$

An operator $\mathcal{L}$ on these sets is defined by

$$\mathcal{L}(X) := \bigcup_{n=0}^{\infty} X^{\mathcal{L}_n}. \tag{5.3}$$

From [BV06c] we know that the operator defined by (5.3) is an $\mathbf{L}^*$-closure operator for a finite set M of attributes and a finite residuated lattice $\mathbf{L}$.

Definition 5.5. For an $\mathbf{L}$-context (G, M, I), a subset $\mathcal{P} \subseteq \mathbf{L}^M$ is called a **system of $\mathcal{L}$-pseudo-intents** of (G, M, I) if for each $P \in \mathbf{L}^M$ the following holds

$$P \in \mathcal{P} \iff (P = \mathcal{L}(P) \neq P^{\downarrow\uparrow} \text{ and } ||Q \Rightarrow Q^{\downarrow\uparrow}||_P = 1 \text{ for each } Q \in \mathcal{P} \text{ with } Q \neq P).$$

As in the case without background knowledge (see (1.23), page 30), the usage of the globalisation simplifies the definition of the system of $\mathcal{L}$-pseudo-intents. We have that $\mathcal{P} \subseteq \mathbf{L}^M$ is a system of pseudo-intents if

$$P \in \mathcal{P} \iff (P = \mathcal{L}(P) \neq P^{\downarrow\uparrow} \text{ and } Q^{\downarrow\uparrow} \subseteq P \text{ for each } Q \in \mathcal{P} \text{ with } Q \subsetneq P).$$

Theorem 5.6. *The set of fuzzy implications*

$$\mathcal{B}_{\mathcal{L}} := \{P \Rightarrow P^{\downarrow\uparrow} \mid P \textit{ is an } \mathcal{L}\textit{-pseudo-intent}\} \tag{5.4}$$

*is an $\mathcal{L}$-base of $\mathbb{K}$. We call it the **$\mathcal{L}$-Duquenne-Guigues-base** or the **$\mathcal{L}$-stem base**.*

Proof. First note that all implications from $\mathcal{B}_{\mathcal{L}}$ are implications of (G, M, I). We start by showing that $\mathcal{B}_{\mathcal{L}}$ is complete, i.e., $||A \Rightarrow B||_{\mathcal{B}_{\mathcal{L}} \cup \mathcal{L}} = ||A \Rightarrow B||_{(G,M,I)}$ for every fuzzy implication $A \Rightarrow B$. We have $||A \Rightarrow B||_{(G,M,I)} = ||A \Rightarrow B||_{\mathrm{Int}(G^*,M,I)}$ by Lemma 1.40 (page 29). Hence, it suffices to prove

$$||A \Rightarrow B||_{\mathcal{B}_{\mathcal{L}} \cup \mathcal{L}} = ||A \Rightarrow B||_{\mathrm{Int}(G^*,M,I)}$$

for every fuzzy attribute implication $A \Rightarrow B$. For any $\mathbf{L}$-set $N \in \mathbf{L}^M$, $N \Rightarrow \mathcal{L}(N)$ is entailed by $\mathcal{L}$, therefore, we have $N = \mathcal{L}(N)$.

Each intent $N \in \mathrm{Int}(G^*, M, I)$ is a model of $\mathcal{B}_{\mathcal{L}}$. Now let $N \in \mathrm{Mod}(\mathcal{B}_{\mathcal{L}})$ and assume that $N \neq N^{\downarrow\uparrow}$, i.e., N is not an intent. Since $N \in \mathrm{Mod}(\mathcal{B}_{\mathcal{L}})$ we have $||Q \Rightarrow Q^{\downarrow\uparrow}||_N = 1$ for every $\mathcal{L}$-pseudo-intent $Q \in \mathcal{P}$. By definition, N is an $\mathcal{L}$-pseudo-intent and therefore $N \Rightarrow N^{\downarrow\uparrow}$ belongs to $\mathcal{B}_{\mathcal{L}}$. But now, we have

$$\begin{aligned} ||N \Rightarrow N^{\downarrow\uparrow}||_N &= \mathrm{S}(N, N)^* \to \mathrm{S}(N^{\downarrow\uparrow}, N) \\ &= 1^* \to \mathrm{S}(N^{\downarrow\uparrow}, N) \\ &= \mathrm{S}(N^{\downarrow\uparrow}, N) \\ &\neq 1, \end{aligned}$$

which is a contradiction because N does not respect this implication.

To finish the proof, we still have to show that $\mathcal{B}_{\mathcal{L}}$ is $\mathcal{L}$-non-redundant. To this end let $P \Rightarrow P^{\downarrow\uparrow} \in \mathcal{B}_{\mathcal{L}}$. We show that this implication is not entailed by $(\mathcal{B}_{\mathcal{L}} \setminus \{P \Rightarrow P^{\downarrow\uparrow}\}) \cup \mathcal{L} =: \underline{\mathcal{L}}$. As $P = \mathcal{L}(P)$, it is obviously a model of $\mathcal{L}$. We have $||Q \Rightarrow Q||_P = 1$ for every $\mathcal{L}$-pseudo-intent $Q \in \mathcal{P}$ different from P, since P is an $\mathcal{L}$-pseudo-intent. Therefore, $P \in \operatorname{Mod}(\underline{\mathcal{L}})$. We also have that $||P \Rightarrow P^{\downarrow\uparrow}||_P = \mathrm{S}(P^{\downarrow\uparrow}, P) \neq 1$ and thus P is not a model of $\mathcal{B}_{\mathcal{L}} \cup \mathcal{L}$. Hence,

$$\begin{aligned} ||P \Rightarrow P^{\downarrow\uparrow}||_{(G,M,I)} &= ||P \Rightarrow P^{\downarrow\uparrow}||_{\mathcal{B}_{\mathcal{L}} \cup \mathcal{L}} \\ &\neq ||P \Rightarrow P^{\downarrow\uparrow}||_{\underline{\mathcal{L}}}, \end{aligned}$$

showing that $\underline{\mathcal{L}}$ is not complete and therefore $\mathcal{B}_{\mathcal{L}} \cup \mathcal{L}$ is non-redundant. □

In general we write $P \Rightarrow P^{\downarrow\uparrow} \setminus \{m \in M \mid P(m) = P^{\downarrow\uparrow}(m)\}$ instead of $P \Rightarrow P^{\downarrow\uparrow}$.

Note that computing the Duquenne-Guigues-base and closing the implications from it with respect to the $\mathcal{L}$ operator will yield a different set of implications than the $\mathcal{L}$-Duquenne-Guigues-base. Let us take a look at the following example.

Example 5.7. Consider the **L**-context given in Figure 5.5. In order to ensure that its stem base and $\mathcal{L}$-stem-base exist, we use the globalisation. Further, we use the Gödel logic. The Duquenne-Guigues-base is displayed in the left column of Figure 5.6. For the background implications $\mathcal{L} := \{b \Rightarrow a,\ d \Rightarrow a,\ \{a, c\} \Rightarrow b\}$ we obtain the $\mathcal{L}$-Duquenne-Guigues-base displayed in the middle column of the figure. If we close the pseudo-intents of the stem

	a	b	c	d
x	1	0.5	0	0
y	1	0	0	0
z	0	0	1	0
t	0	0	0	0.5

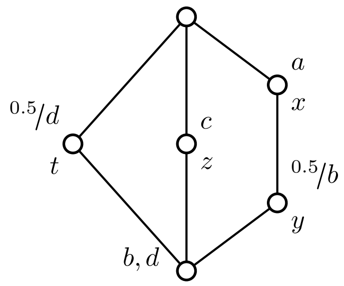

Figure 5.5.: An **L**-context and its **L**-concept lattice with the Gödel logic and the globalisation as hedge

base with respect to the $\mathcal{L}$ operator, we obtain implications of the form $\mathcal{L}(P) \Rightarrow P^{\downarrow\uparrow}$ which are displayed in the right column of the figure. As one easily sees, the latter set of implications and the $\mathcal{L}$-Duquenne-Guigues-base are different. The set

$$\{\mathcal{L}(P) \Rightarrow P^{\downarrow\uparrow} \mid P \text{ is a pseudo-intent with } \mathcal{L}(P) \neq P^{\downarrow\uparrow}\}$$

contains an additional implication, namely $\{a, b, c\} \Rightarrow d$.

As in the crisp case, the fuzzy attribute exploration can be extended to the usage of background knowledge, i.e., the interactive determination of the $\mathcal{L}$-Duquenne-Guigues-base for a given set $\mathcal{L}$ of background implications. We start by showing that the set of all intents and all $\mathcal{L}$-pseudo-intents is an $\mathbf{L}^*$-closure system:

stem base	$\mathcal{L}$-stem base	$\mathcal{L}(P) \Rightarrow P^{\downarrow\uparrow}$
${}^{0.5}/b \Rightarrow a,$	${}^{0.5}/b \Rightarrow a,$	${}^{0.5}/b \Rightarrow a,$
${}^{0.5}/a \Rightarrow a,$	${}^{0.5}/a \Rightarrow a,$	${}^{0.5}/a \Rightarrow a,$
$d \Rightarrow a, b, c,$	$c, {}^{0.5}/d \Rightarrow a, b, d,$	$a, d \Rightarrow b, c,$
$c, {}^{0.5}/d \Rightarrow a, b, d,$	$a, {}^{0.5}/d \Rightarrow b, c, d,$	$c, {}^{0.5}/d \Rightarrow a, b, d,$
$b \Rightarrow a, c, d,$	$a, d \Rightarrow b, c,$	$a, b \Rightarrow c, d,$
$a, {}^{0.5}/d \Rightarrow b, c, d,$	$a, b \Rightarrow c, d.$	$a, {}^{0.5}/d \Rightarrow b, c, d,$
$a, c \Rightarrow b, d.$		$a, b, c \Rightarrow d.$

Figure 5.6.: Different stem bases

Lemma 5.8. *Let (G, M, I) be an $\mathbf{L}$-context, let $\mathcal{L}$ be a set of fuzzy implications of (G, M, I). Further, let P and Q be intents or $\mathcal{L}$-pseudo-intents such that*

$$\mathrm{S}(P, Q)^* \le \mathrm{S}(P^{\downarrow\uparrow}, P \cap Q),$$
$$\mathrm{S}(Q, P)^* \le \mathrm{S}(Q^{\downarrow\uparrow}, P \cap Q).$$

Then, $P \cap Q$ is an intent.

Proof. Obviously, P and Q are models of any fuzzy implication from $\mathcal{B}_{\mathcal{L}} \cup \mathcal{L}$ except of $P \Rightarrow P^{\downarrow\uparrow}$ and $Q \Rightarrow Q^{\downarrow\uparrow}$. First we show that $P \cap Q$ is a model of

$$\underline{\mathcal{L}} := \mathcal{L} \cup (\mathcal{B}_{\mathcal{L}} \setminus \{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}).$$

To this end, let $A \Rightarrow B$ be a fuzzy implication from $\underline{\mathcal{L}}$. Since $P, Q \in \mathrm{Mod}(\underline{\mathcal{L}})$, we have $S(A, P)^* \le \mathrm{S}(B, P)$ and $S(A, Q)^* \le \mathrm{S}(B, Q)$. Hence,

$$\begin{aligned}
\mathrm{S}(A, P \cap Q)^* &= (\mathrm{S}(A, P) \wedge \mathrm{S}(A, Q))^* \\
&\le \mathrm{S}(A, P)^* \wedge \mathrm{S}(A, Q)^* \\
&\le \mathrm{S}(B, P) \wedge \mathrm{S}(B, Q) \\
&= \mathrm{S}(B, P \cap Q),
\end{aligned}$$

showing that $P \cap Q$ is a model of $\underline{\mathcal{L}}$. We still have to prove that $P \cap Q$ is a model of $\{P \Rightarrow P^{\downarrow\uparrow}, Q \Rightarrow Q^{\downarrow\uparrow}\}$. By assumption, we have

$$\mathrm{S}(P, P \cap Q)^* = S(P, Q)^* \le \mathrm{S}(P^{\downarrow\uparrow}, P \cap Q)$$

that is equivalent to $||P \Rightarrow P^{\downarrow\uparrow}||_{P \cap Q} = 1$. Analogously, one can show $||Q \Rightarrow Q^{\downarrow\uparrow}||_{P \cap Q} = 1$, finishing the proof. □

Note that if we choose for $(-)^*$ the globalisation, then $P \cap Q$ is an intent provided that P and Q are ($\mathcal{L}$-pseudo-)intents with $P \notin Q$ and $Q \notin P$.

Now we are interested in the closure of an $\mathbf{L}$-set with respect to the implications of the $\mathcal{L}$-base $\mathcal{B}_{\mathcal{L}}$. Therefore, we first define for each $\mathbf{L}$-set $X \in \mathbf{L}^M$ and each non-negative integer

n the **L**-sets $X^{\mathcal{L}^\bullet}, X^{\mathcal{L}^\bullet_n} \in \mathbf{L}^M$ as follows:

$$X^{\mathcal{L}^\bullet} := X \cup \bigcup \{B \otimes \mathrm{S}(A, X)^* \mid A \Rightarrow B \in \mathcal{B}_\mathcal{L}, A \neq X\}, \tag{5.5}$$

$$X^{\mathcal{L}^\bullet_n} := \begin{cases} X, & n = 0, \\ (X^{\mathcal{L}^\bullet_{n-1}})^{\mathcal{L}^\bullet}, & n \geq 1. \end{cases} \tag{5.6}$$

Further, we define an operator $\mathcal{L}^\bullet$ on these sets by

$$\mathcal{L}^\bullet(X) := \bigcup_{n=0}^{\infty} X^{\mathcal{L}^\bullet_n}. \tag{5.7}$$

Lemma 5.9. *If* $(-)^*$ *is the globalisation, then* $\mathcal{L}^\bullet$ *given by* (5.7) *is an* $\mathbf{L}^*$*-closure operator and* $\{\mathcal{L}^\bullet(X) \mid X \in \mathbf{L}^M\}$ *coincides with the set of all* $\mathcal{L}$*-pseudo-intens and intents of* (G, M, I)*.*

Proof. In order to show that $\mathcal{L}^\bullet$ is an $\mathbf{L}^*$-closure operator we have to prove the conditions

$$X \subseteq \mathcal{L}^\bullet(X), \tag{5.8}$$

$$\mathrm{S}(X_1, X_2)^* \leq \mathrm{S}(\mathcal{L}^\bullet(X_1), \mathcal{L}^\bullet(X_2)), \tag{5.9}$$

$$\mathcal{L}^\bullet(\mathcal{L}^\bullet(X)) = \mathcal{L}^\bullet(X) \tag{5.10}$$

for every **L**-sets $X, X_1, X_2 \in \mathbf{L}^M$. Condition (5.8) follows directly by the definition of $\mathcal{L}^\bullet$.

For (5.9) we have

$$\text{If } \mathrm{S}(X_1, X_2) < 1, \text{ then } \mathrm{S}(X_1, X_2)^* = 0 \leq \mathrm{S}(\mathcal{L}^\bullet(X_1), \mathcal{L}^\bullet(X_2)),$$

and we are done. If, however, $S(X_1, X_2)^* = 1$, then $X_1 \subseteq X_2$. As $\mathcal{B}_\mathcal{L}$ is given by (5.4), we have

$$X^{\mathcal{L}^\bullet} = X \cup \bigcup \{Q^{\downarrow\uparrow} \mid Q \in \mathcal{P}, Q \subset X\}.$$

Hence, if $Q \subset X_1$, then $Q \subset X_2$ and therefore we have $X_1^{\mathcal{L}^\bullet} \subseteq X_2^{\mathcal{L}^\bullet}$. Now this yields $S(\mathcal{L}^\bullet(X_1), \mathcal{L}^\bullet(X_2)) = 1$.

For (5.10), we have to prove the two inclusions. The second one, $\mathcal{L}^\bullet(X) \subseteq \mathcal{L}^\bullet(\mathcal{L}^\bullet(X))$, follows by the definition of $\mathcal{L}^\bullet$. For the first, it suffices to check $(\mathcal{L}^\bullet(X))^{\mathcal{L}^\bullet} \subseteq \mathcal{L}^\bullet(X)$. If $Q \subset \mathcal{L}^\bullet(X)$, then $Q \subset X^{\mathcal{L}^\bullet_n}$ holds for some n. Therefore, we have

$$Q^{\downarrow\uparrow} \subseteq X^{\mathcal{L}^\bullet_{n+1}} \subseteq \mathcal{L}^\bullet(X).$$

Hence, $(\mathcal{L}^\bullet(X))^{\mathcal{L}^\bullet} \subseteq \mathcal{L}^\bullet(X)$, yielding $\mathcal{L}^\bullet(\mathcal{L}^\bullet(X)) \subseteq \mathcal{L}^\bullet(X)$.

Now we show that $\{\mathcal{L}^\bullet(X) \mid X \in \mathbf{L}^M\}$ is exactly the set of intents and $\mathcal{L}$-pseudo-intents. From the properties of the globalisation and the definition of $\mathcal{L}^\bullet$ it follows that

$$\mathcal{P} \cup \mathrm{Int}(G^*, M, I) \subseteq \{\mathcal{L}^\bullet(X) \mid X \in \mathbf{L}^M\}.$$

To prove the converse inclusion, we have to show that $\mathcal{L}^\bullet \in \mathcal{P}$, and it suffices to check $\mathcal{L}^\bullet(X) \neq (\mathcal{L}^\bullet(X))^{\downarrow\uparrow}$. Now let $Q \in \mathcal{P}$ such that $Q \subset \mathcal{L}^\bullet(X)$. Then,

$$Q^{\downarrow\uparrow} \subseteq (\mathcal{L}^\bullet(X))^{\mathcal{L}^\bullet} = \mathcal{L}^\bullet(X).$$

Hence, we have $\|Q \Rightarrow Q^{\downarrow\uparrow}\|_{\mathcal{L}^\bullet(X)} = 1$. □

Remark 5.10. Note that for a general hedge, $\mathcal{L}^\bullet$ does not need be an $\mathbf{L}^*$-closure operator. For instance, choose the Goguen structure and the identity for the hedge $(-)^*$. Further, let $\mathcal{L} := \{{}^{0.3}/y \Rightarrow y\}$. Then,

$$\begin{aligned}\mathcal{L}^\bullet(\{{}^{0.2}/y\})(y) \geq (\{{}^{0.2}/y\})^{\mathcal{L}^\bullet}(y) &= \{{}^{0.2}/y\} \cup \{y \otimes (0.3 \to 0.2)\}\\ &= \{{}^{0.2}/y\} \cup \{{}^{0.(66)}/y\}\\ &= \{{}^{0.(66)}/y\},\end{aligned}$$

and $\mathcal{L}^\bullet(\{{}^{0.3}/y\})(y) = \{{}^{0.3}/y\}$. Hence, $\mathcal{L}^\bullet$ does not satisfy the monotony property, because we have $\{{}^{0.2}/y\} \subseteq \{{}^{0.3}/y\}$ but $\mathcal{L}^\bullet(\{{}^{0.2}/y\}) \nsubseteq \mathcal{L}^\bullet(\{{}^{0.3}/y\})$.

Due to the previous remark and the fact that we are only able to perform a successful attribute exploration if the chosen hedge is the globalisation, we will consider only this hedge in the rest of this section.

The lectic order is defined analogously as in Section 1.5, see (1.21). The only difference lies in the definition of "$\oplus$". This time we are using the $\mathbf{L}^*$-closure operator $(-)^{\mathcal{L}^\bullet}$ instead of $(-)^{\downarrow\uparrow}$. The next theorem follows trivially from [GW96, Běl02a].

Theorem 5.11. *The lectically first intent or $\mathcal{L}$-pseudo-intent is $\varnothing^{\mathcal{L}^\bullet}$. For a given $\mathbf{L}$-set $A \in \mathbf{L}^M$ the lectically next intent or $\mathcal{L}$-pseudo-intent is given by the $\mathbf{L}$-set $A \oplus (m,l)$, where $(m,l) \in M \times L$ is the greatest tuple such that $A <_{(m,l)} A \oplus (m,l)$. The lectically last intent or $\mathcal{L}$-pseudo-intent is M.*

Now we are prepared to present the main proposition regarding attribute exploration with background knowledge in a fuzzy setting.

Proposition 5.12. *Let $\mathbf{L}$ be a finite residuated lattice with globalisation. Further, let $\mathcal{P}$ be the unique system of $\mathcal{L}$-pseudo-intents of a finite $\mathbf{L}$-context $\mathbb{K}$ with $P_1, \dots, P_n \in \mathcal{P}$ being the first n $\mathcal{L}$-pseudo-intents in $\mathcal{P}$ with respect to the lectic order. If $\mathbb{K}$ is extended by an object g, the object intent $g^\uparrow$ of which respects the implications from*

$$\mathcal{L} \cup \{P_i \Rightarrow P_i^{\downarrow\uparrow} \mid i \in \{1, \dots, n\}\},$$

then $P_1, \dots, P_n$ remain the lectically first n $\mathcal{L}$-pseudo-intents of the extended context.

Proof. The proof of this proposition is a consequence of Proposition 5.1 (page 109), its variant without background knowledge, and the results we obtained so far in this section. Let once again $\mathbb{K} = (H, M, J)$ be the initial context and let (G, M, I) be the extended context, namely $G = H \cup \{g\}$ and $J = I \cap (H \times M)$. To put it briefly, since g^I is a model of $P_i \Rightarrow P_i^{JJ}$ for all $i \in \{1, \dots, n\}$ we have that $P_i^{JJ} = P_i^{II}$. By the definition of $\mathcal{L}$-pseudo-intents and the fact that every $\mathcal{L}$-pseudo-intent Q of (H, M, J) with $Q \subset P_i$ is lectically smaller than P_i, we have that $P_1, \dots, P_n$ are the lectically first n $\mathcal{L}$-pseudo-intents of (G, M, I). □

The algorithm for the attribute exploration with background knowledge in the fuzzy setting is the same as the one without background knowledge (Algorithm 2), except that we are using another closure operator, namely the operator given by (5.7). Further, the user has to enter the background implications in advance.

Example 5.13. Let us explore once again the size and distance of planets but this time using background implications. We start with the initial **L**-context from Figure 5.1 (page 112) with the three-element Łukasiewicz logic and the globalisation as hedge. Further, we use the following implications as background knowledge

$$\begin{aligned} n &\Rightarrow s, \\ f, {}^{0.5}/l &\Rightarrow l, \\ l, f &\Rightarrow {}^{0.5}/n, \\ {}^{0.5}/l &\Rightarrow f, \\ {}^{0.5}/l, {}^{0.5}/n &\Rightarrow l, f. \end{aligned}$$

We start the attribute exploration. The first $\mathcal{L}$-pseudo-intent is $\{\varnothing\}$ and we are asked

Do all objects have the attribute small to degree 1*?*

This is of course not true and we provide a counterexample:

	small (s)	large (l)	far (f)	near (n)
Jupiter	0	1	1	0.5

The next $\mathcal{L}$-pseudo-intent is $\{l\}$ and we are asked

Do objects having attribute large to degree 1 *also have attributes far and near to degree* 1 *and* 0.5*, respectively?*

This is true and we confirm the implication. The next $\mathcal{L}$-pseudo-intent is $\{{}^{0.5}/l, f\}$ which yields the following question:

Do objects having attributes large and far to degree 0.5 *and* 1*, respectively, also have attributes large and near to degree* 1 *and* 0.5*, respectively?*

The implication does not hold and we provide a counterexample

	small (s)	large (l)	far (f)	near (n)
Uranus	0.5	0.5	1	0

The following two $\mathcal{L}$-pseudo-intents provide us with true implications. These are

$$\begin{aligned} {}^{0.5}/s, {}^{0.5}/n &\Rightarrow s, n, \\ s, {}^{0.5}/l, f &\Rightarrow l, n. \end{aligned}$$

And the attribute exploration has stopped. Once again we have an extended **L**-context, namely the one containing Jupiter and Uranus besides the objects given in Figure 5.1. Also this time we did not have to include all the planets into the object set, just a representative part of them. The other planets with their attributes are displayed in Figure 5.2 (page 113). The three implications we confirmed during the exploration process represent the $\mathcal{L}$-base of our extended context relative to the background implications we provided at the beginning. We have seen that by the use of background knowledge we were able to shorten the exploration process.

5.4. Conclusion

Attribute exploration is one of the most powerful tools of Formal Concept Analysis having applications both in theoretical and practical research fields. In this chapter we have presented a generalisation of attribute exploration into the fuzzy setting.

For general hedges there is not yet a method to perform an attribute exploration. The problem can be traced back to the issues regarding the stem bases of **L**-contexts with general hedges. In this case, generally, the existence and uniqueness of stem bases is not ensured. Further, the set of intents and pseudo-intents does not form an **L**-closure system.

As it turned out we are able to perform a successful attribute exploration in the fuzzy setting provided the chosen hedge is the globalisation. We carried the investigation of the exploration process forward and presented its variant with background knowledge. However, before arriving at this kind of exploration we had to develop the appropriate structures for obtaining stem bases relative to background knowledge. Our expectations that using background knowledge will shorten the process of the fuzzy attribute exploration confirmed to be true in Section 5.3. We have prepared the theoretical fundamentals of attribute exploration in a fuzzy setting. The method has to be tested on real-world data sets in order to establish itself as its crisp variant has done.

6 User preferences

Our starting point in this chapter are the *attribute dependency formulas* introduced in [BS05]. These reduce the size of the concept lattice by allowing the users to model their preferences within the framework of Formal Concept Analysis. We keep the idea of modelling users' preferences, but deviate from the approach presented in [BS05]. In this chapter we develop a framework that allows the users to state their favouritisms on compound attributes, i.e., on those that contain more than one trait. In order to have a general setting, we develop our method for fuzzy data. We have already published some results of this work in [Glo12a].

First we briefly present the attribute dependencies from [BS05] in Section 6.1. Thereafter, the main work starts in Section 6.2, beginning with the development of our so-called *fuzzy attribute dependency formulas*. After studying some properties of these formulas we turn our attention to the computation of their non-redundant bases. This is motivated by the fact that the users are very likely to enter the formulas redundantly. Having a set of non-redundant formulas facilitates their further investigation and altering. In Section 6.3 we briefly discuss some other alternatives for modelling users' preferences in a fuzzy setting. We also show that an approach to fuzzy functional dependencies studied in the literature arises as a special case of our fuzzy attribute dependency formulas. An overall conclusion of this chapter is given in Section 6.4.

6.1. Attribute Dependencies

Attribute dependency formulas were introduced in [BS05] and further studied in a series of papers, see for instance [BV09b]. They were developed as a method for controlling the size of crisp concept lattices. The most appealing aspect of this technique is that the reduction is done based on the user's preferences. Indeed, the user is allowed to define a sort of order on the attributes. In accordance with these preferences, the user receives just the "interesting" concepts, "interesting" from his point of view based on the preferences he entered.

Example 6.1. Let us give a small example. Suppose we have a formal context about the products of a women clothing store. The attributes can be subdivided into groups, namely: Designer (Prada, Escada, Desigual, etc.), fabric (cotton, silk, jeans, leather, etc.), color (red, blue, yellow, etc.), type of clothing (skirt, t-shirt, blouse, dress, etc.), aim (work, dinner, sport, etc.). A customer may now define her preferences, for instance, "color is less important than designer", "type is less important than aim". Then, the concepts shown to the user will contain mainly information about the designer and the aim. However, another customer may define another set of preferences. Hence, from the same context different sets of formal concepts will be extracted for different customers, based on their preferences.

In [BS05] such preferences were modelled in the language of Formal Concept Analysis as follows: An **attribute dependency formula (AD formula)** over a set M of attributes is

$$A \sqsubseteq B,$$

where $A, B \subseteq M$. The meaning of the formula is "the attributes from A are less important than the attributes from B". The attributes from A and B should be of the same type, respectively, as we have seen above, i.e., "color is less important than designer". The AD formula $A \sqsubseteq B$ is **true** in $N \subseteq M$, written $N \vDash A \sqsubseteq B$, if

$$\text{if } A \cap N \neq \varnothing, \text{ then } B \cap N \neq \varnothing.$$

A formal concept $(C, D) \in \mathfrak{B}(G, M, I)$ **satisfies** $A \sqsubseteq B$ if $D \vDash A \sqsubseteq B$.

Example 6.2. Suppose we have the context about clothes and the AD formula "color is less important than designer". A formal concept of this context may be ({skirt1, skirt3, skirt7}, {red, skirt}). This formal concept does not satisfy the AD formula, because the intersection of its intent with the less important attributes is {red} whereas the intersection with the more important attributes is the empty set. Therefore, we will not show this concept to the user.

There are a lot of interesting results about AD formulas presented in the above cited papers. In this brief outline we will mainly focus on the computation of non-redundant bases of AD formulas, i.e., such non-redundant sets of AD formulas T_1 from which all the AD formulas of a given set T of AD formulas follow.

An attribute set $N \subseteq M$ is called a **model** of a set T of AD formulas if $N \vDash A \sqsubseteq B$, for every $A \sqsubseteq B \in T$. We denote by $\mathrm{Mod}(T)$ the set of all models of T, i.e.,

$$\mathrm{Mod}(T) := \{N \subseteq M \mid N \vDash A \sqsubseteq B, \text{ for every } A \sqsubseteq B \in T\}.$$

Theorem 6.3 ([BV09b])**.** $\mathrm{Mod}(T)$ *is a kernel system.*

Lemma 6.4 ([BV09b])**.** *For $A, B, N \subseteq M$, we have*

$$N \vDash A \sqsubseteq B \Longleftrightarrow \overline{N} \vDash B \Rightarrow A,$$

where $B \Rightarrow A$ is an attribute implication and $\overline{N}$ is the complement of N.

Remark 6.5. This lemma permits the computation of a minimal and non-redundant base of AD formulas for a given set T of AD formulas. The algorithm works as follows: First, one associates to each T a closure operator C. Afterwards, one uses NextClosure ([Gan84, GW96]) to compute a minimal base T_1 (of attribute implications) associated to C. Finally, $T_{\min} := \{B \smallsetminus A \sqsubseteq A \mid A \Rightarrow B \in T_1\}$ is the minimal and non-redundant base of AD formulas we were searching for.

6.2. Fuzzy Attribute Dependencies

In this section we develop fuzzy attribute dependency formulas for compounded attributes, i.e., for qualities which include more than one trait. For instance, the notion "wealth" might be considered as a compounded attribute consisting of "investment" and "fluency". A person who is wealthy has to have high values on both investment and fluency.

Also in this case the user has to enter a sort of order on the attributes, this time on the groups of attributes, and fix the truth values for their relevance.

First, we introduce our formulas and illustrate their usefulness on an example. After we have investigated some of their basic properties, we turn our attention to the problem that builds the core of this section. There we will develop two methods to eliminate redundancies from fAD formulas. This is an important technique, because the user is allowed to enter the formulas and it is therefore very likely that they are somewhat redundant, making their further handling more difficult than necessary. The first method acts in a straightforward way and the second one is based on a connection between the fAD formulas and fuzzy attribute implications.

Definition 6.6. A **fuzzy attribute-dependency formula** (fAD) over a set M of attributes is an expression $A \sqsubseteq B$, where $A, B \in \mathbf{L}^M$ are $\mathbf{L}$-sets of attributes. The fAD $A \sqsubseteq B$ is **true** in an $\mathbf{L}$-set $N \in \mathbf{L}^M$ for $\alpha, \beta \in L \smallsetminus \{0\}$ and $\alpha \leq \beta$, written $N \vDash_{\alpha,\beta} A \sqsubseteq B$, if the following condition is satisfied:

$$\text{if } \mathrm{S}(A, N) \geq \alpha, \text{ then } \mathrm{S}(B, N) \geq \beta. \tag{6.1}$$

For an fAD formula or a set T of fAD formulas, the values α and β are called the **thresholds** of $A \sqsubseteq B$ or T. An $\mathbf{L}$-concept $(C, D) \in \mathfrak{B}(G, M, I)$ **satisfies** $A \sqsubseteq B$ if $D \vDash_{\alpha,\beta} A \sqsubseteq B$.

For notational simplicity we will sometimes omit α and β from $\vDash_{\alpha,\beta}$ provided they are clear from the framework.

The set of all $\mathbf{L}$-concepts from $\mathfrak{B}(G, M, I)$ that satisfy a given set T of fAD formulas is denoted by $\mathfrak{B}_T(G, M, I)$, i.e.,

$$\mathfrak{B}_T(G, M, I) := \{(C, D) \in \mathfrak{B}(G, M, I) \mid D \vDash A \sqsubseteq B \text{ for every } A \sqsubseteq B \in T\}.$$

We call $\mathfrak{B}_T(G, M, I)$ together with the restricted concept order the **L-concept lattice of** (G, M, I) **constrained by** T and denote it by $\underline{\mathfrak{B}}_T(G, M, I)$.

The fAD formulas permit a two-sided modelling of the extracted $\mathbf{L}$-concepts. On the one hand, α and β provide the thresholds to which an intent has to contain all elements of A and B. On the other hand, the truth degrees of the elements contained in A and B fix the thresholds to which we want the attributes to be contained in the intent of a concept satisfying the fAD formula. Such formulas gives us much leeway. Although the thresholds are fixed for all formulas, we may control the importance of the attributes by the **L**-sets A and B. We will illustrate this fact in the forthcoming example.

In applications it is particularly useful to associate to the truth values of a residuated lattice **L** a Likert scale $\mathcal{L}$. This allows the user to have a better understanding of the truth values. For instance, let $L = \{0, 0.25, 0.5, 0.75, 1\}$ be the support set of some residuated lattice. Its associated Likert scale could be $\mathcal{L}$ = {not important, less important, important, very important, most important}, i.e, 0 =not important, 0.25 =less important, etc.

Example 6.7. Let us consider the **L**-context given in Figure 6.1. It represents the evaluation

	good team player		good organi-sational skills			adaptive towards new			confi-dential		computer skills	
	a: collaborative	b: not discriminative	c: time management	d: problem solving	e: analytical thinking	f: environment	g: assignment	h: priority	i: judgement	j: discretion	k: word processing	l: databases
1	0	0.5	0.5	1	1	0	0	0.5	0.5	0.5	1	1
2	1	1	1	1	1	1	1	1	0.5	0.5	0	0
3	0.5	0.5	0.5	1	1	0	1	1	1	1	0.5	0.5
4	1	0.5	1	1	1	0	1	1	0.5	0.5	1	1
5	0	0.5	0	0.5	0.5	0	0	0.5	0.5	0.5	0	0
6	1	1	0.5	1	1	1	1	1	0.5	0.5	0.5	0.5
7	0	0.5	0	0	0.5	0	0	0.5	0	0.5	0	0

Figure 6.1.: **L**-context about employees

of the employees of a small business regarding some qualities. Here, each quality is a compound of two or more traits. For example, an employee is a "good team player" if he/she is collaborative and not discriminative. Another example is "good organisational skills" which is evaluated based on the attributes "time management", "problem solving" and "analytical thinking". The context has 44 **L**-concepts with the Gödel logic which are far too many to be analysed by a busy manager. The manager however knows how good or bad the employees do their jobs and he is interested more in their collaboration than their organisational skills and more in their adaptivity than in their confidentiality. Therefore, he chooses the following two fAD formulas

$$\{{}^{0.5}/c, d, e\} \sqsubseteq \{a, b\} \text{ and } \{{}^{0.5}/i, {}^{0.5}/j\} \sqsubseteq \{f, g, h\}, \tag{6.2}$$

with $\alpha = 0.5$ and $\beta = 1$. Then, he obtains 11 **L**-concepts, which are displayed in Figure 6.2. The manager reconsiders his choices and realises that the company does neither send its employees to business trips nor to other companies and the employees should know their

	Extent							Intent											
	1	2	3	4	5	6	7	a	b	c	d	e	f	g	h	i	j	k	l
1	0	0	0	0	0	0.5	0	1	1	1	1	1	1	1	1	1	1	1	1
2	0	0.5	0	0	0	0.5	0	1	1	1	1	1	1	1	1	1	1	0	0
3	0	1	0	0	0	0.5	0	1	1	1	1	1	1	1	1	0.5	0.5	0	0
4	0	0	0	0	0	1	0	1	1	0.5	1	1	1	1	1	0.5	0.5	0.5	0.5
5	0	1	0	0	0	1	0	1	1	0.5	1	1	1	1	1	0.5	0.5	0	0
6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	1	0	0	1	0	0	1	0	1	0	0
7	0.5	1	0.5	0.5	0.5	1	0.5	0	1	0	0	1	0	0	1	0	0.5	0	0
8	0.5	0.5	1	0.5	0.5	0.5	0.5	0	0.5	0	0	1	0	0	1	0	1	0	0
9	0.5	1	1	1	0.5	1	0.5	0	0.5	0	0	1	0	0	1	0	0.5	0	0
10	1	1	1	1	0.5	1	0.5	0	0.5	0	0	1	0	0	0.5	0	0.5	0	0
11	1	1	1	1	1	1	1	0	0.5	0	0	0.5	0	0	0.5	0	0.5	0	0

Figure 6.2.: **L**-concepts satisfying the fAD formulas from (6.2)

	Extent							Intent											
	1	2	3	4	5	6	7	a	b	c	d	e	f	g	h	i	j	k	l
12	0	0	0.5	0.5	0	0.5	0	1	1	1	1	1	0	1	1	1	1	1	1
13	0	0.5	0.5	0.5	0	0.5	0	1	1	1	1	1	0	1	1	1	1	0	0
14	0	1	0.5	0.5	0	0.5	0	1	1	1	1	1	0	1	1	0.5	0.5	0	0
15	0	0	0.5	0.5	0	1	0	1	1	0.5	1	1	0	1	1	0.5	0.5	0.5	0.5
16	0	1	0.5	0.5	0	1	0	1	1	0.5	1	1	0	1	1	0.5	0.5	0	0

Figure 6.3.: **L**-concepts satisfying the fAD formula from (6.3) and the first fAD formula from (6.2)

priorities. Therefore, he changes the second fAD formula into

$$\{{}^{0.5}/i, {}^{0.5}/j\} \sqsubseteq \{g, {}^{0.5}/h\}. \tag{6.3}$$

Obviously the concepts from the first set of fAD formulas are a subset of the concepts from the second set of fAD formulas. The latter yields 16 concepts that are displayed in Figure 6.2 and 6.3.

Remark 6.8. The fAD formulas may be used with the Gödel and Goguen logic independent from the nature of the formulas. However, this is less advisable for the Łukasiewicz logic, as we will show in the following. Suppose we have the fAD formula ${}^{0.5}/x \sqsubseteq {}^{0.3}/z$[1] with thresholds $\alpha = \beta = 0.5$ and an **L**-set $N = \{{}^{0.5}/x, y\}$. Obviously, we have $N \models {}^{0.5}/x \sqsubseteq {}^{0.3}/z$ (because $\mathrm{S}({}^{0.5}/x, N) = 1$ and $\mathrm{S}({}^{0.3}/z, N) = 0.7$) while $N(z) = 0$. Hence, such formulas are of little avail for the user with the Łukasiewicz logic. However, if every component of the fAD formula has the truth value 1, the situation changes. Using the Łukasiewicz logic we then have $1 \to l = \min\{1 - 1 + l, 1\} = \min\{l, 1\}$ for every $l \in L$.

We have developed a framework for modelling users' preferences in a fuzzy setting. Further, we illustrated how the method works in practice. Now let us investigate some properties of these formulas.

[1]We sometimes omit curly brackets around singletones.

Proposition 6.9. *Let T be a set of fAD formulas. Then, $\underline{\mathfrak{B}}_T(G,M,I)$ is a complete lattice, which is a $\bigvee$-sublattice of $\underline{\mathfrak{B}}(G,M,I)$.*

Proof. Clearly, $\mathfrak{B}_T(G,M,I) \subseteq \mathfrak{B}(G,M,I)$ and $\underline{\mathfrak{B}}_T(G,M,I)$, with the restricted concept order, is a partially ordered subset of $\underline{\mathfrak{B}}(G,M,I)$. Further, note that $\underline{\mathfrak{B}}_T(G,M,I)$ is bounded from below because the least **L**-concept of $\underline{\mathfrak{B}}(G,M,I)$ is $(M^{\downarrow},M)$, concept which is compatible with every fAD formula. Now, we have to show that $\underline{\mathfrak{B}}_T(G,M,I)$ is closed under arbitrary suprema in $\underline{\mathfrak{B}}(G,M,I)$. To this end let $(A_j,B_j) \in \underline{\mathfrak{B}}_T(G,M,I)$ $(j \in J)$ be **L**-concepts. For any fAD formula $A \sqsubseteq B \in T$ we have $B_j \vDash A \sqsubseteq B$ for every $j \in J$. Now, if there exists $j \in J$ such that $B_j(a) < \alpha$ for some $a \in M$ with $A(a) > 0$, then $\cap_{j\in J} B_j(a) < \alpha$ and we are done because then $\cap_{j\in J} B_j \vDash A \sqsubseteq B$. Contrary, if for all $j \in J$ and all $a \in M$ such that $A(a) > 0$ we have $B_j(a) \geq \alpha$, then $\cap_{j\in J} B_j(a) \geq \alpha$ for all $a \in M$ satisfying $A(a) > 0$. Since $B_j \vDash A \sqsubseteq B$ holds for all $j \in J$, then we also have that $B_j(b) \geq \beta$ for all $j \in J$ and $b \in M$ such that $B(b) > 0$. Due to the same argument as before, we have $\cap_{j\in J} B_j(b) \geq \beta$ for all $b \in M$ such that $B(b) > 0$ and so it follows that $\cap_{j\in J} B_j \vDash A \sqsubseteq B$, showing that $\underline{\mathfrak{B}}_T(G,M,I)$ is closed under arbitrary suprema. □

Thus, after selecting the relevant formal concepts for the users we still have a complete lattice. This can be used in the further analysis of the data. For instance, the user can browse between the formal concepts, going from the more general concepts to the more concrete ones.

Remark 6.10.
1. Note that in general $\underline{\mathfrak{B}}_T(G,M,I)$ is not closed under arbitrary infima in $\underline{\mathfrak{B}}(G,M,I)$. In order to show this, let $(A_j,B_j) \in \underline{\mathfrak{B}}_T(G,M,I)$ $(j \in J)$ be **L**-concepts. For any fAD formula $A \sqsubseteq B \in T$, we have that $B_j \vDash A \sqsubseteq B$ for every $j \in J$. If there is at least one $j \in J$ such that $\mathrm{S}(A,B_j) \geq \alpha$ and $\mathrm{S}(B,B_j) \geq \beta$, then $\mathrm{S}(A,\bigcup_{j\in J} B_j) \geq \alpha$ and $\mathrm{S}(B,\bigcup_{j\in J} B_j) \geq \beta$ hold, and due to $\bigcup_{j\in J} B_j \subseteq (\bigcup_{j\in J} B_j)^{\downarrow\uparrow}$, we also have that $(\bigcup_{j\in J} B_j)^{\downarrow\uparrow} \vDash A \sqsubseteq B$. However, if we have $\mathrm{S}(A,B_j) < \alpha$ for all $j \in J$, then $\mathrm{S}(A,(\bigcup_{j\in J} B_j)^{\downarrow\uparrow}) \geq \alpha$ and $\mathrm{S}(B,(\bigcup_{j\in J} B_j)^{\downarrow\uparrow}) < \beta$ may happen, which yields $(\bigcup_{j\in J} B_j)^{\downarrow\uparrow} \nvDash A \sqsubseteq B$.
2. The top elements of $\underline{\mathfrak{B}}_T(G,M,I)$ and $\underline{\mathfrak{B}}(G,M,I)$ may be different from each other. They are the same if T does not contain an fAD formula $A \sqsubseteq B$ such that every attribute from A is shared by all objects from G with a the truth degree $\geq \alpha$ and there is at least one attribute from B which is not shared by all objects from G with a truth value $\geq \beta$. If there is no such fAD formula in T, then $(G,G^{\uparrow})$ satisfies every fAD formula from T and hence is the upper bound of $\underline{\mathfrak{B}}_T(G,M,I)$.

Denote by fADF the set of all fAD formulas. Then, $\vDash$ induces two mappings between fADF (the formulas) and all **L**-concepts (the models/structures). The two mappings are

$$\begin{aligned}\mathrm{Str} &: \mathfrak{P}(fADF) \to \mathfrak{P}(\mathfrak{B}(G,M,I)),\\ \mathrm{Fml} &: \mathfrak{P}(\mathfrak{B}(G,M,I)) \to \mathfrak{P}(fADF).\end{aligned}$$

Let T be a set of fAD formulas and let $C \in \mathfrak{P}(\mathfrak{B}(G,M,I))$ be a subset of concepts. Define

$$\begin{aligned}\mathrm{Str}(T) &:= \{(A,B) \in \mathfrak{B}(G,M,I) \mid (A,B) \vDash \varphi \text{ for each } \varphi \in T\},\\ \mathrm{Fml}(C) &:= \{\varphi \in \mathrm{fADF} \mid (A,B) \vDash \varphi \text{ for each } (A,B) \in C\}.\end{aligned}$$

Obviously, the mappings Str and Fml form a crisp Galois connection between (the power sets of) fADF and $\mathfrak{B}(G,M,I)$.

The next lemma shows further properties of the fAD formulas. On the one hand, it allows us to reduce the number of formulas in a simple way. On the other hand, it permits the testing of semantic entailment on simpler formulas. Before presenting this lemma, we need to introduce some further notions.

Definition 6.11. An **L**-set $N \in \mathbf{L}^M$ is a **model** of a set T of fAD formulas if $N \vDash A \sqsubseteq B$ holds for each $A \sqsubseteq B \in T$. Let $\mathrm{Mod}(T)$ denote the set of all models of T, i.e.,

$$\mathrm{Mod}(T) := \{N \in \mathbf{L}^M \mid N \vDash A \sqsubseteq B, \text{ for each } A \sqsubseteq B \in T\}.$$

An fAD formula $A \sqsubseteq B$ **follows semantically** from T, written $T \vDash A \sqsubseteq B$, if for each $N \in \mathrm{Mod}(T)$, we have $N \vDash A \sqsubseteq B$.

Lemma 6.12. *i) $N \vDash A \sqsubseteq \{^{l_1}/y_1, \ldots, ^{l_n}/y_n\}$ if and only if $N \vDash A \sqsubseteq \{^{l_i}/y_i\}$ for all $i \in \{1, \ldots, n\}$. ii) For each set T of fAD formulas and each fAD formula φ, we have $T \vDash \varphi$ if and only if $\lfloor T \rfloor \vDash \varphi$, where $\lfloor T \rfloor := \{A \sqsubseteq \{^l/y\} \mid A \sqsubseteq B \in T$ and $B(y) = l\}$.*

Proof. i) If N trivially satisfies the formula, $\mathrm{S}(A,N) < \alpha$, then we are done. Now suppose it does satisfy the formula in a non-trivial way. Thus, we have $\mathrm{S}(A,N) \geq \alpha$ and $\mathrm{S}(\{^{l_1}/y_1, \ldots, ^{l_n}/y_n\}, N) \geq \beta$. By the definition of S, the latter holds if and only if we have $\mathrm{S}(\{^{l_i}/y_i\}, N) \geq \beta$ for all $i \in \{1, \ldots, n\}$. Thus, $N \vDash A \sqsubseteq \{^{l_i}/y_i\}$ for all $i \in \{1, \ldots, n\}$.
ii) We have to show that $\mathrm{Mod}(T) = \mathrm{Mod}(\lfloor T \rfloor)$. Suppose this is not the case. Then, there must be a model $N \in \mathrm{Mod}(T)$ such that $N \notin \mathrm{Mod}(\lfloor T \rfloor)$. Let $A \sqsubseteq B \in T$ be an fAD formula. Since N is a model of T, we have $N \vDash A \sqsubseteq B$. By i), for any $^l/b \in B$ it holds that $M \vDash A \sqsubseteq {^l/b}$, a contradiction. Therefore, $\mathrm{S}(\mathrm{Mod}(T), \mathrm{Mod}(\lfloor T \rfloor)) = 1$. Analogously, one can show $\mathrm{S}(\mathrm{Mod}(\lfloor T \rfloor), \mathrm{Mod}(T)) = 1$, yielding that $\mathrm{Mod}(\lfloor T \rfloor) = \mathrm{Mod}(T)$. ☐

Due to Lemma 6.12 we may merge fAD formulas with the same left-hand side into a single fAD formula. The new formula is true in a model if and only if all its component fAD formulas are true in that model. Further, this lemma allows us also to test semantic entailment in fAD formulas $A \sqsubseteq \{^l/y\}$ rather than on the whole $A \sqsubseteq B$.

As we already stated at the beginning, it is very likely that the formulas entered by the user, are redundant. Wishful thinking suggests to have a set of non-redundant formulas, because these are then easier to follow and to modify. Therefore, in the following we will develop methods for removing such redundancies. In order to do so, we will first study the connection between the models of fAD formulas and $\mathbf{L}^*$-closure systems in a series of propositions. It will turn out that any $\mathbf{L}^*$-closure system can be described by a set of fAD formulas.

Proposition 6.13. *Let T be a set of fAD formulas. Then,* $\mathrm{Mod}(T)$ *is an* $\mathbf{L}^*$*-closure system with* $(-)^*$ *being the globalisation.*

Proof. Let $\{N_j \in \mathrm{Mod}(T) \mid j \in J\}$. We will show that $\mathrm{Mod}(T)$ is closed under arbitrary intersection, i.e., $\bigcap_{j \in J} N_j$ is a model of T. For any fAD formula $A \sqsubseteq B \in T$, we have $N_j \vDash A \sqsubseteq B$ for every $j \in J$. But now, if there is $j \in J$ such that $N_j(a) < \alpha$ for some $a \in M$

with $A(a) > 0$, then $\cap_{j \in J} N_j(a) < \alpha$ and we are done since then $\cap_{j \in J} N_j \models A \sqsubseteq B$. Contrary, if for all $j \in J$ and all $a \in M$ with $A(a) > 0$ we have $N_j(a) \geq \alpha$, then $\cap_{j \in J} N_j(a) \geq \alpha$ for all $a \in M$ with $A(a) > 0$. Since $N_j \models A \sqsubseteq B$ holds for all $j \in J$, we also have that $N_j(b) \geq \beta$ for all $j \in J$ and $b \in M$ with $B(b) > 0$. Due to the same argument as before, we obtain $\cap_{j \in J} N_j(b) \geq \beta$ for all $b \in M$ with $B(b) > 0$ and hence we have $\cap_{j \in J} N_j \models A \sqsubseteq B$, showing that $\mathrm{Mod}(T)$ is closed under arbitrary intersections.

According to Theorem 1.4 (page 9) $\mathrm{Mod}(T)$ is an $\mathbf{L}^*$-closure system if and only if it is closed under arbitrary intersections and $a^* \to N$ is a model of T for any $N \in \mathrm{Mod}(T)$ and any $a \in L$. Due to the first part of this proof, we just have to show the latter. However, this condition only holds if $(-)^*$ is the globalisation. Then, we have

$$\mathrm{S}(A, a^* \to N) = a^* \to \mathrm{S}(A, N) = \begin{cases} 1, & a = 0, \\ \mathrm{S}(A, N), & a = 1, \end{cases}$$

i.e., $a^* \to N$ trivially satisfies any fAD formula if $a = 0$ or we do not gain anything new to N in the case that $a = 1$. □

In order to use general hedges but still have the result of the previous proposition we have to impose some restrictions on the thresholds.

Corollary 6.14. *Let T be a set of fAD formulas with thresholds $\alpha = \beta = 1$. Then,* $\mathrm{Mod}(T)$ *is an $\mathbf{L}^*$-closure system.*

Proof. The first part from the proof of Proposition 6.13 still holds. For the second part we still have to show that $a^* \to N$ is a model of T for any $N \in \mathrm{Mod}(T)$ and any $a \in L$. Let $A \sqsubseteq B \in T$. Then, we have

$$\mathrm{S}(A, a^* \to N) = a^* \to \mathrm{S}(A, N) = \begin{cases} 1, & a^* = 0 \text{ or } a = \mathrm{S}(A, a^* \to N) = 1, \\ \mathrm{S}(A, N), & \text{else.} \end{cases}$$

Due to the fact that we chose $\alpha = \beta = 1$ we have $\mathrm{S}(A, N) = 1$ or $\mathrm{S}(A, N) = 0$. Since the same applies to $\mathrm{S}(B, a^* \to N)$, we are done. □

Based on the previous two results we have the following:

Proposition 6.15. *Let $\mathcal{S}$ be an $\mathbf{L}^*$-closure system on M. The following hold:*

(i) *There is a set T of fAD formulas over M with thresholds $\alpha = \beta = 1$ such that $\mathcal{S} = \mathrm{Mod}(T)$.*

(ii) *There is a set T of fAD formulas over M such that $\mathcal{S} = \mathrm{Mod}(T)$ provided that $(-)^*$ is the globalisation.*

Proof. i) Define a set T of fAD formulas by $T := \{A \sqsubseteq C_{\mathcal{S}}(A) \mid A \in \mathbf{L}^M\}$, where $C_{\mathcal{S}}(A)$ is the closure of A given by the $\mathbf{L}^*$-closure operator $C_{\mathcal{S}}$. Let $N \in \mathcal{S}$, i.e., $N = C_{\mathcal{S}}(N)$. We have to show that N is a model of T. Thus let $N \models A \sqsubseteq C_{\mathcal{S}}(A)$ for every $A \sqsubseteq C_{\mathcal{S}}(A) \in T$. If $\mathrm{S}(A, N) < 1$, then $N \models A \sqsubseteq C_{\mathcal{S}}(A)$ and we are done. Now take $\mathrm{S}(A, N) \geq 1$, meaning that $A \subseteq N$. Since $C_{\mathcal{S}}$ is a closure operator we have $C_{\mathcal{S}}(A) \subseteq C_{\mathcal{S}}(N) = N$, hence $\mathrm{S}(C_{\mathcal{S}}(A), N) \geq 1$, i.e., $N \models A \sqsubseteq C_{\mathcal{S}}(A)$. Thus, N is a model of T and we have the first

inclusion, namely $\mathcal{S} \subseteq \mathrm{Mod}(T)$.
For the converse, let $N \in \mathrm{Mod}(T)$. Since $\mathrm{S}(N,N) \geq 1$ obviously holds, we must also have $\mathrm{S}(C_{\mathcal{S}}(N),N) \geq 1$, yielding that $N = C_{\mathcal{S}}(N)$, i.e., $N \in \mathcal{S}$ and hence $\mathrm{Mod}(T) \subseteq \mathcal{S}$.
ii) Is similar to i), with the necessary changes as shown in Proposition 6.13 and Corollary 6.14. □

According to Proposition 6.13, $\mathrm{Mod}(T)$ is an $\mathbf{L}^*$-closure system, so it must exist an $\mathbf{L}^*$-closure operator $C_{\mathrm{Mod}(T)} : \mathbf{L}^M \to \mathbf{L}^M$ such that $N = C_{\mathrm{Mod}(T)}(N)$ if and only if $N \in \mathrm{Mod}(T)$. Hence, by definition, $C_{\mathrm{Mod}(T)}(N)$ is the least model in $\mathrm{Mod}(T)$ which contains N. This definition of the $\mathbf{L}^*$-closure operator does not provide a useful method for computing the closure of a given N. First, because one has to iterate over all models in $\mathrm{Mod}(T)$ and second, such an iteration may be impossible if $\mathbf{L}$ is infinite, because then $\mathrm{Mod}(T)$ is infinite.

Similarly to (fuzzy) attribute implications we proceed as follows: For any set T of fAD formulas with thresholds α, β and for any $\mathbf{L}$-set $N \in \mathbf{L}^M$ of attributes, we define an $\mathbf{L}$-set $N^T \in \mathbf{L}^M$ of attributes as follows:

$$N^T := N \cup \bigcup \{\beta \otimes B \mid A \sqsubseteq B \in T,\, \mathrm{S}(A,N) \geq \alpha\}. \tag{6.4}$$

Further, we define an $\mathbf{L}$-set $N^{T_n} \in \mathbf{L}^Y$ of attributes for each non-negative integer by

$$N^{T_n} := \begin{cases} N, & n = 0, \\ (N^{T_{n-1}})^T, & n \geq 1, \end{cases} \tag{6.5}$$

and define an operator $\mathrm{cl_T} : \mathbf{L}^M \to \mathbf{L}^M$ by

$$\mathrm{cl_T}(N) := \bigcup_{n=0}^{\infty} N^{T_n}. \tag{6.6}$$

Proposition 6.16. *For each $N \in \mathrm{Mod}(T)$ we have $\mathrm{cl_T}(N) = N$.*

Proof. By definition $N \subseteq N^T$ holds. Conversely, for any $A \sqsubseteq B \in T$ and any $N \in \mathrm{Mod}(T)$ we have $N \models A \sqsubseteq B$. If $\mathrm{S}(A,N) < \alpha$, then $N^T = N$. If $\mathrm{S}(A,N) \geq \alpha$, then $\mathrm{S}(B,N) \geq \beta$ must hold since $N \in \mathrm{Mod}(T)$. From $\mathrm{S}(B,N) \geq \beta$ we get $\beta \otimes B \subseteq N$ by the adjointness property and hence $N^T = N \cup \{\beta \otimes B\} = N$. The definitions of N^{T_n} and $\mathrm{cl_T}$ yield $N = N^{T_0} = N^{T_1} = \cdots$ for every $N \in \mathrm{Mod}(T)$. Thus, $N = \bigcup_{n=0}^{\infty} N^{T_n} = \mathrm{cl_T}(N)$. □

The next lemma shows that the $\mathbf{L}^*$-closure operator defined on the models of T coincides with the $\mathrm{cl_T}$-operator defined in (6.6).

Lemma 6.17. *Let T be a set of fAD formulas over M. Further let both M and $\mathbf{L}$ be finite. Then, $\mathrm{cl_T}$ is an $\mathbf{L}^*$-closure operator such that $C_{\mathrm{Mod}(T)}(N) = \mathrm{cl_T}(N)$ for each $N \in \mathbf{L}^M$.*

Proof. $C_{\mathrm{Mod}(T)}$ is an $\mathbf{L}^*$-closure operator, therefore it suffices to check that $C_{\mathrm{Mod}(T)}$ and $\mathrm{cl_T}$ coincide. To this end let $N \in \mathbf{L}^M$ be an $\mathbf{L}$-set of attributes. By the definition of $\mathrm{cl_T}$ we have $N \subseteq \mathrm{cl_T}(N)$. We still have to show that $\mathrm{cl_T}(N)$ belongs to $\mathrm{Mod}(T)$ and that $\mathrm{cl_T}(N)$ is the least model containing N. First of all note that the finiteness of $\mathbf{L}$ and M imply that $\mathbf{L}^M$ is finite and that there exists a non-negative integer k such that $\mathrm{cl_T}(N) = N^{T_k}$.

Hence there can be only finitely many proper inclusions in $N^{T_0} \subseteq N^{T_1} \subseteq \cdots \subseteq N^{T_k} \subseteq \cdots$. Thus, there always exists some k satisfying $\mathrm{cl}_\mathrm{T}(N) = N^{T_k}$ and $N^{T_k} = N^{T_{k+1}} = \cdots$.

It remains to show that $\mathrm{cl}_\mathrm{T}(N) \in \mathrm{Mod}(T)$, i.e., for any fAD formula $A \sqsubseteq B \in T$ we have that $\mathrm{cl}_\mathrm{T}(N) \models A \sqsubseteq B$. If $\mathrm{S}(A, \mathrm{cl}_\mathrm{T}(N)) < \alpha$, we are done. Now suppose that we have $\mathrm{S}(A, \mathrm{cl}_\mathrm{T}(N)) \geq \alpha$. It follows that $\mathrm{cl}_\mathrm{T}(N) = N \cup \{\beta \otimes B\}$. Obviously, $\mathrm{S}(B, N \cup \{\beta \otimes B\}) \geq \beta$, proving that $\mathrm{cl}_\mathrm{T}(N)$ is a model of T which contains N. For any $X \in \mathrm{Mod}(T)$ such that $N \subseteq X$ we have to show that $\mathrm{cl}_\mathrm{T}(N) \subseteq X$. This easily follows by the properties of closure operators and by Proposition 6.16. In fact, we have $\mathrm{cl}_\mathrm{T}(N) \subseteq \mathrm{cl}_\mathrm{T}(X) = X$. □

Based on the previous result we present Algorithm 3 for the computation of the closure $C_{\mathrm{Mod}(T)}(N)$ of an $\mathbf{L}$-set $N \in \mathbf{L}^M$ of attributes with respect to a set T of fAD formulas provided that $\mathbf{L}$ and M are finite.

Algorithm 3: Closure(N, T)

```
repeat
    take A ⊑ B ∈ T such that S(A, N) ≥ α and S(B, N) < β;
    set N to N ∪ {β ⊗ B};
until forall A ⊑ B ∈ T, (S(A, N) < α) or (S(A, N) ≥ α and S(B, N) ≥ β);
return N
```

Now let us investigate the connection between fAD formulas and implications. Therefore, denote by $\mathrm{Imp}(T)$ the set of implications obtained by the set of fAD formulas T, i.e.,

$$\mathrm{Imp}(T) := \{A \Rightarrow B \mid \text{ for all } A \sqsubseteq B \in T\}.$$

If we choose more restrictive values for α and β, we have the following connection between fuzzy implications and fAD formulas:

Proposition 6.18. *If we choose $\alpha = \beta = 1$, then for every set T of fAD formulas it holds*

$$\mathrm{Mod}(\mathrm{Imp}(T)) \subseteq \mathrm{Mod}(T).$$

Proof. Let T be a set of fAD formulas and $A \Rightarrow B$ an implication from $\mathrm{Imp}(T)$. Further, let $N \in \mathrm{Mod}(\mathrm{Imp}(T))$. Since N is a model of $\mathrm{Imp}(T)$, we have

$$||A \Rightarrow B||_N = \mathrm{S}(A, N) \to \mathrm{S}(B, N) = 1.$$

The latter holds whenever $\mathrm{S}(A, N) \leq \mathrm{S}(B, N)$. Since we chose $\alpha = \beta = 1$ this means that N is a model of T. □

Definition 6.19. Two sets T_1 and T_2 of fAD formulas are called **equivalent**, written $T_1 \equiv T_2$, if for each $\varphi_1 \in T_1$ and each $\varphi_2 \in T_2$ we have $T_1 \models \varphi_2$ and $T_2 \models \varphi_1$.

Proposition 6.20. *Let T_1 and T_2 be sets of fAD formulas. Then, the following are equivalent:*

(i) $\mathrm{Mod}(T_1) = \mathrm{Mod}(T_2)$*;*

(ii) *for any fAD formula φ we have $T_1 \models \varphi \iff T_2 \models \varphi$;*

(iii) $T_1 \equiv T_2$.

Proof. "i)⇒ ii)" follows directly by the definition of $\models$.
"ii)⇒ iii)" Let $\varphi_1 \in T_1$. Then, $T_1 \models \varphi_1$ because φ_1 is true in every model of T_1. By ii) we have that $T_2 \models \varphi_1$. Dually, one can show that $T_1 \models \varphi_2$ follows from $\varphi_2 \in T_2$. Hence, $T_1 \equiv T_2$.
"iii)⇒i)" Assume that $\mathrm{Mod}(T_1)$ and $\mathrm{Mod}(T_2)$ are two different $\mathbf{L}$-sets. Then, there must be an $\mathbf{L}$-set $N \in \mathbf{L}^M$ such that $N \in \mathrm{Mod}(T_1)$ and $N \notin \mathrm{Mod}(T_2)$. This means that there is $A \sqsubseteq B \in T_2$ such that $N \not\models A \sqsubseteq B$. However, by iii) we have that $T_1 \equiv T_2$, which yields $T_1 \models A \sqsubseteq B$, and because N is a model of T_1 we have $N \models A \sqsubseteq B$, a contradiction. For $\mathrm{S}(\mathrm{Mod}(T_2), \mathrm{Mod}(T_1)) < 1$ one can proceed analogously. Thus, we have that $\mathrm{Mod}(T_1) = \mathrm{Mod}(T_2)$ showing i). □

Now we are prepared to introduce non-redundant bases.

Definition 6.21. A set T_1 of fAD formulas is called a **non-redundant base** of T if $T \equiv T_1$ and there is no $T_2 \subset T_1$ with $T_2 \equiv T$. A set T_1 of fAD formulas is called a **minimal base** of T if $T \equiv T_1$ and for each T_2 with $T \equiv T_2$, we have $|T_1| \leq |T_2|$.

Obviously, if T_1 is a minimal base of T, then T_1 is a non-redundant base of T. The converse is not true in general.

For a given set T of fAD formulas we may compute a non-redundant base as follows: First note that if $T_1 := T \setminus \{A \sqsubseteq B\}$ and $T_1 \models A \sqsubseteq B$, then $T \equiv T_1$. We may then remove fAD formulas $A \sqsubseteq B$ from T step-by-step until there is no $T_1 \subset T$ such that $T_1 \equiv T$. The computation of a non-redundant base with this method is quite laborious. In what follows we present another connection between fuzzy attribute implications and fAD formulas which will considerably simplify this task, provided that $(-)^*$ is the globalisation.

Lemma 6.22. *Let T be a set of fAD formulas. We have*

$$\mathrm{Mod}(T) = \mathrm{Mod}(\mathrm{Imp}(T^*)),$$

where

$$\mathrm{Imp}(T^*) := \{\alpha \otimes A \Rightarrow \beta \otimes B \mid \text{ for all } A \sqsubseteq B \in T,\ \alpha, \beta \text{ thresholds of } T\} \tag{6.7}$$

where the truth values of the implications from $\mathrm{Imp}(T^*)$ *are computed by using the globalisation.*

Proof. Let $N \in \mathrm{Mod}(T)$ and $A \sqsubseteq B \in T$. There are two cases:
1) $\mathrm{S}(A, N) \geq \alpha$ and $\mathrm{S}(B, N) \geq \beta$ both hold. Then, for every attribute $m \in M$, we have $A(m) \to N(m) \geq \alpha$ which by the adjointness property gives us $\alpha \otimes A(m) \leq N(m)$ and therefore $\mathrm{S}(\alpha \otimes A, N) = 1$. Thus, $\mathrm{S}(\beta \otimes B, N) = 1$. Hence, we have

$$||\alpha \otimes A \Rightarrow \beta \otimes B||_N = \mathrm{S}(\alpha \otimes A, N)^* \to \mathrm{S}(\beta \otimes B, N) = 1^* \to 1 = 1.$$

2) We have $\mathrm{S}(A, N) < \alpha$, which is equivalent to $\mathrm{S}(\alpha \otimes A, N) < 1$. Therefore,

$$||\alpha \otimes B \Rightarrow \beta \otimes B||_N = \mathrm{S}(\alpha \otimes A, N)^* \to \mathrm{S}(\beta \otimes B, N) = 0 \to \mathrm{S}(\beta \otimes B, N) = 1.$$

Cases 1) and 2) show that N is a model of $\mathrm{Imp}(T^*)$.

For the converse let $N \in \mathrm{Mod}(\mathrm{Imp}(T^*))$. Then, we have

$$||\alpha \otimes A \Rightarrow \beta \otimes B||_N = \mathrm{S}(\alpha \otimes A, N)^* \to \mathrm{S}(\beta \otimes B, N) = 1 \tag{6.8}$$

for any fuzzy attribute implication $A \Rightarrow B \in \mathrm{Imp}(T^*)$. Equation 6.8 holds if and only if one of the following cases apply:

1) $(\mathrm{S}(\alpha \otimes A, N)^* = 1 \text{ and } \mathrm{S}(\beta \otimes B, N) = 1) \Longleftrightarrow (\mathrm{S}(A, N) \geq \alpha \text{ and } \mathrm{S}(B, N) \geq \beta)$,
2) $\mathrm{S}(\alpha \otimes A, N)^* = 0 \Longleftrightarrow \mathrm{S}(\alpha \otimes A, N) < 1 \Longleftrightarrow \mathrm{S}(A, N) < \alpha$.

In both cases, it follows that $N \vDash_{\alpha,\beta} A \sqsubseteq B$. □

Thus for an **L**-set $N \in \mathbf{L}^M$ we have

$$N \vDash_{\alpha,\beta} A \sqsubseteq B \Longleftrightarrow (||\alpha \otimes A \Rightarrow \beta \otimes B||_N = 1 \text{ with } (-)* \text{ being the globalisation}).$$

With this link between fAD formulas and fuzzy attribute implications we may easily compute a minimal base for any set T of fAD formulas. First we build the set $\mathrm{Imp}(T^*)$ associated to T as given by (6.7). Afterwards, we compute a minimal base of attribute implications $\mathcal{B}_{T^*}$ for this set. Finally, from $\mathcal{B}_{T^*}$ we obtain a minimal base of fAD formulas for T by

$$\mathcal{B}_T := \{A^\blacklozenge \sqsubseteq B^\blacklozenge \setminus A^\blacklozenge \mid A \Rightarrow B \in \mathcal{B}_{T^*}\},$$

where

$$A^\blacklozenge := \bigvee\{C \in \mathbf{L}^M \mid \alpha \otimes C = \alpha \otimes A\} \text{ and } B^\blacklozenge := \bigvee\{D \in \mathbf{L}^M \mid \alpha \otimes D = \alpha \otimes B\}. \tag{6.9}$$

Example 6.23. Let (G, M, I) be an **L**-context with $M = \{a, b \ldots, h\}$. We will use the three-element Gödel logic. Suppose the user enters the following fAD formulas:

$$\begin{array}{ll} \{a,b\} \sqsubseteq \{^{0.5}/c, {}^{0.5}/d\}, & \{a,b\} \sqsubseteq \{e,f\}, \\ \{e,f\} \sqsubseteq \{c,d\}, & \{g,h\} \sqsubseteq \{c,d\}, \\ \{e,f\} \sqsubseteq \{^{0.5}/g, {}^{0.5}/h\}, & \end{array}$$

with thresholds $\alpha = 0.5$ and $\beta = 1$. We may now use (6.7) to transform these formulas into attribute implications in a fuzzy setting. Afterwards, we may compute a stem base for these implications. We obtain:

$$\begin{array}{l} \{^{0.5}/e, {}^{0.5}/f\} \Rightarrow \{c, d, {}^{0.5}/g, {}^{0.5}/h\}, \\ \{^{0.5}/a, {}^{0.5}/b\} \Rightarrow \{c, d, e, f, {}^{0.5}/g, {}^{0.5}/h\}, \\ \{^{0.5}/g, {}^{0.5}/h\} \Rightarrow \{c, d\}. \end{array}$$

Using (6.9) we obtain the following minimal non-redundant base of fAD formulas:

$$\begin{array}{l} \{e,f\} \sqsubseteq \{c, d, {}^{0.5}/g, {}^{0.5}/h\}, \\ \{a,b\} \sqsubseteq \{c, d, e, f, {}^{0.5}/g, {}^{0.5}/h\}, \\ \{g,h\} \sqsubseteq \{c,d\}. \end{array}$$

The possibility of computing a non-redundant base allows the user to review his choices and alter them conveniently.

6.3. Related Works

We presented here an approache for modelling users' preferences in a fuzzy setting. However, there might be other formulas/models of formulas which are meaningful for the user. For instance, a variant might be the following: An **L**-set $N \in \mathbf{L}^M$ is a model of the fAD formula $A \sqsubseteq B$ with $A, B \in \mathbf{L}^M$ if $\bigvee_{m\in M}(A \cap N)(m) \leq \bigvee_{m\in M}(B \cap N)(m)$. Another alternative could be: An **L**-set $N \in \mathbf{L}^M$ is a model of the fAD formula $A \sqsubseteq B$ if $\bigvee_{a\in A} N(a) \leq \bigvee_{b\in B} N(b)$. In both cases an **L**-concept satisfies the formula if its intent is a model of the formula.

A somehow different approach of modelling users' preferences within the framework of crisp Formal Concept Analysis was presented in [Obi12]. Starting with the users preferences on objects, one obtains a preference relation on concepts and afterwards on the attributes. The method embeds preference logic into the tools of Formal Concept Analysis.

Straightforward Generalisation

Let us go back now to the starting point of our fuzzy attribute dependency formulas. Unlike those formulas ours were developed for compound attributes. Thus, one could ask how the method developed in [BS05] translates into the language of Formal Fuzzy Concept Analysis. Let us briefly discuss this matter.

A *fuzzy attribute-dependency formula* (fAD formula) over a set M of attributes is an expression $A \sqsubseteq B$, where $A, B \in \mathbf{L}^M$ are **L**-sets of attributes. The fAD formula $A \sqsubseteq B$ is *true* in an **L**-set $N \in \mathbf{L}^M$, if the following condition is satisfied:

$$\text{if } A \cap N \neq \varnothing^2, \text{ then } B \cap N \neq \varnothing.$$

The verbal meaning of $A \sqsubseteq B$ is once again "the attributes from A are less important than the attributes from B". However, such formulas may not be sufficient in yielding the interesting concepts in accordance with the users' preferences. For instance, there might be some attribute $m \in M$ such that $m \in A \cap N$ but there might be just one attribute $n \in M$ such that $(B \cap N)(n) > 0$ and further $(B \cap N)(n) = 0.000\ldots001$. This is of little use for our knowledge discovery and concept reduction. Indeed, the user's interest might not be awaken by concepts which contain just a small fragment of his preferences. Therefore, we have to use some thresholds to control the values in the intersection. For two **L**-sets $C, D \in \mathbf{L}^M$ and a truth value $\alpha \in L \setminus \{0\}$, we say that $C \cap D$ is α-**true** if there is at least one attribute $m \in M$ such that $(C \cap D)(m) \geq \alpha$. Now, we may redefine the notion of *truth*. The fAD formula $A \sqsubseteq B$ is *true* in an **L**-set $N \in \mathbf{L}^M$ for $\alpha, \beta \in L \setminus \{0\}$ and $\alpha \leq \beta$, written $N \vDash_{\alpha,\beta} A \sqsubseteq B$, if the following condition is satisfied:

$$\text{if } A \cap N \text{ is } \alpha\text{-true, then } B \cap N \text{ is } \beta\text{-true.} \tag{6.10}$$

In applications the user has to enter a set of fAD formulas and values for α and β which will then be used for every entered fAD formula. A *model* of a set T of fAD formulas is an **L**-set $N \in \mathbf{L}^M$ such that $N \vDash_{\alpha,\beta} A \sqsubseteq B$ for all $A \sqsubseteq B \in T$. We denote by

$$\mathrm{Mod}(T) := \{N \in \mathbf{L}^M \mid N \vDash_{\alpha,\beta} A \sqsubseteq B \text{ for each } A \sqsubseteq B \in T\}$$

[2] $\varnothing \in \mathbf{L}^M$ is an **L**-set such that $\varnothing(m) = 0$ for all $m \in M$.

the set of all models of T. In the discrete case we have seen that $\mathrm{Mod}(T)$ is a kernel system. What about this result in the fuzzy setting? Due to Theorem 1.6 (page 9) it suffices to show that $\mathrm{Mod}(T)$ is closed under arbitrary unions and that $a^* \otimes N \in \mathrm{Mod}(T)$ holds for all $a \in L$ and $N \in \mathrm{Mod}(T)$. The first is easy to see. Thus, we have

$$\{N_j \in \mathrm{Mod}(T) \mid j \in J\} \Longrightarrow \bigcup_{j \in J} N_j \in \mathrm{Mod}(T).$$

The second condition is satisfied if and only if * is the globalisation. Then,

$$(a^* \otimes N)(m) = a^* \otimes N(m) = \begin{cases} 0, & a = 0, \\ N(m), & a = 1, \end{cases}$$

holds for all $m \in M$, i.e., we either have 0 and $a^* \otimes N$ trivially satisfies any fAD formula, or we do not gain anything new. Thus, we have a similar situation as in Proposition 6.13, namely that the models form an $\mathbf{L}^*$-closure and $\mathbf{L}^*$-kernel system, provided the hedge is the globalisation. For such formulas the requirement that $\mathrm{Mod}(T)$ should form an $\mathbf{L}^*$-closure or $\mathbf{L}^*$-kernel system with general hedges seems to be too strong.

In this setting we have the dual of Lemma 6.12. Indeed, $N \models \{^{l_1}/y_1, \dots, {}^{l_n}/y_n\} \sqsubseteq B$ holds if and only if $N \models \{^{l_i}/y_i\} \sqsubseteq B$ holds for all $i \in \{1, \dots, n\}$. Note that such an assertion does not hold for the fAD formulas for compound attributes. Suppose N satisfies an fAD formula $\{^{l_1}/y_1, {}^{l_2}/y_2\} \sqsubseteq B$ such that $\mathrm{S}(\{^{l_1}/y_1, {}^{l_2}/y_2\}, N) < \alpha$ and $\mathrm{S}(B, N) < \beta$. Further, let $\mathrm{S}(\{^{l_1}/y_1\}, N) < \alpha$ and $\mathrm{S}(\{^{l_2}/y_2\}, N) \geq \alpha$. Obviously, $N \models \{^{l_1}/y_1\} \sqsubseteq B$ but $N \not\models \{^{l_2}/y_2\} \sqsubseteq B$, because $\mathrm{S}(B, N) < \beta$.

Connection to Functional Dependencies

Let us now we present connections between our fuzzy attribute dependency formulas and other research areas. We will give a very brief overview about functional dependencies in order for the reader to get familiar with the notion and to follow the established connection.[3]

Codd's relational model of data ([Dat00]) is the core of relational data bases. The model has many extensions, some of which are obtained through fuzzy logics. In [BV11] several of these fuzzy extensions were studied and compared with a general framework ([BV05a]), called *functional dependencies*. These functional dependencies are of the form $A \Rightarrow B$, where A and B are $\mathbf{L}$-sets of attributes. Such formulas are interpreted in so-called *ranked tables over domains with similarities* ([BV11]). The meaning of the formula is "for any two table rows: Similar values of attributes from A imply similar values of attributes from B".

The data table coincides with a data table of a classical relational model. The fuzzification is obtained through the domain similarities and the ranking. To each pair of values from a given domain, the domain similarities assign them a degree of similarity. The degree to which a row (tuple) of the data table satisfies a query is given by the ranking, value which belongs to the unit interval.

A *(fuzzy) functional dependence* (FD) is defined by $A \Rightarrow B$, where A and B are $\mathbf{L}$-sets of attributes, i.e., $A, B \in \mathbf{L}^Y$. Further, for a ranked data table $\mathcal{D}$, tuples t_1, t_2 and an $\mathbf{L}$-set

[3] Special thanks to Bělohlávek and Vychodil for pointing this out to me.

$C \in \mathbf{L}^Y$ of attributes, the degree $t_1(C) \approx_D t_2(C)$ to which t_1 and t_2 have similar values on the attributes from C is given by

$$t_1(C) \approx_{\mathcal{D}} t_2(C) := (\mathcal{D}(t_1) \otimes \mathcal{D}(t_2)) \to \bigwedge_{y \in Y} (C(y) \to (t_1[y] \approx_y t_2[y])). \tag{6.11}$$

The degree $||A \Rightarrow B||_{\mathcal{D}}$ to which a FD $A \Rightarrow B$ is *true* in $\mathcal{D}$ is defined by

$$||A \Rightarrow B||_{\mathcal{D}} := \bigwedge_{t_1, t_2} ((t_1(A) \approx_{\mathcal{D}} t_2(A))^* \to (t_1(B) \approx_{\mathcal{D}} t_2(B))) \,. \tag{6.12}$$

Then, $t_1(C) \approx_{\mathcal{D}} t_2(C)$ is the truth degree of the statement "if t_1, t_2 are from $\mathcal{D}$, then t_1 and t_2 have similar values on y for each attribute y from C". Further, $||A \Rightarrow B||_{\mathcal{D}}$ stands for the truth degree of "for any tuples t_1, t_2: If t_1 and t_2 have similar values on attributes from A, then t_1 and t_2 have similar values on attributes from B".

The extension of Codd's model by Raju and Majumdar ([RM88]) is said to be one of the most influential one on FDs over domains with similarities. In the literature there are several approaches based on this work. Most of them consider FD formulas $A \Rightarrow B$, where both A and B are crisp. The authors in [BV11] use the globalisation and the identity for the hedge $(-)^*$ in order to show that some methods developed by various authors regarding fuzzy relational data bases are just special cases of their approach.

We are particularly interested in the method presented in [CVC94], where the authors use thresholds to determine the truth value of an FD in a given relational data table $\mathcal{D}$. A FD is of the form $A \Rightarrow_{\alpha,\beta} B$ with $\alpha = (c_y)_{y \in A}$ and $\beta = (c_y)_{y \in B}$. Then, $A \Rightarrow_{\alpha,\beta} B$ is true in $\mathcal{D}$ if

$$\begin{aligned} &\text{if for each } y \in A \text{ we have } t_1[y] \approx_y t_2[y] \geq c_y, \\ &\text{then for each } y \in B \text{ we have } t_1[y] \approx_y t_2[y] \geq c_y. \end{aligned}$$

The following lemma presents the connection between these kind of FDs and the ones defined by [BV11].

Lemma 6.24 ([BV11])**.** *For $(-)^*$ being the globalisation, $L = [0,1]$, and any $\to$, $A \Rightarrow B$ is true in $\mathcal{D}$ according to [CVC94] if and only if $||A_c \Rightarrow B_c||_D = 1$ according to* (6.12)*, where $A_c(y) = c_y$ for $y \in A$ and $A_c(y) = 0$ for $y \notin A$.*

From [BV05a] we know that it is possible to construct a ranked table over domains with similarities $\mathcal{D}$ for each $\mathbf{L}$-context $\mathbb{K}$ such that the stem base of $\mathbb{K}$ coincides with the "stem base" of $\mathcal{D}$ and vice versa. Therefore, using Lemma 6.24, we conclude that the FD's developed in [CVC94] represent a special case of our fuzzy attribute dependencies, provided that the hedge is the globalisation.

6.4. Conclusion

In this chapter we presented a new method of modelling users' preferences in a fuzzy setting. The preferences are expressed by the users in the form of formulas on compound attributes. These allow the users to express their preferences on groups of attributes, i.e., on features that contain more than just one trait. Based on these preferences the users obtain only the formal fuzzy concepts that are "interesting" for them. These concepts form again a complete lattice that represents the further basis of the data analysis.

After investigating some properties of these formulas, we turned our attention to the computation of non-redundant bases. Such methods are useful as the users enter the formulas and it is therefore very likely that these are redundant. Having a set of non-redundant formulas makes it possible for the user to handle them more easily and alter them conveniently. As it turned out, there is a close connection between fAD formulas and attribute implications in a fuzzy setting. Finally, we showed a connection between our framework and some approaches to functional dependencies.

7 Fuzzy-valued Triadic Concept Analysis

The usual way to fuzzify Triadic Concept Analysis is to consider all three components of a triadic concept as **L**-sets. For instance, such an approach was presented in [BO10]. A more general approach was developed in [OK10], where different residuated lattices were considered for the objects, attributes and conditions. A somehow different strategy was considered in [Cla09] using alpha-cuts.

In [Glo11d] we presented a different approach to Triadic Concept Analysis in a fuzzy setting. This work, together with the technical report [Glo11c], represents the basis of this chapter. Our approach differs from the other ones in considering just two components in a triadic concept as fuzzy and one as crisp. This is motivated by the fact that in some situations it is not appropriate to regard all sets as fuzzy. For example, it is not natural to say that "half of a person is old", however we may say "a person is half old".

First, we translate the notions of triadic context, concept and concept lattice to our setting in Section 7.1. Unlike the other works, we generalise all triadic derivation operators. Afterwards, we focus on implications. These have a wide-spread applicability, describe the concept lattice and support knowledge discovery. Thus, it arises as a natural wish to generalise implications to the fuzzy-valued triadic setting. This was done, as it seems, for the first time in [Glo11d], the results of which we present in Section 7.2. At the end, in Setion 7.3, we present another valuable method of Formal Concept Analysis for our fuzzy-valued setting: the conceptual factorisation.

7.1. Context and Concepts

The starting point of each variant of Formal Concept Analysis is the formal context. Therefore, it will be our first concern to introduce fuzzy-valued tricontexts. Afterwards, we turn our attention to the definition of the derivation operators for our setting. The $(-)^{A_k}$-derivation operators can be obtained in various ways, see [Glo11d, Glo11c]. However, here we will present just one of them, which has the advantage of leading to a unified framework. In order to generalise all triadic derivation operators to our setting, we will need the double-scaling of fuzzy-valued tricontexts. This allows us to transform the fuzzy-valued tricontext into a crisp tricontext. The trilattices of the two are isomorphic. Having generalised all triadic derivation operators to the fuzzy-valued setting, we will focus on how they interact with each other. Further, we will investigate some properties of the concepts and present different characterisations of them. We conclude the theory of our fuzzy-valued triadic setting by the *Basic Theorem of Fuzzy-valued Triadic Concept Analysis*.

Note that there are some overlaps between this section and the work presented in [OK10]. The latter contains a more general approach to the properties of fuzzy triadic concepts since it that setting all three components of the concept are fuzzy sets.

Definition 7.1. A **fuzzy-valued triadic context** (shortly **f-valued tricontext**) is a quadruple $\mathbb{K} := (K_1, K_2, K_3, Y)$, where Y is a ternary $\mathbf{L}$-relation between the sets K_i $(i \in \{1,2,3\})$, i.e., $Y\colon K_1 \times K_2 \times K_3 \to L$ and L is the support set of some residuated lattice. The elements of K_1, K_2 and K_3 are called **objects**, **attributes** and **conditions**, respectively. To every triple $(k_1, k_2, k_3) \in K_1 \times K_2 \times K_3$, the $\mathbf{L}$-relation Y assigns a truth value $Y(k_1, k_2, k_3)$ to which the object k_1 has the attribute k_2 under the condition k_3.

An f-valued tricontext can be represented as a three-dimensional table, the entries of which are fuzzy values (see Figure 7.1). These contexts can be understood as follows: They are tricontexts with each condition d-cut $(K_1, K_2, Y^{12}_{k_3})$ (see Definition 3.1, page 68) being an $\mathbf{L}$-context.

Until explicitly said otherwise, $\mathbb{K}$ will denote an f-valued tricontext for the remainder of this chapter.

For our f-valued setting we want to obtain the corresponding $(-)^{A_k}$ and $(-)^{(i)}$-derivation operators, whose crisp variants we have seen in Definition 1.28 and 1.29 (page 19). Since the first two components of an f-valued triconcept will be $\mathbf{L}$-sets and the third one crisp (see Definition 7.2), we lose the symmetry of the derivation operators from the crisp setting. For the $(-)^{A_k}$-derivation operators we distinguish between two cases, namely when A_k is a crisp set and when it is fuzzy. There are various ways of defining these operators, some of which we investigated in [Glo11c]. However, here we will show a unified framework for them.[1] Let $X_i, A_i, A_j \in \mathbf{L}^{K_i}$ $(\{i,j\} = \{1,2\})$ and $X_3, A_3 \subseteq K_3$. We define

$$X_i^{A_3}(k_j) := \bigwedge_{k_i \in K_i} \Big(X_i(k_i) \to Y^{ij}_{A_3}(k_i, k_j) \Big), \tag{7.1}$$

$$X_i^{A_j}(k_3) := \bigwedge_{k_i \in K_i} \Big(X_i(k_i) \to Y^{i3}_{A_j}(k_i, k_3) \Big)^*, \tag{7.2}$$

$$X_3^{A_i}(k_j) := \bigwedge_{k_3 \in K_3} \Big(X_3(k_3) \to Y^{j3}_{A_i}(k_j, k_3) \Big), \tag{7.3}$$

where $(-)^*$ is the globalisation in order to ensure that $X_i^{A_j}$ is a crisp set and

$$Y^{ij}_{A_k}(k_i, k_j) := \bigwedge_{k_k \in K_k} \left(A_k(k_k) \to Y(k_i, k_j, k_k) \right) \tag{7.4}$$

[1] A similar approach, developed independently, was presented in [BO10].

for $\{i,j,k\} = \{1,2,3\}$. We obtain the dyadic $\mathbf{L}$-context $\mathbb{K}^{ij}_{A_k} := (K_i, K_j, Y^{ij}_{A_k})$ with its derivation operators given by (7.1), (7.2) and (7.3).

Having the first kind of fuzzy-valued triadic derivation operators, we may now define what a formal concept in this setting is:

Definition 7.2. A **fuzzy-valued triadic concept** (shortly **f-valued triconcept**) of an f-valued tricontext (K_1, K_2, K_3, Y) is a triple (A_1, A_2, A_3) with $A_1 \in \mathbf{L}^{K_1}$, $A_2 \in \mathbf{L}^{K_2}$ and $A_3 \subseteq K_3$ such that $A_i^{A_j} = A_k$ for all $\{i,j,k\} = \{1,2,3\}$. The components A_1, A_2 and A_3 are called the (**f-valued**) **extent**, the (**f-valued**) **intent**, and the **modus** of (A_1, A_2, A_3), respectively. We denote by $\mathfrak{T}(\mathbb{K})$ the set of all f-valued triconcepts.

An f-valued triconcept (A_1, A_2, A_3) can be understood as follows: The tuple (A_1, A_2) is an $\mathbf{L}$-concept and A_3 contains all the conditions under which this $\mathbf{L}$-concept exists.

Let us illustrate the so-far introduced notions by an example.

Example 7.3. Figure 7.1 shows an f-valued tricontext with values from the 3-element chain $\{0, 0.5, 1\}$. The object set $K_1 = \{S_1, S_2, S_3, S_4, S_5\}$ contains 5 groups of students, the attribute set $K_2 = \{f, s, v\}$ contains 3 feelings, namely fevered (f), serious (s), vigilant (v), and the condition set $K_3 = \{E, P, F\}$ contains the events: doing an exam (E), giving a presentation (P) and meeting friends (F). Using the Łukasiewicz logic, we obtain 28 f-valued

	E			P			F		
	f	s	v	f	s	v	f	s	v
S_1	1	1	1	1	0.5	0.5	0	0.5	1
S_2	1	0.5	1	0.5	0	0	0	0	0.5
S_3	0.5	0.5	0.5	0.5	0.5	0	0	0	0.5
S_4	0.5	0	0.5	0.5	0.5	0.5	0	0.5	0.5
S_5	1	1	1	1	0.5	0.5	0	0.5	1

Figure 7.1.: F-valued tricontext

triconcepts and with the Gödel logic 34. For example, $(\{S_1, S_2, S_3, S_4, S_5\}, \{{}^{0.5}/v\}, \{E, F\})$ is an f-valued triconcept meaning that all students are partially vigilant while doing an exam and meeting their friends. Another example is $(\{S_1, S_2, {}^{0.5}/S_3, S_5\}, \{f, {}^{0.5}/s, v\}, \{E\})$ meaning that the first, second and fifth group of students are fevered, vigilant and partially serious while doing an exam, whereas this description applies just partially to the third group of students.

The set $\mathfrak{T}(\mathbb{K})$ of all f-valued triconcepts of $\mathbb{K} = (K_1, K_2, K_3, Y)$ is structured by the crisp quasiorders $\lesssim_i$ and their corresponding equivalence relations $\sim_i$ defined by

$$(A_1, A_2, A_3) \lesssim_i (B_1, B_2, B_3) :\Longleftrightarrow A_i \subseteq B_i \text{ and} \tag{7.5}$$

$$(A_1, A_2, A_3) \sim_i (B_1, B_2, B_3) :\Longleftrightarrow A_i = B_i \ (i = 1,2,3). \tag{7.6}$$

Note that for $i \in \{1,2\}$ we are dealing with fuzzy subsethood, i.e., $A_i, B_i \in \mathbf{L}^{K_i}$ and $A_i \subseteq B_i$ means $\mathrm{S}(A_i, B_i) = 1$, whereas for $\lesssim_3$ we have $A_3, B_3 \subseteq K_3$ and the crisp subsethood is considered. Analogous remarks hold for the equivalence relations $\sim_i$. By $[(A_1,A_2,A_3)]_i$ we denote the equivalence class of $\sim_i$ containing the f-valued triconcept (A_1,A_2,A_3). The quasiorder $\lesssim_i$ induces an order $\leq_i$ on the factor set $\mathfrak{T}(\mathbb{K})/\sim_i$ of all equivalence classes of $\sim_i$ which is characterised by

$$[(A_1,A_2,A_3)]_i \leq_i [(B_1,B_2,B_3)]_i \iff A_i \subseteq B_i.$$

Note that in [BO10] the authors also considered **L**-quasiorders and **L**-equivalence relations. However, for the structure of the f-valued trilattice, our crisp quasiorders are sufficient.

As we will see later, $\underline{\mathfrak{T}}(\mathbb{K}) := (\mathfrak{T}(\mathbb{K}),\lesssim_1,\lesssim_2,\lesssim_3)$ is the fuzzy-valued counterpart of the trilattice from Triadic Concept Analysis. But before we come to this matter, let us first investigate the connection between f-valued tricontexts and crisp tricontexts.

According to [Pol97] (see also Theorem 1.37, page 26), we may transform an **L**-context through double-scaling into a crisp context, the concept lattice of which is isomorphic to the concept lattice of the **L**-context. By double-scaling each condition d-cut we obtain the corresponding double-scaled triadic crisp context $\widetilde{\mathbb{K}}$ for an f-valued tricontext $\mathbb{K}$. We will now present the construction of $\widetilde{\mathbb{K}} := (K_1^{\square},K_2^{\square},K_3,\widetilde{Y})$, the double-scaled tricontext, for a given f-valued tricontext $\mathbb{K} = (K_1,K_2,K_3,Y)$. Let $X_i \in \mathbf{L}^{K_i}$ and $Z_i \subseteq K_i \times L$ with $i \in \{1,2\}$ and let L be the support set of some residuated lattice. We define

$$\begin{aligned} X_i^{\square} &:= \{(k_i,\nu) \mid k_i \in K_i, \nu \in L, \nu \leq X_i(k_i)\} \subseteq K_i^{\square} := K_i \times L,\\ Z_i^{\diamond}(k_i) &:= \bigvee\{\nu \mid (k_i,\nu) \in Z_i\} \in \mathbf{L}^{K_i} \end{aligned}$$

for each $k_i \in K_i$. Then, $\widetilde{Y} \subseteq K_1^{\square} \times K_2^{\square} \times K_3$ is a crisp ternary relation defined by

$$((k_1,\nu),(k_2,\lambda),k_3) \in \widetilde{Y} :\iff \nu \otimes \lambda \leq Y(k_1,k_2,k_3).$$

Theorem 7.4. $\mathfrak{T}(\mathbb{K}) \cong \mathfrak{T}(\widetilde{\mathbb{K}})$.

Proof. For the proof we will use the adjointness property (1.2) (page 6) and the properties of residuated lattices given on page 7.

Let $\widetilde{\mathbb{K}} := (K_1^{\square},K_2^{\square},K_3,\widetilde{Y})$ be the double-scaled tricontext of the f-valued tricontext $\mathbb{K} := (K_1,K_2,K_3,Y)$. We will show that an isomorphism is given by

$$\varphi : \underline{\mathfrak{T}}(\mathbb{K}) \to \underline{\mathfrak{T}}(\widetilde{\mathbb{K}}) \text{ with } \varphi(A_1,A_2,A_3) := (A_1^{\square},A_2^{\square},A_3)$$

by proving that its inverse is given by

$$\psi : \underline{\mathfrak{T}}(\widetilde{\mathbb{K}}) \to \underline{\mathfrak{T}}(\mathbb{K}) \text{ with } \psi(X_1,X_2,X_3) := (X_1^{\diamond},X_2^{\diamond},X_3).$$

In order to do so, we need to show the following: For all f-valued triconcepts (A_1,A_2,A_3), $(B_1,B_2,B_3) \in \underline{\mathfrak{T}}(\mathbb{K})$ and for all (crisp) triconcepts $(X_1,X_2,X_3) \in \underline{\mathfrak{T}}(\widetilde{\mathbb{K}})$ we have

$$\varphi(A_1,A_2,A_3) \in \underline{\mathfrak{T}}(\widetilde{\mathbb{K}}), \tag{7.7}$$

$$\psi(X_1,X_2,X_3) \in \underline{\mathfrak{T}}(\mathbb{K}), \tag{7.8}$$

$$\psi\varphi(A_1,A_2,A_3) = (A_1,A_2,A_3), \tag{7.9}$$

$$\varphi\psi(X_1,X_2,X_3) = (X_1,X_2,X_3), \tag{7.10}$$

$$(A_1,A_2,A_3) \lesssim_i (B_1,B_2,B_3) \iff \varphi(A_1,A_2,A_3) \lesssim_i \varphi(B_1,B_2,B_3) \tag{7.11}$$

for all $i \in \{1,2,3\}$.[2] For statement (7.7) we have

$$\begin{aligned}
(A_1^{\Box})^{A_3} &= \left\{(k_2,\lambda) \in K_2^{\Box} \mid \forall((k_1,\nu),k_3) \in A_1^{\Box} \times A_3 : ((k_1,\nu),(k_2,\lambda),k_3) \in \widetilde{Y}\right\} \\
&= \left\{(k_2,\lambda) \in K_2^{\Box} \mid \forall k_1 \in K_1, \forall \nu \leq A_1(k_1) : \lambda \otimes \nu \leq Y^{12}_{A_3}(k_1,k_2)\right\} \\
&= \left\{(k_2,\lambda) \in K_2^{\Box} \mid \forall k_1 \in K_1 : \lambda \otimes A_1(k_1) \leq Y^{12}_{A_3}(k_1,k_2)\right\} \\
&= \left\{(k_2,\lambda) \in K_2^{\Box} \mid \lambda \leq \bigwedge_{k_1 \in K_1} \left(A_1(k_1) \to Y^{12}_{A_3}(k_1,k_2)\right)\right\} && \text{(by (1.2))} \\
&= \left\{(k_2,\lambda) \in K_2^{\Box} \mid k_2 \in K_2,\ \lambda \leq A_1^{A_3}(k_2)\right\} \\
&= \left(A_1^{A_3}\right)^{\Box} \\
&= A_2^{\Box}.
\end{aligned}$$

One proceeds similarly for $(A_2^{\Box})^{A_3} = A_1^{\Box}$, $(A_i^{\Box})^{A_j^{\Box}} = A_3$ and $A_3^{A_i^{\Box}} = A_j^{\Box}$ for all $\{i,j\} = \{1,2\}$. Thus, we have $(A_1^{\Box}, A_2^{\Box}, A_3) \in \underline{\mathfrak{T}}(\widetilde{\mathbb{K}})$.

For statement (7.8) we have

$$\begin{aligned}
(X_1^{\Diamond})^{X_3}(k_2) &= \bigwedge_{k_1 \in K_1} \left(X_1^{\Diamond}(k_1) \to Y^{12}_{X_3}(k_1,k_2)\right) \\
&= \bigwedge_{k_1 \in K_1} \left(\bigvee\{\nu \mid (k_1,\nu) \in X_1\} \to Y^{12}_{X_3}(k_1,k_2)\right) \\
&= \bigwedge \left\{\nu \to Y^{12}_{X_3}(k_1,k_2) \mid (k_1,\nu) \in X_1\right\} && \text{(by (1.7))} \\
&= \bigvee \left\{\lambda \in L \mid \forall (k_1,\nu) \in X_1 : \lambda \leq \nu \to Y^{12}_{X_3}(k_1,k_2)\right\} \\
&= \bigvee \left\{\lambda \in L \mid \forall((k_1,\nu),k_3) \in X_1 \times X_3 : ((k_1,\nu),(k_2,\lambda),k_3) \in \widetilde{Y}\right\} && \text{(by (1.2))} \\
&= \bigvee \left\{\lambda \in L \mid (k_2,\lambda) \in X_1^{X_3}\right\} \\
&= (X_1^{X_3})^{\Diamond}(k_2)
\end{aligned}$$

for all $k_2 \in K_2$. Therefore, $(X_1^{\Diamond})^{X_3} = (X_1^{X_3})^{\Diamond} = X_2^{\Diamond}$. Analogously, one can show that $(X_2^{\Diamond})^{X_3} = X_1^{\Diamond}$, $(X_i^{\Diamond})^{X_j^{\Diamond}} = X_3$ and $X_3^{X_i^{\Diamond}} = X_j^{\Diamond}$ for all $\{i,j\} = \{1,2\}$. Thus, we have $(X_1^{\Diamond}, X_2^{\Diamond}, X_3) \in \underline{\mathfrak{T}}(\mathbb{K})$.

For statement (7.9) we have

$$\begin{aligned}
A_1^{\Box\Diamond}(k_1) &= \bigvee \left\{\nu \in L \mid (k_1,\nu) \in A_1^{\Box}\right\} \\
&= \bigvee \left\{\nu \in L \mid \nu \leq A_1(k_1)\right\} \\
&= A_1(k_1)
\end{aligned}$$

for all $k_1 \in K_1$. Thus, $A_1^{\Box\Diamond} = A_1$ holds and analogously $A_2^{\Box\Diamond} = A_2$. Hence, we have $\psi\varphi(A_1,A_2,A_3) = (A_1,A_2,A_3)$.

Statement (7.10): For all triconcepts $(X_1,X_2,X_3) \in \underline{\mathfrak{T}}(\widetilde{\mathbb{K}})$ it holds that $X_i^{X_j} = X_k$ for

[2] By abuse of notation we use $\lesssim_i$ both for the quasiorders on the f-valued triconcepts as well as on the crisp triconcepts.

all $\{i,j,k\} = \{1,2,3\}$, and therefore we have

$$\begin{aligned}
X_1^{\diamond}(k_1) &= \left(X_2^{X_3}\right)^{\diamond}(k_1) \\
&= \bigvee \left\{\nu \mid (k_1,\nu) \in X_2^{X_3}\right\} \\
&= \bigvee \left\{\nu \in L \mid \forall (k_2,\lambda) \in X_2 : ((k_1,\nu),(k_2,\lambda)) \in \widetilde{Y}_{X_3}^{12}\right\} \\
&= \bigvee \left\{\nu \in L \mid \forall (k_2,\lambda) \in X_2 : \nu \leq \lambda \to Y_{X_3}^{12}(k_1,k_2)\right\} \\
&= \bigwedge_{(k_2,\lambda)\in X_2} \left(\lambda \to Y_{X_3}^{12}(k_1,k_2)\right)
\end{aligned}$$

for all $k_1 \in K_1$. Hence,

$$\begin{aligned}
X_1^{\diamond\square} &= \left\{(k_1,\nu) \mid k_1 \in K_1, \nu \in L, \nu \leq X_1^{\diamond}(k_1)\right\} \\
&= \left\{(k_1,\nu) \mid \forall (k_2,\lambda) \in X_2 : k_1 \in K_1, \nu \in L, \nu \leq \lambda \to Y_{X_3}^{12}(k_1,k_2)\right\} \\
&= \left\{(k_1,\nu) \mid \forall (k_2,\lambda) \in X_2 : ((k_1,\nu),(k_2,\lambda)) \in \widetilde{Y}_{X_3}^{12}\right\} \\
&= X_2^{X_3} = X_1.
\end{aligned}$$

Thus, $X_1^{\diamond\square} = X_1$ holds and analogously $X_2^{\diamond\square} = X_2$, i.e., $\varphi\psi(X_1,X_2,X_3) = (X_1,X_2,X_3)$.

For the last item (7.11), let $(A_1,A_2,A_3),(B_1,B_2,B_3) \in \underline{\mathfrak{T}}(\mathbb{K})$. We have

$$\begin{aligned}
& (A_1,A_2,A_3) \lesssim_1 (B_1,B_2,B_3) \\
\iff\; & A_1 \subseteq B_1 \\
\iff\; & A_1(k_1) \leq B_1(k_1) \text{ for all } k_1 \in K_1 \\
\iff\; & \{(k_1,\nu) \mid k_1 \in K_1, \nu \in L, \nu \leq A_1(k_1)\} \subseteq \{(k_1,\nu) \mid k_1 \in K_1, \nu \in L, \nu \leq B_1(k_1)\} \\
\iff\; & A_1^{\square} \subseteq B_1^{\square} \\
\iff\; & \varphi(A_1,A_2,A_3) \lesssim_1 \varphi(B_1,B_2,B_3),
\end{aligned}$$

where the last equivalence is due to (7.7). One proceeds analogously for $\lesssim_2$, and the proof for $\lesssim_3$ is straightforward as A_3 remains a crisp set. □

This theorem allows the proper generalisation of the $(-)^{(i)}$-derivation operators into our setting, which seems to be the first generalisation of these operators into the fuzzy setting. And finally, the most important consequence is the fuzzy-valued triadic version of the Basic Theorem of Triadic Concept Analysis (Theorem 7.8) which we will present at the end of this section.

Let us now turn our attention to the generalisation of the $(-)^{(i)}$-derivation operators to our setting. As in the case of the $(-)^{A_k}$-derivation operators we will distinguish between different cases for the $(-)^{(i)}$-derivation operators. In case of the **$(-)^{(i)}$-derivation operators** with $Z = X_j \times X_3 \subseteq \mathbf{L}^{K_j} \times K_3$ and $X_i \in \mathbf{L}^{K_i}$ for $\{i,j\} = \{1,2\}$, the situation is easy. They are defined by

$$Z^{(i)}(k_i) := \bigwedge_{k_j \in K_j, k_3 \in K_3} \left(Z(k_j,k_3) \to Y^{(i)}(k_i,k_j,k_3)\right),$$

$$X_i^{(i)}(k_j,k_3) := \left(\bigwedge_{k_i \in K_i} (X_i(k_i) \to Y^{(i)}(k_i,k_j,k_3)), k_3\right),$$

where $Z(k_j, k_3) := X_j(k_j) \cdot X_3(k_3)$. These derivation operators correspond to the derivation operators of the dyadic $\mathbf{L}$-contexts defined by

$$\mathbb{K}^{(i)} := (K_i, K_j \times K_3, Y^{(i)}) \text{ with } Y^{(i)}(k_i, k_j, k_3) := \begin{cases} Y(k_i, k_j, k_3), & i < j, \\ Y(k_j, k_i, k_3), & i > j. \end{cases}$$

The $(-)^{(3)}$**-derivation operator** for $Z := X_1 \times X_2 \in \mathbf{L}^{K_1} \times \mathbf{L}^{K_2}$ and $X_3 \subseteq K_3$ is defined by

$$Z^{(3)} := \{k_3 \in K_3 \mid \forall (k_1, k_2) \in Z : Z(k_1, k_2) \leq Y(k_1, k_2, k_3)\} \tag{7.12}$$

where $Z(k_1, k_2) := X_1(k_1) \otimes X_2(k_2)$. The corresponding dyadic $\mathbf{L}$-context is given by $\mathbb{K}^{(3)} := (K_1 \times K_2, K_3, Y^{(3)})$ with the $\mathbf{L}$-relation $Y^{(3)}(k_1, k_2, k_3) := Y(k_1, k_2, k_3)$. One can easily show that

$$Z^{(3)}(k_3) = \bigwedge_{(k_1,k_2) \in K_1 \times K_2} \left(Z(k_1, k_2) \to Y^{(3)}(k_1, k_2, k_3)\right)^*,$$

where $(-)^*$ is the globalisation in order to ensure that $Z^{(3)}$ is crisp. We basically search for the condition d-cuts which contain the maximal rectangle generated by Z.

The situation for $X_3^{(3)}$ is quite tricky. Applying the derivation operators in $\mathbb{K}^{(3)}$ for X_3, we get a truth value $l \in L$ such that $l = l_1 \otimes l_2$ instead of a tuple $({}^{l_1}/k_1, {}^{l_2}/k_2)$ consisting of $\mathbf{L}$-sets. To obtain such a tuple, first we have to compute the double-scaled context $\widetilde{\mathbb{K}}$. Afterwards, we use the crisp $(-)^{(3)}$-derivation operator in $\widetilde{\mathbb{K}}$ to find the components of the triconcept. Finally, we transform these into $\mathbf{L}$-sets as described in the construction of $\widetilde{\mathbb{K}}$. This way, we obtain tuples of the form $({}^{l_1}/k_1, {}^{l_2}/k_2)$ consisting of object and attribute $\mathbf{L}$-sets instead of the truth value $l_1 \otimes l_2$. The operator is well-defined due to the isomorphism between $\mathbb{K}$ and $\widetilde{\mathbb{K}}$.

Having generalised all derivation operators from the crisp case into our setting, we may now investigate the interplay between them. Further, we will study some characterisations and properties of f-valued triconcepts, the fuzzy-valued counterparts of those presented in [Wil95].

Proposition 7.5. *For all f-valued triconcepts (A_1, A_2, A_3) and (B_1, B_2, B_3) of $\mathbb{K}$ and for $\{i, j, k\} = \{1, 2, 3\}$ we have*

$$(A_1, A_2, A_3) \lesssim_i (B_1, B_2, B_3) \text{ and } (A_1, A_2, A_3) \lesssim_j (B_1, B_2, B_3)$$
$$\text{imply } (A_1, A_2, A_3) \gtrsim_k (B_1, B_2, B_3).$$

Further, $\sim_i \cap \sim_j$ is the identity on $\mathfrak{T}(\mathbb{K})$ whenever $i \neq j$.

Proof. From $A_i \subseteq B_i$ and $A_j \subseteq B_j$ it follows that $A_k = (A_i \times A_j)^{(k)} \supseteq (B_i \times B_j)^{(k)} = B_k$, showing the first part of the proposition. The definition of f-valued triconcepts implies that two components of an f-valued triconcept uniquely determine its third component. This, on the other hand, shows the second part of the proposition. □

Proposition 7.6. *Let $(A_1, A_2, A_3) \in \underline{\mathfrak{T}}(\mathbb{K})$ be an f-valued triconcept. Then, the following hold:*

i) (A_1, A_2, A_3) *is maximal with respect to component-wise set inclusion;*

ii) $A_i = (A_j \times A_k)^{(i)}$ *for* $\{i,j,k\} = \{1,2,3\}$ *with* $j < k$.

Proof. i) Let $(A_1, A_2, A_3) \in \underline{\mathfrak{T}}(\mathbb{K})$. Further let $(X_1, X_2, X_3) \in \mathbf{L}^{K_1} \times \mathbf{L}^{K_2} \times \mathfrak{P}(K_3)$ such that $A_i \subseteq X_i$ for all $i \in \{1,2,3\}$. Thus, if $X_1 \times X_2 \times X_3 \subseteq Y$, then $X_i \subseteq (A_j \times A_k)^{(i)} = A_i$ for all $\{i,j,k\} = \{1,2,3\}$. Hence, we have $(A_1, A_2, A_3) = (X_1, X_2, X_3)$.
ii) For the proof of this item we use some of the properties of residuated lattices listed on page 7. Let (A_1, A_2, A_3) be an f-valued triconcept. Then, we have

$$\begin{aligned}
A_1^{A_3}(k_2) &= \bigwedge_{k_1 \in K_1} \left(A_1(k_1) \to Y_{A_3}^{12}(k_1,k_2)\right) \\
&= \bigwedge_{k_1 \in K_1} \left(A_1(k_1) \to \bigwedge_{k_3 \in K_3} (A_3(k_3) \to Y(k_1,k_2,k_3))\right) \\
&= \bigwedge_{k_1 \in K_1} \bigwedge_{k_3 \in K_3} (A_1(k_1) \to (A_3(k_3) \to Y(k_1,k_2,k_3))) && \text{(by (1.5))} \\
&= \bigwedge_{k_1 \in K_1, k_3 \in K_3} ((A_1(k_1) \otimes A_3(k_3)) \to Y(k_1,k_2,k_3)) && \text{(by (1.3))} \\
&= (A_1 \times A_3)^{(2)}(k_2)
\end{aligned}$$

for all $k_2 \in K_2$. Thus, $A_2 = A_1^{A_3} = (A_1 \times A_3)^{(2)}$. In a similar way we obtain $A_1 = (A_2 \times A_3)^{(1)}$. For the $(-)^{(3)}$-derivation operator the situation is straightforward because we obtain it through the double-scaled context. Hence, for all $\{i,j,k\} = \{1,2,3\}$ with $j < k$ we have $A_i = (A_j \times A_k)^{(i)}$. □

Proposition 7.7. *For* $\{i,j,k\} = \{1,2,3\}$ *let there be* $\mathbf{L}$*-sets* $X_i \in \mathbf{L}^{K_i}$ *(crisp set* $X_i \subseteq K_i$*, if* $i = 3$*) and* $X_k \in \mathbf{L}^{K_k}$ *(crisp set* $X_k \subseteq K_k$*, if* $k = 3$*) such that* $A_j := X_i^{X_k}$*,* $A_i := A_j^{X_k}$ *and* $A_k := (A_i \times A_j)^{(k)}$ *(if* $i < j$*) or* $A_k := (A_j \times A_i)^{(k)}$ *(if* $i > j$*). Then,* (A_1, A_2, A_3) *is an f-valued triconcept denoted by* $\mathfrak{b}_{ik}(X_i, X_k)$ *having the smallest* k*-th component under all f-valued triconcepts* (B_1, B_2, B_3) *with the largest* j*-th component satisfying* $X_i \subseteq B_i$ *and* $X_k \subseteq B_k$*. Particularly,* $\mathfrak{b}_{ik}(A_i, A_k) = (A_1, A_2, A_3)$ *for each f-valued triconcept* (A_1, A_2, A_3) *of* $\mathbb{K}$*.*

Proof. Without loss of generality we can assume $(i,j,k) = (1,2,3)$. Obviously, $X_1 \subseteq A_1$ and $X_3 \subseteq A_3$. We start by proving that (A_1, A_2, A_3) is indeed an f-valued triconcept. From Proposition 7.6 we know that $A_3 = (A_1 \times A_2)^{(3)} = A_1^{A_2}$. Hence, $A_2 \subseteq A_1^{A_3} \subseteq X_1^{X_3} = A_2$ and $A_2 = A_1^{A_3}$. Similarly, we obtain $A_1 = A_2^{A_3}$. Therefore, (A_1, A_2, A_3) satisfies the conditions from the definition of an f-valued triconcept.

For the second statement of the proposition, let $(B_1, B_2, B_3) \in \mathfrak{T}(\mathbb{K})$ with $X_1 \subseteq B_1$ and $X_3 \subseteq B_3$. Then, $B_2 \subseteq A_2$, because $B_2 = B_1^{B_3} \subseteq X_1^{X_3} = A_2$. If $B_2 = A_2$, then by similar considerations as before, we obtain $B_1 \subseteq A_1$. Hence, we have $A_3 = A_1^{A_2} \subseteq B_1^{B_2} = B_3$, which proves the first part of the statement. Now, if (A_1, A_2, A_3) is an f-valued triconcept, then $A_1^{A_3} = A_2$ and $A_2^{A_3} = A_1$. Therefore, $\mathfrak{b}_{ik}(A_1, A_3) = (A_1, A_2, A_3)$ follows by the first part of the proposition. □

Now we may present the fuzzy-valued version of the Basic Theorem of Triadic Concept Analysis (Theorem 1.31).

Theorem 7.8. *Let* $\mathbb{K} = (K_1, K_2, K_3, Y)$ *be an f-valued tricontext. Then,* $\underline{\mathfrak{T}}(\mathbb{K})$ *is a complete trilattice of* $\mathbb{K}$ *with the ik-joins for two sets* $\mathcal{X}_i$ *and* $\mathcal{X}_k$ *of f-valued triconcepts given by*

$$\nabla_{ik}(\mathcal{X}_i, \mathcal{X}_k) := \mathfrak{b}_{ik}(\bigcup\{A_i \mid (A_1, A_2, A_3) \in \mathcal{X}_i\}, \bigcup\{A_k \mid (A_1, A_2, A_3) \in \mathcal{X}_k\})$$

for $\{i, j, k\} = \{1, 2, 3\}$*. In general, every complete trilattice* $\underline{V} := (V, \lesssim_1, \lesssim_2, \lesssim_3)$ *is isomorphic to* $\underline{\mathfrak{T}}(\mathbb{K})$ *if and only if there exist mappings* $\widetilde{\kappa}_i : K_i \times L \to \mathcal{F}_i(\underline{V})$ *and* $\widetilde{\kappa}_3 : K_3 \to \mathcal{F}_3(\underline{V})$ *such that* $\widetilde{\kappa}_i(K_i \times L)$ *and* $\widetilde{\kappa}_3(K_3)$ *are i-dense* $(i = 1, 2)$ *and* 3*-dense with respect to* $\underline{V}$*, respectively, and*

$$(A_1 \bigotimes A_2) \times A_3 \subseteq Y \Longleftrightarrow \bigcap_{a_1 \in K_1} \widetilde{\kappa}_1(a_1, A_1(a_1)) \cap \bigcap_{a_2 \in A_2} \widetilde{\kappa}_1(a_2, A_2(a_2)) \cap \bigcap_{a_3 \in A_3} \widetilde{\kappa}_1(a_3) \neq \varnothing$$

for all $A_1 \in \mathbf{L}^{K_1}, A_2 \in \mathbf{L}^{K_2}$ *and* $A_3 \subseteq K_3$*.*

Proof. The proof is straightforward by applying the isomorphism between the f-valued tricontext and its double-scaled context given by Theorem 7.4 (page 142). ☐

Example 7.9. To keep things manageable we will draw the complete f-valued trilattice of a smaller f-valued tricontext, namely of the following:

	A		B	
	a	b	a	b
x	1	0	1	1
y	1	0.5	1	0.5
z	0.5	0	0	1

With the Gödel logic the f-valued tricontext has 11 f-valued triconcepts, which can be seen in the complete f-valued trilattice displayed in Figure 7.2. The diagram can be read in an analogous way to its crisp counterpart. See the text next to Figure 1.4 (page 23).

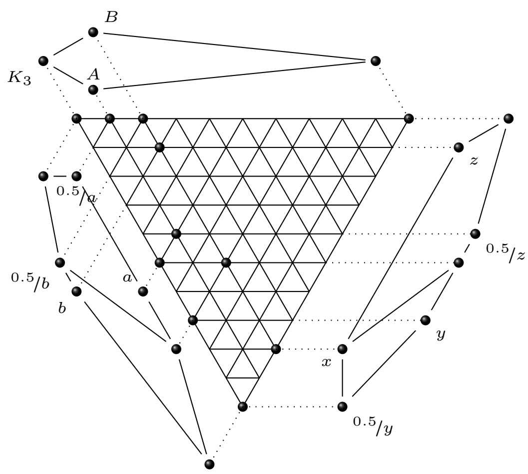

Figure 7.2.: F-valued trilattice

7.2. Implications

In this section we study f-valued implications as generalisations of those elaborated for the crisp case in [GO04]. There, the authors presented various triadic implications, which are stronger than the ones developed in [Bie88]. For a given crisp tricontext $\mathbb{K} = (K_1, K_2, K_3, Y)$ and for $A, B \subseteq K_2$ and $X \subseteq K_3$, the expression $A \xrightarrow{X} B$ is called *conditional attribute implication*. Further, for $A, B \subseteq K_3$ and $X \subseteq K_2$, the expression $A \xrightarrow{X} B$ is called *attributional condition implication*. Implications of the form $A \to B$ with $A, B \subseteq K_2 \times K_3$ are called *attribute* $\times$ *condition implications*. Our main aim in the upcoming subsections is to generalise such implications to our setting. This was done, as it seems, for the first time in [Glo11d].

7.2.1. F-valued Conditional Attribute vs. Attributional Condition Implications

In this subsection we study two kinds of implications in a tricontext that we introduce formally in a moment. For our running example from Figure 7.1 (page 141) they look as follows:

- *If we are moderately vigilant during an exam, then we are also fevered.*
- *If we are serious during an exam, then we feel the same during our presentation.*

We start with the first kind of implications, the so-called *f-valued conditional attribute implications*. After showing how such implications can be computed, we present a method, based on [GO04], to describe all such implications. Since we may arbitrarily interchange the roles of objects, attributes and conditions in a tricontext, we come to the second kind of implications. For both families of implications we also show different ways of handling them, and we illustrate the notions on our running example from Figure 7.1.

Definition 7.10. Let $\mathbb{K} = (K_1, K_2, K_3, Y)$ be an f-valued tricontext. For $A, B \in \mathbf{L}^{K_2}$ and $X \subseteq K_3$ the truth value $||A \overset{X}{\Rightarrow} B||_{\mathbb{K}}$ in $\mathbb{K}$ of the expression $A \overset{X}{\Rightarrow} B$ is given by

$$||A \overset{X}{\Rightarrow} B||_{\mathbb{K}} := ||A \Rightarrow B||_{\mathbb{K}_X^{12}}.$$

Such an implication is called an **f-valued conditional attribute implication**.

Note that these implications are ordinary fuzzy implications since we are working in the $\mathbf{L}$-context $\mathbb{K}_X^{12}$. Therefore, using the theory presented in Subsection 1.5.1, we may further characterise these implications by

$$\begin{aligned} ||A \overset{X}{\Rightarrow} B||_{\mathbb{K}} &= \mathrm{S}(A, \mathrm{Int}(\mathbb{K}_X^{12}))^* \to \mathrm{S}(B, \mathrm{Int}(\mathbb{K}_X^{12})) \\ &= \mathrm{S}(B, A^{XX}), \end{aligned}$$

where $\mathrm{Int}(\mathbb{K}_X^{12})$ denotes the set of all intents of $\mathbb{K}_X^{12}$.

Example 7.11. Consider the f-valued tricontext $\mathbb{K}$ given in Figure 7.1 (page 141) with the Gödel logic and the identity as hedge. We have for instance the f-valued conditional attribute implications $||{}^{0.5}/s \overset{E}{\Rightarrow} f||_{\mathbb{K}} = ||{}^{0.5}/s \overset{P}{\Rightarrow} f||_{\mathbb{K}} = 0.5$. The first implication means that whenever the students are partially serious during an exam, they might also be fevered. The same holds for this implication during a presentation given by the students. The implication does not hold (holds with truth value zero) when they are meeting their friends. Another implication which does not hold is $\|v \overset{P}{\Rightarrow} {}^{0.5}/s\|_{\mathbb{K}}$. However, if we replace the identity hedge on the objects with the globalisation, then we obtain $\|v \overset{P}{\Rightarrow} {}^{0.5}/s\|_{\mathbb{K}} = 1$.

For an f-valued tricontext $\mathbb{K}$ denote by $\mathrm{Imp}(K_2)$ the set of all fuzzy implications on K_2, i.e.,

$$\mathrm{Imp}(K_2) := \{A \Rightarrow B \mid A, B \in \mathbf{L}^{K_2}\}.$$

We construct the dyadic context

$$\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K}) := (\mathrm{Imp}(K_2), K_3, I)$$

with $I(A \Rightarrow B, x) := \|A \overset{x}{\Rightarrow} B\|_{\mathbb{K}}$ where $A \Rightarrow B \in \mathrm{Imp}(K_2)$ and $x \in K_3$. It is sufficient to consider only implications of the form $A \Rightarrow m$ where $m \in \mathbf{L}^M \setminus A$. In $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ we may use the derivation operators for dyadic $\mathbf{L}$-contexts given by (1.16) and (1.17) (page 24) to obtain the $\mathbf{L}$-concepts of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$. Then, $(C, D) \in \underline{\mathfrak{B}}(\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K}))$ contains in its extent all the implications that hold under all conditions from D, i.e., the implications from (K_1, K_2, Y_d^{12}) with $d \in D$. In the crisp case, each extent is an implicational theory, i.e., the set of all implications of some formal context. Hence, every extent has a stem base. Thus, the implicational theories in $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ are ordered by the conditions under which they hold. In order to ensure in our setting that each extent has a stem base, we need to use the globalisation for $(-)^{*_G}$ and may use an arbitrary hedge for $(-)^{*_M}$, as discussed in Subsection 1.5.2.

In accordance with the idea presented in [GO04] we label the $\mathbf{L}$-lattice of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ as follows: The attribute labelling is done in the usual way. For the object labelling the situation is more cumbersome. Each set of implications from $\mathrm{Imp}(K_2)$ generates an extent of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ and an implicational theory, where the latter is always contained in the first. The object labels shall be distributed such that every extent is generated as an implicational theory by the labels attached to it and to its subconcepts. Therefore, the bottom element of the lattice will contain the stem base of all f-valued conditional attribute implications. The other nodes are labelled with the $\mathcal{L}$-stem base relative to the union of the extents below it. We have studied such relative stem bases in Section 5.3. By labelling the lattice in the usual way we would overload the diagram. Another variant would be to clarify the context. However, afterwards it would be difficult to reconstruct the initial extents.

The $\mathbf{L}$-lattice of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ is displayed in Figure 7.3. For instance, the implication $\{s, {}^{0.5}/v\} \Rightarrow v^3$ at the bottom of the lattice means that whenever the students are serious and partially vigilant they are also vigilant during any of the three events. The implication ${}^{0.5}/v \Rightarrow {}^{0.5}/f$ on the left side means that whenever the students are partially vigilant during a presentation or an exam, they are also partially fevered during the same event. Yet another example is the top most implication $\{f, {}^{0.5}/s\} \Rightarrow v$ on the right side of the lattice. It means that whenever the students are fevered and partially serious while giving

[3] Once again we sometimes omit curly brackets around singletones.

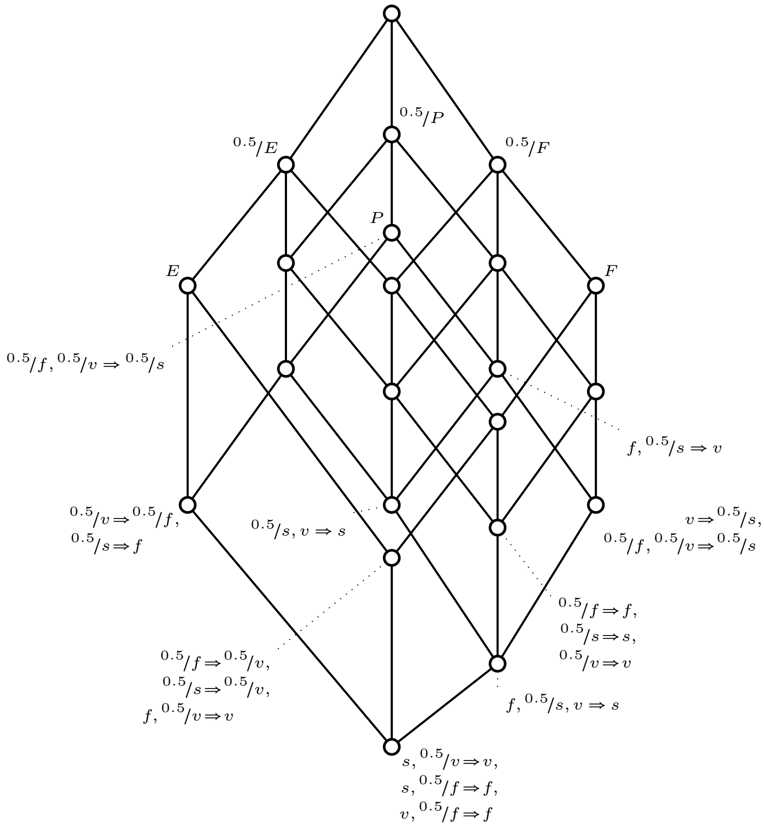

Figure 7.3.: **L**-concept lattice of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ for the context from Figure 7.1 with the Gödel logic and the globalisation on the objects

a presentation, they are also vigilant. This description applies only partially while they are meeting their friends.

An implication $R \Rightarrow S$ between the intents of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ means that if $A \overset{R}{\Rightarrow} B$ holds, then $A \overset{S}{\Rightarrow} B$ must hold as well. The implication $R \Rightarrow S$ is a fuzzy attribute implication and therefore its truth value may be computed by (1.22) (page 29). Since we used the globalisation in $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ for $(-)^{*_G}$, the stem base of the context exists and is uniquely determined. This is

$$E, {}^{0.5}/P \Rightarrow P,$$
$$E, {}^{0.5}/V \Rightarrow V.$$

The drawback of this approach is that the intents (subsets of K_3) are **L**-sets unlike the modi of (K_1, K_2, K_3, Y) are crisp. In order to have crisp intents for the concepts of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ we may use the dual version of the approach presented in [Kra03]. There, for an **L**-context (G, M, I), the derivation operators $^{\upharpoonright} : \mathfrak{P}(G) \to \mathbf{L}^M$ and $^{\downharpoonright} : \mathbf{L}^M \to \mathfrak{P}(G)$ are defined by

$$A^{\upharpoonright}(m) := \bigwedge_{g \in A} I(g, m),$$
$$B^{\downharpoonright} := \{g \in G \mid \text{for all } m \in M : B(m) \leq I(g, m)\}$$

for $A \in \mathfrak{P}(G)$ and $B \in \mathbf{L}^M$. A formal concept in this setting is a tuple (A, B) with $A \in \mathfrak{P}(G)$ and $B \in \mathbf{L}^M$ such that $A^{\upharpoonright} = B$ and $B^{\downharpoonleft} = A$. The set of all formal concepts ordered by the set inclusion on one component forms a complete lattice. One may show, see [BV05e], that the following connection exists between the derivation operators just presented and the ordinary fuzzy derivation operators $(-)^{\uparrow}, (-)^{\downarrow}$ ((1.16) and (1.17), page 24):

$$A^{\upharpoonright} = A^{\uparrow} \text{ and } B^{\downharpoonleft} = {}^1(B^{\downarrow}) = (B^{\downarrow})^*$$

where ${}^1(-)$ is the 1-cut (see (1.8), page 7) and $(-)^*$ is the globalisation. Further, the concept lattice obtained through the method from [Kra03] is isomorphic to the so-called crisply generated concept lattice, see [BSZ05].

With the dual of this method we may now reanalyse the context $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$. The outcome is displayed in Figure 7.4. The labelling is done in the same way as for the concept lattice from Figure 7.3. The difference between the two methods is that in the first case we consider the identity on the attributes, whereas in the second case we use the globalisation. Hence, implications which held in the first case just under some $\mathbf{L}$-set can now be found in the top concept. Further, the f-valued conditional attribute implications in the second case are not separated as much as in the first one. For instance, for implications $A_1 \Rightarrow B_1, A_2 \Rightarrow B_2$ from $\mathrm{Imp}_2(\mathbb{K})$ we might have $\{A_1 \Rightarrow B_1\}^{\uparrow} \neq \{A_2 \Rightarrow B_2\}^{\uparrow}$, while $\{A_1 \Rightarrow B_1\}^{\upharpoonright} = \{A_2 \Rightarrow B_2\}^{\upharpoonright}$. This happens whenever the 1-cuts of the $(-)^{\uparrow}$-derivation operators of the two sets are equal.

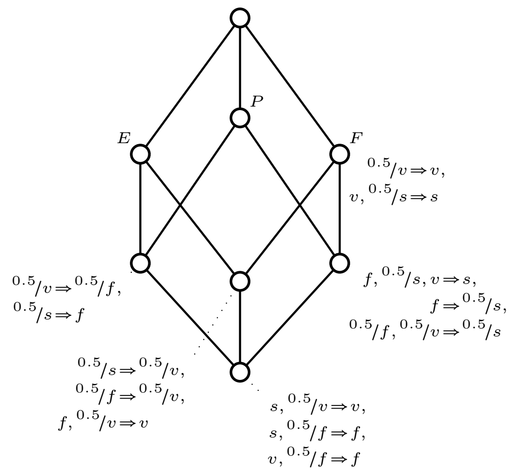

Figure 7.4.: $\mathbf{L}$-concept lattice of $\mathfrak{C}_{\mathrm{Imp}_2}(\mathbb{K})$ for the context from Figure 7.1 with the Gödel logic and the method from [Kra03]

In a tricontext we may arbitrarily interchange the roles of objects, attributes and conditions. Therefore, a tricontext has a sixfold symmetry. By interchanging attributes with conditions in Definition 7.10, we obtain the *attributional condition implications* defined as follows:

Definition 7.12. Let $\mathbb{K} = (K_1, K_2, K_3, Y)$ be an f-valued tricontext. For $A, B \subseteq K_3$ and $X \in \mathbf{L}^{K_2}$ the truth value $||A \overset{X}{\Rightarrow} B||_{\mathbb{K}}$ in $\mathbb{K}$ of the expression $A \overset{X}{\Rightarrow} B$ is given by

$$||A \overset{X}{\Rightarrow} B||_{\mathbb{K}} := ||A \Rightarrow B||_{\mathbb{K}_X^{13}}.$$

Such an implication is called **f-valued attributional condition implication**.

Once again we have

$$\begin{aligned} ||A \overset{X}{\Rightarrow} B||_{\mathbb{K}} &= \mathrm{S}(A, \mathrm{Int}(\mathbb{K}_X^{13}))^* \to \mathrm{S}(B, \mathrm{Int}(\mathbb{K}_X^{13})) \\ &= \mathrm{S}(B, A^{XX}). \end{aligned}$$

Example 7.13. Consider the context $\mathbb{K}$ given in Figure 7.1 with the Gödel logic and the identity as hedge. We have for instance the f-valued attributional condition implications $||E \overset{v}{\Rightarrow} F||_{\mathbb{K}} = 0.5$. The implication means that whenever the students are vigilant during an exam, the same applies partially also while they are meeting their friends. However, if we replace the identity hedge on the objects with the globalisation, then we have that $||E \overset{v}{\Rightarrow} F||_{\mathbb{K}} = 1$.

Similarly to the case of f-valued conditional attribute implications, we may also construct for an f-valued tricontext $\mathbb{K}$ the dyadic $\mathbf{L}$-context

$$\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K}) := (\mathrm{Imp}(K_3), K_2, I)$$

where

$$\mathrm{Imp}(K_3) := \{A \Rightarrow B \mid A, B \subseteq K_3\} \text{ and} \tag{7.13}$$

$$I(A \Rightarrow B, x) := ||A \overset{x}{\Rightarrow} B||_{\mathbb{K}} \tag{7.14}$$

with $A \Rightarrow B \in \mathrm{Imp}(K_3)$ and $x \in K_2$. The concept lattice of $\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K})$ for our running example from Figure 7.1 is displayed in Figure 7.5. We used the same trick for the labelling as for the f-valued conditional attribute implications. Now, $(C, D) \in \underline{\mathfrak{B}}(\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K}))$ contains in its extent all the implications (between the conditions from K_3) that hold under all attributes (elements of K_2) from D. Note that in $\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K})$ as well as for the computation of the truth values of the implications from $\mathrm{Imp}(K_3)$, we used the globalisation for $(-)^{*G}$ in order to ensure that the stem bases exist. Take for instance the right most implication $E \Rightarrow P$. It says: If the students during an exam are partially fevered and partially serious, that they have the same emotion during a presentation.

The stem base of $\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K})$ consists of the following implications:

$$\begin{aligned} \{\,\} &\Rightarrow {}^{0.5}/s, \\ {}^{0.5}/s, f &\Rightarrow f, \\ {}^{0.5}/f, {}^{0.5}/s &\Rightarrow s, \\ {}^{0.5}/f, {}^{0.5}/s, {}^{0.5}/v &\Rightarrow f, s. \end{aligned}$$

Once again there are more variants of defining $\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K})$ depending on the outcome we desire. On the one hand we can use the method from [Kra03] as we did in the case of

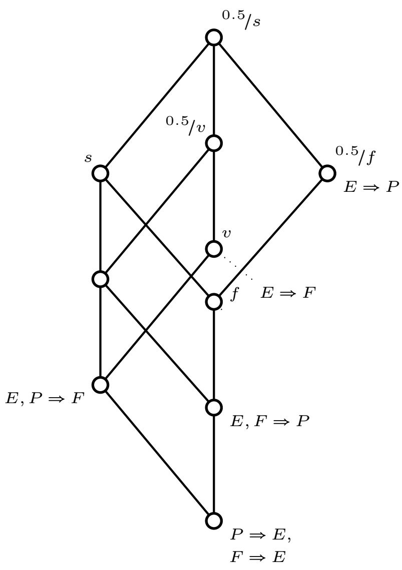

Figure 7.5.: **L**-concept lattice of $\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K})$ for the context from Figure 7.1 with the Gödel logic and the globalisation on the objects

f-valued conditional attribute implications. On the other hand we may use the following definition of $\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K})$ since the extents of the f-valued tricontext $\mathbb{K}$ are **L**-sets:

$$\mathfrak{C}_{\mathrm{Imp}_3}(\mathbb{K}) := (\mathrm{Imp}(K_3), \mathbf{L}^{K_2}, I)$$

where $\mathrm{Imp}(K_3)$ and I are given by (7.13) and (7.14) with $A \Rightarrow B \in \mathrm{Imp}(K_3)$ and $x \in \mathbf{L}^{K_2}$. However, we are missing the theory for contexts, whose attribute sets are **L**-sets.

An open question, also in the crisp case, is the interplay between the f-valued conditional attribute implications and the f-valued attributional condition implications. A further research topic is the exploration of such (f-valued) implications.

7.2.2. F-valued Attribute × Condition Implications

As presented for the discrete case, the two classes of implications studied so far are not powerful enough to express all possible kinds of implications in a tricontext. Therefore, we will generalise the so-called *attribute × condition implications* to our fuzzy-valued setting. For our running example these express implications of the form

- *If we are moderately serious during our presentation, then we are fevered during the exam.*
- *If we are fevered during the exam, then we are partially serious while meeting our friends.*

First, let us introduce some notations. For $A_1, A \in \mathbf{L}^{K_i}$ and $X_1, X \in \mathbf{L}^{K_j}$ $(\{i,j\} = \{1,2\})$ the degree $\mathrm{S}(A_1 \times X_1, A \times X)$ to which $A_1 \times X_1$ is a fuzzy subset of $A \times X$ is defined by

$$\mathrm{S}(A_1 \times X_1, A \times X) := \bigwedge_{(k_i,k_j)\in K_i \times K_j} \Big((A_1 \times X_1)(k_i,k_j) \to (A \times X)(k_i,k_j)\Big) \tag{7.15}$$

where

$$(A \times X)(k_i, k_j) := A(k_i) \otimes X(k_j). \tag{7.16}$$

We can use this definition even if X and X_1 are crisp sets, however with some slight modification. We define

$$(A \times X)(k_i, k_j) := A(k_i) \cdot X(k_j). \tag{7.17}$$

Definition 7.14. Let (K_1, K_2, K_3, Y) be an f-valued tricontext. For $A, B \in \mathbf{L}^{K_2} \times K_3$ the truth value $||A \Rightarrow B||_{\mathbb{K}}$ in $\mathbb{K}$ of the expression $A \Rightarrow B$ is given by

$$||A \Rightarrow B||_{\mathbb{K}} := ||A \Rightarrow B||_{\mathbb{K}^{(1)}}.$$

We call such an implication an **f-valued attribute × condition implication**.

These are the attribute implications of the $\mathbf{L}$-context $\mathbb{K}^{(1)} = (K_1, K_2 \times K_3, Y^{(1)})$. Hence, we have

$$\begin{aligned} ||A \Rightarrow B||_{\mathbb{K}} &= \mathrm{S}(A, \mathrm{Int}(\mathbb{K}^{(1)}))^* \to \mathrm{S}(B, \mathrm{Int}(\mathbb{K}^{(1)})) \\ &= \mathrm{S}(B, A^{(1)(1)}). \end{aligned}$$

For our running example with the Gödel logic and the identity as the hedge, we have for instance the implication $(P, {}^{0.5}/v) \Rightarrow (E, f)$ which holds with the truth values 0.5. It expresses that the following holds only partially: If the students are partially vigilant during a presentation, then they are also fevered during an exam. An f-valued attribute × condition implication which does not hold is $(P, v) \Rightarrow (E, s)$. However, if we replace the identity on the objects with the globalisation, then the latter implication holds with truth value 1.

The $\mathbf{L}$-concept lattice of $\mathbb{K}^{(1)}$ with the globalisation on the objects is displayed in Figure 7.6. We use the globalisation in order to ensure the existence of the stem base. For a better legibility we simplified the attribute labelling as follows: If two tuples have the same first component, then we make a single tuple from them by taking the union of their second components, i.e., if we have (A, a) and (A, b), then we write $(A, \{a, b\})$. The same remark holds for the implications from the stem base for this context.

The $\mathbf{L}$-concept lattice of $\mathbb{K}^{(1)}$ contains eight concepts, whereas the stem base contains fourteen implications! This is not very helpful. We displayed the implications in Figure 7.7. As one can see, they are numerous and hard to handle.

In the crisp case one could determine the attribute × condition implications from the conditional attribute or attributional condition implications. Indeed, from the crisp conditional attribute implication $A \overset{X}{\Rightarrow} B$ with $A, B \subseteq K_2$ and $X \subseteq K_3$, we could compute the corresponding attribute × condition implication by

$$A \times \{x\} \Rightarrow B \times \{x\} \text{ for all } x \in X.$$

However this is not the case in our f-valued setting. We obtain the truth value of the f-valued conditional attribute implication $A \overset{X}{\Rightarrow} B$ with $A, B \in \mathbf{L}^{K_2}$ and $X \subseteq K_3$ from the context $\mathbb{K}^{12}_X$, i.e., Y^{12}_X is given as the infimum over all elements from X. Therefore,

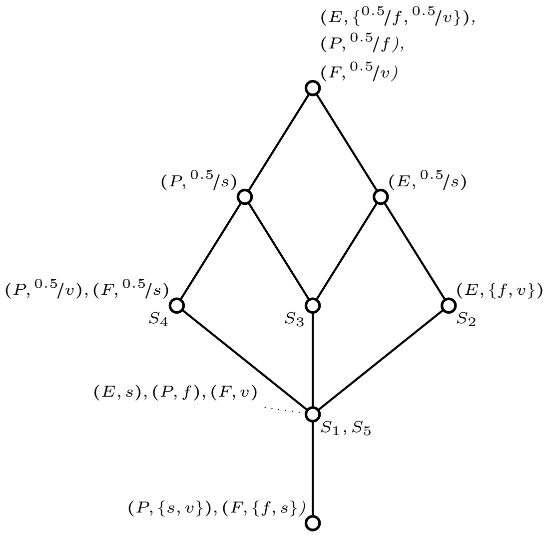

Figure 7.6.: **L**-concept lattice of $\mathbb{K}^{(1)}$ for the context from Figure 7.1 with the Gödel logic and the globalisation on the objects

$$\begin{aligned}
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,{}^{0.5}/f),(F,\{{}^{0.5}/s,{}^{0.5}/v\}) &\Rightarrow (P,\{{}^{0.5}/s,{}^{0.5}/v\}),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,{}^{0.5}/f),(F,{}^{0.5}/v) &\Rightarrow (E,\{f,s,v\}),(P,\{f,{}^{0.5}/s,{}^{0.5}/v\}),(F,{}^{0.5}/s),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,f),(F,{}^{0.5}/v) &\Rightarrow (E,\{f,s,v\}),(P,\{{}^{0.5}/s,{}^{0.5}/v\}),(F,\{f,{}^{0.5}/s\}),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,\{s,{}^{0.5}/f\}),(F,{}^{0.5}/v) &\Rightarrow (E,\{f,s,v\}),(P,\{f,v\}),(F,\{f,s,v\}),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,\{{}^{0.5}/f,{}^{0.5}/s,{}^{0.5}/v\}),(F,\{s,{}^{0.5}/v\}) &\Rightarrow (E,\{f,s,v\}),(P,\{f,s,v\}),(F,\{f,v\}),\\
(E,\{s,{}^{0.5}/f,{}^{0.5}/v\}),(P,{}^{0.5}/f),(F,{}^{0.5}/v) &\Rightarrow (E,\{f,v\}),(P,\{f,{}^{0.5}/s,{}^{0.5}/v\}),(F,\{v,{}^{0.5}/s\}),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,\{{}^{0.5}/f{}^{0.5}/v\}),(F,{}^{0.5}/v) &\Rightarrow (P,{}^{0.5}/s),(F,{}^{0.5}/s),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,\{{}^{0.5}/f,{}^{0.5}/s,v\}),(F,\{{}^{0.5}/s,{}^{0.5}/v\}) &\Rightarrow (E,\{f,s,v\}),(P,\{f,s\}),(F,\{f,s,v\}),\\
\{\,\} &\Rightarrow (E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,{}^{0.5}/f),(F,{}^{0.5}/v),\\
(E,\{f,v,{}^{0.5}/s\}),(P,\{{}^{0.5}/f,{}^{0.5}/v\}),(F,{}^{0.5}/v) &\Rightarrow (E,s),(P,\{f,{}^{0.5}/v),(F,\{v,{}^{0.5}/s\}),\\
(E,\{f,{}^{0.5}/v\}),(P,{}^{0.5}/f),(F,{}^{0.5}/v) &\Rightarrow (E,\{{}^{0.5}/s,v\}),\\
(E,\{{}^{0.5}/f,{}^{0.5}/v\}),(P,{}^{0.5}/f),(F,\{{}^{0.5}/f,{}^{0.5}/v\}) &\Rightarrow (E,\{f,s,v\}),(P,\{f,s,v\}),(F,\{f,s,v\}),\\
(E,\{{}^{0.5}/f,{}^{0.5}/s,{}^{0.5}/v\}),(P,\{{}^{0.5}/f,{}^{0.5}/s,{}^{0.5}/v\}),(F,\{{}^{0.5}/s,{}^{0.5}/v\}) &\Rightarrow (E,\{f,s,v\}),(P,f),(F,v),\\
(E,\{{}^{0.5}/f,v\}),(P,{}^{0.5}/f),(F,{}^{0.5}/v) &\Rightarrow (E,\{f,{}^{0.5}/s\}).
\end{aligned}$$

Figure 7.7.: Stem base of the **L**-context $\mathbb{K}^{(1)}$

we cannot directly compute the truth value $||A \overset{x}{\Rightarrow} B||_{\mathbb{K}}$ for $x \in X$ and hence also not the truth value of $A \times \{x\} \Rightarrow B \times \{x\}$. Going the other way around, namely transforming the f-valued attribute × condition implications into f-valued conditional attribute and attributional condition implications, yields a similar situation.

Taking into account the previous two remarks on f-valued attribute × condition implications, one probably wants to avoid them, as they lose the smooth handling their crisp variants have.

One could also be interested in f-valued object × attribute or object × condition implications. For our example, the first kind would be for instance *If the first group of students is fevered, then the second one is serious.* One should always keep in mind the expressiveness of such implications. Thus, the second kind of implications, between tuples of attributes and conditions, do not make much sense for our example.

7.3. Fuzzy-valued Triconceptual Factorisation

Now we turn our attention to the conceptual factorisation problem and generalise the results from Section 3.1 to our fuzzy-valued setting.

We start by defining what a conceptual factorisation in our fuzzy-valued setting is. Afterwards, we investigate some properties of these factorisations, for instance, their optimality and size. Further, we show how the f-valued triconceptual factorisation works by illustrating it on our running example. For practice, we need an appropriate algorithm doing the job, which we present thereafter. We conclude the section with the transformations between the space of attributes × conditions and the space of factors.

Definition 7.15. A **factorisation** of an f-valued tricontext (K_1, K_2, K_3, Y) consists of two **L**-contexts (K_1, F, I_1) and (K_2, F, I_2) and a formal context (K_3, F, I_3) such that

$$Y(a_1, a_2, a_3) = l \iff (I_1(a_1, f) \otimes I_2(a_2, f)) \cdot I_3(a_3, f) = l \text{ for some } f \in F.$$

The set F is called the **factor set**, its elements the **(f-valued triadic) factors**, and (K_i, F, I_i) $(i = 1, 2, 3)$ are said to be the **factorisation contexts**. We write

$$(K_1, K_2, K_3, Y) = \circ((K_1, F, I_1), (K_2, F, I_2), (K_3, F, I_3))$$

to indicate a factorisation.

Similarly to the crisp triadic case, the factorisation contexts represent relationships between objects and factors, attributes and factors, conditions and factors, respectively. Hence, the object a_1 has the attribute a_2 under the condition a_3 if there is a factor f which applies to a_1, for which a_2 is one of its manifestations and which exists under condition a_3. As we have seen in Section 2.4, the first two are graded relations, whereas the third one is crisp.

We may carry on with the translations from the crisp triadic case into the f-valued triadic setting. To each factorisation we associate a **factorising family**

$$\{(A_f^1, A_f^2, A_f^3) \mid f \in F\}$$

given by the **L**-sets $A^1_f \in \mathbf{L}^{K_1}$, $A^2_f \in \mathbf{L}^{K_2}$ and the crisp set $A_3 \subseteq K_3$ defined by

$$A^1_f(k_1) := I_1(k_1, f),\ \ A^2_f(k_2) := I_2(k_2, f)\ \text{ and }\ A^3_f := \{a_3 \in K_3 \mid a_3\ I_3\ f\}$$

for all $k_1 \in K_1$ and $k_2 \in K_2$. Once more, a family $\{(A^1_f, A^2_f, A^3_f) \mid f \in F\}$ is a factorising family of (K_1, K_2, K_3, Y) if and only if

$$Y = \bigcup_{f \in F} (A^1_f \bigotimes A^2_f) \times A^3_f, \tag{7.18}$$

where $A^1_f \otimes A^2_f$ is the **L**-set given by (2.6) (page 65). Further, (7.18) is equivalent to

$$Y = \bigcup_{f \in F} (f^{I_1} \bigotimes f^{I_2}) \times f^{I_3}.$$

Once again the factorising families are precisely those families of f-valued triadic preconcepts[4] of (K_1, K_2, K_3, Y) covering the **L**-relation Y. By enlarging these f-valued preconcepts we obtain a factorising family of f-valued triconcepts. In the following we will call such factorisations **f-valued triconceptual**. Similar to the crisp dyadic and crisp triadic case the enlargement is not unique.

We may uniquely determine a factorisation context through the other two. Indeed, in $\mathbb{K}$ from A^i_f and A^j_f we find the set $A^k_f = (A^i_f)^{A^j_f}$ for each $f \in F$.

Proposition 7.16. *For any f-valued triconceptual factorisation with factor set F the intersection of the attribute orders of (K_i, F, I_i) and (K_j, F, I_j) is contained in the dual attribute order of (K_k, F, I_k) for $\{i, j, k\} = \{1, 2, 3\}$.*

Proof. The proof goes exactly in same way as the one for its crisp version (Proposition 3.11, page 74), we just have to replace crisp sets and derivation operators by fuzzy ones. □

This time we may also choose between different representations among the factorisation contexts. For instance, we might be interested in the relationship between objects and attributes for each factor independent from the condition. Then, we may define the tricontext $\mathbb{K}^J_{12} := (K_1, K_2, F, J)$, where

$$J\ (a_1, a_2, f) := I_1(a_1, f) \otimes I_2(a_2, f).$$

Each factor d-cut in $\mathbb{K}^J_{12}$ represents the relationship between the objects and attributes for that factor. However, one may also be interested in the relationship between attributes and conditions independent from the objects. Then, in a similar manner as before, one can build $\mathbb{K}^J_{23}$ with

$$J(a_2, a_3, f) := I_2(a_2, f) \cdot I_3(a_3, f).$$

It is easy to reconstruct the factorisation from these contexts. Indeed, we have

$$Y(a_i, a_j, a_k) = l \iff l = \begin{cases} J(a_i, a_j, f) \cdot I_k(a_k, f), & k = 3, \\ J(a_i, a_j, f) \otimes I_k(a_k, f), & \text{else.} \end{cases} \tag{7.19}$$

[4]A triple (A_1, A_2, A_3) with $A_1 \in \mathbf{L}^{K_1}$, $A_2 \in \mathbf{L}^{K_2}$ and $A_3 \subseteq K_1$ is called an *f-valued triadic preconcept* if $A_i^{A_j} \subseteq A_k$ for all $\{i, j, k\} = \{1, 2, 3\}$.

In the previous section we have seen that the f-valued triadic derivation operators satisfy the same properties as triadic derivation operators. A similar remark holds for the f-valued triconcepts. Therefore, we may generalise the results about triconceptual factorisations from Section 3.1 into our f-valued triadic setting. Even more, the results need not to be re-proven since the proofs were entirely based on the properties of the derivation operators and the maximality of the concepts. Therefore, we will just list these results without proofs.

Theorem 7.17. *Fuzzy-valued triconceptual factorisations yield optimal factorisations, i.e., the smallest possible number of factors.*

Theorem 7.18. *Let $\mathbb{K} = (K_1, K_2, K_3, Y)$ be an f-valued tricontext with $|K_1| = p, |K_2| = q$ and $|K_3| = r$ and let F be the factor set of an f-valued triconceptual factorisation. Then, $|F| \le \min\{pq, pr, qr\}$.*

Trivially, the f-valued triconceptual factorisation is "invariant" under clarification and interchanging of objects, attributes and conditions in f-valued tricontexts in the sense described for the crisp dyadic case in Proposition 2.6 and Proposition 2.7 (page 42).

Example 7.19. We illustrate the f-valued triconceptual factorisation on the f-valued tricontext given in Figure 7.1 (page 141). Recall that the number of f-valued triconcepts is 34 with the Gödel logic and 28 with the Łukasiewicz logic. The factorisation contexts with both logics are displayed in Figure 7.8 and 7.9. One immediately sees that we have six factors in both cases. Hence, by using the triconceptual factorisation we are able to describe the objects through six factors instead of three attributes under three conditions. In our toy example the number of objects is higher than the number of attributes and conditions. If we switch K_1 and K_2 in the context, we would still have six factors and the data reduction would be even higher. In lieu of describing the elements from K_2 through five times three items, we would have just six factors.

	f_1	f_2	f_3	f_4	f_5	f_6
S_1	1	1	1	1	0.5	1
S_2	0.5	1	0.5	0	0	0.5
S_3	0.5	0.5	0.5	0	0.5	0.5
S_4	0	0.5	0.5	1	1	0.5
S_5	1	1	1	1	0.5	1

	f	s	v
f_1	1	1	1
f_2	1	0	1
f_3	0	0	1
f_4	0	0.5	0.5
f_5	1	1	0
f_6	1	0	0

d

	E	P	F
f_1	×		
f_2	×		
f_3	×		×
f_4		×	×
f_5		×	
f_6	×	×	

d

Figure 7.8.: Factorisation contexts using the Gödel logic

The factors from the two factorisations are quite similar. For instance, f_1 is the same in both cases. Denote by $(-)^{I_{i\text{G}}}$ and $(-)^{I_{i\text{Ł}}}$ $(i = 1, 2, 3)$ the derivation operators in the factorisation contexts for the Gödel logic and the Łukasiewicz logic, respectively. We see that the factors usually satisfy

$$f^{I_{1\text{G}}} = f^{I_{1\text{Ł}}}, f^{I_{3\text{G}}} = f^{I_{3\text{Ł}}} \text{ and } f^{I_{2\text{G}}} \subseteq f^{I_{2\text{Ł}}}$$

	f_1	f_2	f_3	f_4	f_5	f_6
S_1	1	1	1	1	0.5	1
S_2	0.5	1	0.5	0.5	0	0.5
S_3	0.5	0.5	0.5	0.5	0	0.5
S_4	0	0.5	0.5	1	0.5	0.5
S_5	1	1	1	1	0.5	1

	f	s	v
f_1	1	1	1
f_2	1	0.5	1
f_3	0	0.5	1
f_4	0	0.5	0.5
f_5	1	1	1
f_6	1	0.5	0.5

d

	E	P	F
f_1	×		
f_2	×		
f_3	×		×
f_4		×	×
f_5		×	
f_6	×	×	

d

Figure 7.9.: Factorisation contexts using the Łukasiewicz logic

or

$$f^{I_{2\mathrm{G}}} = f^{I_{2\text{Ł}}}, f^{I_{3\mathrm{G}}} = f^{I_{3\text{Ł}}} \text{ and } f^{I_{1\mathrm{G}}} \subseteq f^{I_{1\text{Ł}}}.$$

Hence, we basically obtain the "same" f-valued triconcepts in both cases. The difference mainly lies in the different logic and its consequence on some elements of the concepts.

Let us now look at the interpretation of the factors. The first two are typical for the case when the students give an exam and they describe the emotions of the students during this situation. The third factor shows a connection between "giving an exam" and "meeting friends". During both events the students are vigilant and, depending on the logic, partially serious. The fourth and sixth factor establish a link between the emotions of the students during a presentation with the emotions while meeting their friends and during an exam, respectively. The fifth factor is characteristic for the presentation. Further, we can see that there are two factors which are typical for the exam situation and one for the presentation. The emotions the students have while meeting their friends are subsumed under the emotions of the other events.

In order to apply the f-valued triconceptual factorisation in practice we need an appropriate algorithm. We obtain it by combining the algorithm from [Běl08] for computing conceptual factorisations of dyadic **L**-contexts and the one from [BGV12] for the computation of crisp triconceptual factorisations. We already know that finding a conceptual factorisation, independent from the nature of the data, is reducible to the set covering problem. Therefore, it is NP-hard and the corresponding decision problem is NP-complete. Recall that there exists a greedy approximation algorithm for the set covering optimisation problem which achieves an approximation ratio $\leq \ln(|U|) + 1$, where U is to be covered, see [CLRS01]. This time, the universe U corresponds to Y, the ternary **L**-relation of the f-valued tricontext $\mathbb{K}$. The family $\mathcal{S}$ of subsets of the universe U that is used for finding a cover corresponds to the set of all f-valued triconcepts $\mathfrak{T}(\mathbb{K})$. In this setting, we are looking for $\mathcal{C} \subseteq \mathcal{S}$ as small as possible such that $\bigcup \mathcal{C} = U$.

Algorithm 4, the implementation of the above-mentioned greedy approach in our setting, works as follows: It determines the f-valued triconceptual factorisation families of the triconceptual factorisation by first computing the set of all f-valued triconcepts which are stored in $\mathcal{S}$ (lines 1–8) and then iteratively selecting triconcepts from $\mathcal{S}$, maximising their overlap with the remaining tuples in U and computing the triconceptual factorisation families (lines 9–21). Note that the f-valued triconcepts are computed by a reduction to

Algorithm 4: $\textsc{ComputeFactors}(K_1, K_2, K_3, Y)$

1 **set** $\mathcal{S}$ **to** $\varnothing$;
2 **foreach** $(D_1, J) \in \mathfrak{B}\big(K_1, K_2 \times K_3, Y^{(1)}\big)$ **do**
3   **foreach** $(D_2, D_3) \in \mathfrak{B}(K_2, K_3, J)$ **do**
4   **if** $D_1 = (D_2 \times D_3)'$ **then**
5   **add** (D_1, D_2, D_3) **to** $\mathcal{S}$;
6   **end**
7   **end**
8 **end**
9 **set** $\mathcal{F}$ **to** $\varnothing$;
10 **set** U **to** Y;
11 **while** $U \neq \varnothing$ **do**
12   **set** j **to** 1;
13   **select** $(D_1, D_2, D_3) \in \mathcal{S}$ **which maximises** $|U \cap ((D_1 \otimes D_2) \times D_3)|$;
14   **foreach** $i \in \{1, 2, 3\}$ **do**
15   **set** $A^i_{f_j}$ **to** D^i;
16   **end**
17   **set** j **to** $j + 1$;
18   **set** U **to** $U \smallsetminus ((D_1 \otimes D_2) \times D_3)$;
19   **remove** (D_1, D_2, D_3) **from** $\mathcal{S}$;
20 **end**
21 **return** $(A^1_{f_j}, A^2_{f_j}, A^3_{f_j})$ *for all j*

the dyadic case, as it was done for crisp tricontexts in [JHS+06]. In line 2, we iterate over all dyadic $\mathbf{L}$-concepts in $\mathfrak{B}\big(K_1, K_2 \times K_3, Y^{(1)}\big)$. In line 3, we iterate over all concepts in $\mathfrak{B}(K_2, K_3, J)$ where J was obtained as an intent in the previous line. The condition in line 4 is needed to check whether D_1 is maximal (note that $(-)'$ in line 4 denotes a derivation operator induced by the dyadic $\mathbf{L}$-context $(K_1, K_2 \times K_3, Y^{(1)})$), i.e., whether (D_1, D_2, D_3) is an f-valued triconcept.

As we have already pointed out in Section 2.1 transformations between the space of attributes and the space of factors proved themselves to be useful in the crisp dyadic case, see for instance [Out10]. Therefore, we generalise these mappings into our f-valued triadic setting. Define $\varphi : \mathfrak{P}(\mathbf{L}^{K_2} \times K_3) \to \mathbf{L}^F$ and $\psi : \mathbf{L}^F \to \mathfrak{P}(\mathbf{L}^{K_2} \times K_3)$ by

$$\varphi(P)(f) := \mathrm{S}(f^{I_2} \times f^{I_3}, P),$$

$$\psi(Q)(k_2, k_3) := \begin{cases} (0, \{\varnothing\}), & f^{I_3}(k_3) = 0 \text{ for all } f \in F, \\ \Big(\bigvee_{f \in F} \big((Q(f) \otimes f^{I_2}(k_2))\big) \cdot f^{I_3}(k_3), \{k_3\}\Big), & \text{else} \end{cases}$$

for $P \in \mathfrak{P}(\mathbf{L}^{K_2} \times K_3)$ and $Q \in \mathbf{L}^F$.

Theorem 7.20. *For any $g \in K_1$ we have*

$$\varphi(g^{(1)}) = g^{I_1} \quad \text{and} \quad \psi(g^{I_1}) = g^{(1)}.$$

That is, φ maps the object dyadic cuts of (K_1, K_2, K_3, Y) to the rows of (K_1, F, I_1) and ψ maps the rows of (K_1, F, I_1) to the object dyadic cuts of (K_1, K_2, K_3, Y).

Proof. In this proof we use the properties of residuated lattices listed on page 7 and the definitions of the f-valued triadic $(-)^{A_k}$-derivation operators. Evidently, $\psi(g^{I_1}) = g^{(1)}$ follows directly from the definition of the f-valued triconceptual factorisation and of ψ. For $\varphi(g^{(1)}) = g^{I_1}$ first note that

$$\begin{aligned}
g^{(1)}(k_2, k_3) &= \Big(\bigwedge_{k_1 \in K_1} (g(k_1) \to Y^{(1)}(k_1, k_2, k_3)), k_3\Big)(k_2, k_3) \\
&= (Y(g, k_2, k_3), k_3)(k_2, k_3) \\
&= Y(g, k_2, k_3)
\end{aligned}$$

for all $k_2 \in K_2$ and $k_3 \in K_3$. Then, using the fuzzy subsethood from (7.15) (page 153) we have

$$\begin{aligned}
\varphi(g^{(1)})(f) &= \mathrm{S}(f^{I_2} \times f^{I_3}, g^{(1)}) \\
&= \bigwedge_{(k_2,k_3) \in K_2 \times K_3} \Big((f^{I_2} \times f^{I_3})(k_2, k_3) \to g^{(1)}(k_2, k_3)\Big) && \text{(by (7.15))} \\
&= \bigwedge_{(k_2,k_3) \in K_2 \times K_3} \Big((f^{I_2} \times f^{I_3})(k_2, k_3) \to Y(g, k_2, k_3)\Big) \\
&= \bigwedge_{k_2 \in K_2} \bigwedge_{k_3 \in K_3} \Big((f^{I_2}(k_2) \cdot f^{I_3}(k_3)) \to Y(g, k_2, k_3)\Big) && \text{(by (7.17))} \\
&= \bigwedge_{k_2 \in K_2} \bigwedge_{k_3 \in K_3} \Big((A^2_f(k_2) \cdot A^3_f(k_3)) \to Y(g, k_2, k_3)\Big) \\
&= \bigwedge_{k_2 \in K_2} \bigwedge_{k_3 \in K_3} \Big(A^2_f(k_2) \to (A^3_f(k_3) \to Y(g, k_2, k_3))\Big) && \text{(by (1.3))} \\
&= \bigwedge_{k_2 \in K_2} \left(A^2_f(k_2) \to \bigwedge_{k_3 \in K_3} (A^3_f(k_3) \to Y(g, k_2, k_3))\right) && \text{(by (1.5))} \\
&= \bigwedge_{k_2 \in K_2} \left(A^2_f(k_2) \to Y^{12}_{A^3_f}(g, k_2)\right) && \text{(by (7.4))} \\
&= (A^2_f)^{A^3_f}(g) \\
&= (A^1_f)(g) \\
&= g^{I_1}(f)
\end{aligned}$$

finishing the proof. □

Hence, we may further generalise the results from Section 3.2. Indeed, we have the following:

Lemma 7.21. *For $P, Q \in \mathfrak{P}(\mathbf{L}^{K_2} \times K_3)$ and $S, T \in \mathbf{L}^F$, we have*

$$\begin{aligned}
&P \subseteq Q \Longrightarrow \varphi(P) \subseteq \varphi(Q), \\
&S \subseteq T \Longrightarrow \psi(S) \subseteq \psi(T), \\
&\psi(\varphi(P)) \subseteq P, \\
&S \subseteq \varphi(\psi(S)).
\end{aligned}$$

Corollary 7.22. *For $P, P_j \in \mathfrak{P}(\mathbf{L}^{K_2} \times K_3)$ and $S, S_j \in \mathbf{L}^F$ with $j \in J$, we have*

$$\varphi(P) = \varphi\psi\varphi(P),$$
$$\psi(S) = \psi\varphi\psi(S),$$
$$\varphi\Big(\bigcap_{j \in J} P_j\Big) = \bigcap_{j \in J} \varphi(P_j),$$
$$\psi\Big(\bigcup_{j \in J} S_j\Big) = \bigcup_{j \in J} \psi(S_j).$$

Once again we find the geometry behind the transformations through their inverse mappings. Therefore, define

$$\varphi^{-1}(S) := \{P \in \mathfrak{P}(\mathbf{L}^{K_2} \times K_3) \mid \varphi(P) = S\},$$
$$\psi^{-1}(P) := \{S \in \mathbf{L}^F \mid \psi(S) = P\}$$

for $P \in \mathfrak{P}(\mathbf{L}^{K_2} \times K_3)$ and $S \in \mathbf{L}^F$.

Theorem 7.23. (1) *$\varphi^{-1}(S)$ is a convex partially ordered subspace of the attribute and condition space and $\psi(S)$ is the least element of $\varphi^{-1}(S)$.*
(2) *$\psi^{-1}(P)$ is a convex partially ordered subspace of the factor space and $\varphi(P)$ is the largest element of $\psi^{-1}(P)$.*

7.4. Conclusion

In this chapter we presented a new framework for fuzzy-valued triadic data together with some useful methods for handling them. In Section 7.1 we mainly focused on the fundamentals, presenting fuzzy-valued triadic contexts, concepts and derivation operators and investigated some of their properties. In this connection, the link between the fuzzy-valued tricontext and its double-scaled context proved itself to be very useful. In Section 7.2 we developed various kinds of fuzzy-valued triadic implications, which arose by generalisations from the crisp case presented in [GO04]. A valuable method of Formal Concept Analysis, the conceptual factorisation, was generalised into our setting in Section 7.3. The motivation for the last two sections is given by the wish of having as many methods from Formal Concept Analysis as possible for our setting. Such methods allow a deeper investigation of fuzzy-valued data and enhance their applicability.

This chapter may be considered as a pioneer work in connecting Triadic Concept Analysis and Formal Concept Analysis with fuzzy attributes. Although there are also other works in this direction ([BO10, OK10]), there are many aspects and methods from Formal Concept Analysis left unexplored. Further, yet unanswered questions regarding the connection between the different kinds of crisp triadic implications could shed light on this matter for the fuzzy-valued triadic implications.

Bibliography

[Adl85] K.-P. Adlassnig, *Approaches to computer-assisted medical diagnosis*, Computers in Biology and Medicine **15** (1985), 315–335.

[BCV04] R. Bělohlávek, M. Chlupová, and V. Vychodil, *Implications from data with fuzzy attributes*, AISTA 2004 in Cooperation with the IEEE Computer Society Proceedings, 2004.

[Běl01] R. Bělohlávek, *Fuzzy closure operators*, Journal of Mathematical Analysis and Applications **262** (2001), 473–489.

[Běl02a] ———, *Algorithms for fuzzy concept lattices*, International Conference on Recent Advances in Soft Computing, 2002, pp. 200–205.

[Běl02b] ———, *Fuzzy relational systems: Foundations and principles*, Systems Science and Engineering, Kluwer Academic/Plenum Press, 2002.

[Běl08] ———, *Optimal decompositions of matrices with grades*, International IEEE Conference on Intelligent Systems, IEEE IS 2008, 2008, pp. 152–157.

[BFG75] A. Burusco and R. Fuentes-Gonzáles, *The study of the L-fuzzy concept lattice*, International Journal of Man-Machine Studies **7** (1975), no. 1, 1–13.

[BFV05] R. Bělohlávek, T. Funioková, and V. Vychodil, *Fuzzy closure operators with truth stressers*, Logic Journal of the IGPL **13** (2005), no. 5, 503–513.

[BG96] L. Biacino and G. Gerla, *An extension principle for closure operators*, Journal of Mathematical Analysis and Applications **198** (1996), 1–24.

[BGV12] R. Bělohlávek, C. V. Glodeanu, and V. Vychodil, *Optimal factorization of three-way binary data using triadic concepts*, Order (2012), online first.

[Bie88] K. Biedermann, *A foundation of the theory of trilattices*, Shaker Verlag, Aachen, 1988.

[Bir67] G. Birkhoff, *Lattice Theory*, Colloquium Publications, vol. 25, American Mathematical Society, Providence, RI, USA, 1967, Third edition.

[BN97] H.-G. Bartel and M. Nofz, *Exploration of NMR data of glasses by means of Formal Concept Analysis*, Chemometrics and Intelligent Laboratory Systems **36** (1997), 53–63.

[BO10] R. Bělohlávek and P. Osicka, *Triadic Concept Analysis of data with fuzzy attributes*, IEEE International Conference on Granular Computing (X. Hu, T. Y. Lin, V. V. Raghavan, J. W. Grzymala-Busse, Q. Liu, and A. Z. Broder, eds.), IEEE Computer Society, 2010, pp. 661–665.

[BS05] R. Bělohlávek and V. Sklenář, *Formal Concept Analysis constrained by attribute-dependency formulas*, International Conference on Formal Concept Analysis (B. Ganter and R. Godin, eds.), Lecture Notes in Computer Science, Springer, 2005, pp. 176–191.

[BSZ05] R. Bělohlávek, V. Sklenář, and J. Zacpal, *Crisply generated fuzzy concepts*, International Conference on Formal Concept Analysis, Lecture Notes in Computer Science, Springer, 2005, pp. 269–284.

[BV05a] R. Bělohlávek and V. Vychodil, *Functional dependencies of data tables over domains with similarity relations*, Indian International Conference on Artificial Intelligence, 2005, pp. 2486–2504.

[BV05b] ———, *Fuzzy attribute implications: Computing non-redundant bases using maximal independent sets*, Australian Conference on Artificial Intelligence (S. Zhang and R. Jarvis, eds.), 2005, pp. 1126–1129.

[BV05c] ———, *Fuzzy attribute logic: attribute implications, their validity, entailment, and non-redundant basis*, Eleventh International Fuzzy Systems Association World Congress, (Y. Liu, G. Chen, and M. Ying, eds.), Fuzzy Logic, Soft Computing & Computational Intelligence, vol. 1, Tsinghua University Press and Springer, 2005, pp. 622–627.

[BV05d] ———, *Implications from data with fuzzy attributes vs. scaled binary attributes*, FUZZ-IEEE: IEEE International Conference on Fuzzy Systems, 2005, pp. 1050–1055.

[BV05e] ———, *What is a fuzzy concept lattice?*, Concept Lattices and Their Applications (Olomouc), vol. 162, CEUR-WS.org, 2005, pp. 34–45.

[BV06a] ———, *Attribute implications in a fuzzy setting*, International Conference on Formal Concept Analysis (R. Missaoui and J. Schmid, eds.), Lecture Notes in Computer Science, Springer, 2006, pp. 45–60.

[BV06b] ———, *Dense rectangles in object-attribute data*, IEEE International Conference on Granular Computing (T. Y. Lin and Y. Q. Zhang, eds.), 2006, pp. 586–591.

[BV06c] ———, *Properties of models of fuzzy attribute implications*, International Conference on Soft Computing and Intelligent Systems and International Symposium on Intelligent Systems, 2006, pp. 20–24.

[BV07] ———, *Fuzzy concept lattices constrained by hedges*, Journal of Advanced Computational Intelligence and Intelligent Informatics **11** (2007), no. 6, 536–545.

[BV09a] ———, *Factor Analysis of incidence data via novel decomposition of matrices*, International Conference on Formal Concept Analysis (S. Ferré and S. Rudolph, eds.), Lecture Notes in Computer Science, vol. 5548, Springer, 2009, pp. 83–97.

[BV09b] ———, *Formal Concept Analysis with background knowledge: Attribute priorities*, IEEE Transactions on Systems, Man, and Cybernetics, Part C **39** (2009), no. 4, 399–409.

[BV10a] ———, *Discovery of optimal factors in binary data via a novel method of matrix decomposition*, Journal of Computer and System Sciences **76** (2010), no. 1, 3–20.

[BV10b] ———, *Optimal factorization of three-way binary data*, IEEE International Conference on Granular Computing (X. Hu, T. Y. Lin, V. Raghavan, J. Grzymala-Busse, Q. Liu, and A. Broder, eds.), 2010, pp. 61–66.

[BV11] ———, *Codd's relational model from the point of view of fuzzy logic*, Journal of logic and computation **21** (2011), no. 5, SI, 851–862.

[BVG01] J. Berkhof, I. Van Mechelen, and A. Gelman, *A bayesian approach to the selection and testing of mixture models*, Tech. report, Statistica Sinica, 2001.

[CH08] P. Caputi and D. Hennessy, *Using Formal Concept Analysis to analyse repertory grid data*, Personal Construct Theory and Practice (2008), 165–173.

[CK55] C. H. Coombs and C.-W. Kao, *Nonmetric Factor Aanalysis*, University of Michigan; Engineering Research Institute; Bulletin, University of Michigan Engineering Research Institute, 1955.

[Cla09] N. Clara, *Hierarchies generated for data represented by fuzzy ternary relations*, WSEAS international conference on Systems (Stevens Point, Wisconsin, USA), World Scientific and Engineering Academy and Society (WSEAS), 2009, pp. 121–126.

[CLRS01] Th. H. Cormen, Ch. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms, Second Edition*, McGraw-Hill Science/Engineering/Math, jul 2001.

[Coo64] C. H. Coombs, *A theory of data*, Wiley, 1964.

[CVC94] J. C. Cubero, M. A. Vila, and F. Ciencias, *A new definition of fuzzy functional dependency in fuzzy relational databases*, International Journal of Intelligent Systems **9** (1994), 441–448.

[CVL03] E. Ceulemans, I. Van Mechelen, and I. Leenen, *Tucker3 Hierarchical Classes Analysis*, Psychometrika **68** (2003), 413–433.

[CY05] Y. H. Chen and Y. Y. Yao, *Formal Concept Analysis and Hierarchical Classes Analysis*, Fuzzy Information Processing Society, 2005, pp. 276–281.

[Dat00] C. J. Date, *Database relational model: A retrospective review and analysis*, Addison-Wesley, 2000.

[DDF84] J.-P. Doignon, A. Ducamp, and J.-C. Falmagne, *On realizable biorders and the biorder dimension of a relation*, Journal of Mathematical Psychology **28** (1984), 73–109.

[DP02] B. A. Davey and H. A. Priestley, *Introduction to lattices and order*, second ed., Cambridge University Press, Cambridge, 2002.

[DR88] P. De Boeck and S. Rosenberg, *Hierarchical Classes: model and data analysis*, Psychometrika **53** (1988), 361–81.

[DRV93] P. De Boeck, S. Rosenberg, and I. Van Mechelen, *The Hierarchical Classes approach: a review*, Cognitive Science Series, ch. 11, pp. 265–286, Academic Press, London, 1993.

[DS11] F. Distel and B. Sertkaya, *On the complexity of enumerating pseudo-intents*, Discrete Applied Mathematics **159** (2011), no. 6, 450–466.

[Duq87] V. Duquenne, *Contextual implications between attributes and some representation properties for finite lattices*, Beitrage zur Begriffsanalyse (B. Ganter, R. Wille, and K. E. Wolff, eds.), BI Wissenschaftsverlag, 1987, pp. 213–239.

[DW39] R. P. Dilworth and M. Ward, *Residuated lattices*, Trans. A.M.S. **45** (1939), 335–354.

[DW01] F. Dau and R. Wille, *On the modal understanding of triadic contexts*, Classification and Information Processing at the Turn of the Millenium (R. Decker and W. Gaul, eds.), Gesellschaft fü Klassifikation, 2001.

[EKSW99] D. Eschenfelder, W. Kollewe, M. Skorsky, and R. Wille, *Ein Erkundungssystem zum Baurecht: Methoden der Entwicklung eines TOSCANA-Systems*, vol. 2036, Techn. Univ., FB 4, Darmstadt, Januar 1999, Ersch. ebenf. in: Begriffliche Wissensverarbeitung: Methoden und Anwendungen. Hrsg.: G. Stumme, R. Wille. - Berlin, Heidelberg (u.a.): Springer, 2000. S. 254-272.

[Ern82] M. Erné, *Einführung in die Ordnungstheorie*, BI Wissenschaftsverlag, 1982.

[Gan84] B. Ganter, *Two basic algorithms in Concept Analysis*, Tech. Report 831, Technische Hochschule Darmstadt, 1984.

[Gan99] ———, *Attribute exploration with background knowledge*, Theoretical Computer Science **217** (1999), no. 2, 215–233.

[GD86] J.-L. Guigues and V. Duquenne, *Familles minimales d'implications informatives resultant d'un tableau de donnes binaires*, Mathématiques et Sciences Humaines **24** (1986), no. 95, 5–18.

[GG12] B. Ganter and C. V. Glodeanu, *Ordinal Factor Analysis*, International Conference on Formal Concept Analysis (F. Domenach, D.I. Ignatov, and J. Poelmans, eds.), Lecture Notes in Computer Science, vol. 7278, Springer-Verlag Berlin Heidelberg, 2012, pp. 128–139.

[Glo10] C. V. Glodeanu, *Triadic factor analysis*, Concept Lattices and Their Applications (M. Kryszkiewicz and S. A. Obiedkov, eds.), CEUR Workshop Proceedings, CEUR-WS.org, 2010, pp. 127–138.

[Glo11a] ______, *Factorization methods of binary, triadic, real and fuzzy data*, Knowledge Engineering: Principles and Techniques Conference, Informatica, Studia Universitatis Babes-Bolyai, 2011, pp. 81–86.

[Glo11b] ______, *Factorization with Hierarchical Classes Analysis and with Formal Concept Analysis*, International Conference on Formal Concept Analysis (R. Jäschke and P. Valtchev, eds.), Lecture Notes in Computer Science, vol. 6628, Springer, 2011, pp. 107–118.

[Glo11c] ______, *Fuzzy-valued triadic concept analysis and its applications*, Tech. Report MATH-AL-07-2011, TU Dresden, September 2011.

[Glo11d] ______, *Fuzzy-valued triadic implications*, Concept Lattices and Their Applications (A. Napoli and V. Vychodil, eds.), INRIA Nancy - Grand Est and LORIA, 2011, pp. 159–173.

[Glo12a] ______, *Attribute dependencies in a fuzzy setting*, Concept Lattices and Their Applications (U. Priss and L. Szathmary, eds.), CLA Conference Series, Universidad de Malaga (Dept. Matemática Aplicada), Spain, 2012, pp. 127–138.

[Glo12b] ______, *Attribute exploration in a fuzzy setting*, Contributions to International Conference on Formal Concept Analysis (F. Domenach, D.I. Ignatov, and J. Poelmans, eds.), Katholieke Universiteit Leuven, 2012.

[GO04] B. Ganter and S. A. Obiedkov, *Implications in triadic formal contexts*, International Conference on Conceptual Structures (K. E. Wolff, H. D. Pfeiffer, and H. S. Delugach, eds.), Lecture Notes in Computer Science, 2004, pp. 186–195.

[Göd32] K. Gödel, *Zum intuitionistischen Aussagenkalkül*, Anzeiger der Akademie der Wissenschaftischen in Wien **69** (1932), 65–66.

[Gog67] J. A. Goguen, *L-fuzzy sets*, Journal of Mathematical Analysis and Applications **18** (1967), 145–174.

[Gog69] ______, *The logic of inexact concepts*, Synthese **19** (1969), 325–373.

[Got88] S. Gottwald, *Mehrwertige Logik*, Akademie Verlag, Berlin, 1988.

[Grä03] G. Grätzer, *General Lattice Theory*, Birkhäuser Verlag, 2003.

[GSE+98] M. A. Gara, R. C. Silver, J. I. Escobar, A. Holman, and H. Waitzkin, *A Hierarchical Classes Analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms*, Psychiatry Research **81** (1998), 77–86.

[GV93] A. Guenoche and I. Van Mechelen, *Galois approach to the induction of concepts*, Cognitive Science Series, ch. 12, pp. 287–308, Academic Press, London, 1993.

[GW96] B. Ganter and R. Wille, *Formale Begriffsanalyse: Mathematische Grundlagen*, Springer, Berlin, Heidelberg, 1996.

[Háj98] P. Hájek, *The Metamathematics of Fuzzy Logic*, Kluwer, 1998.

[Háj01] ———, *On very true*, Fuzzy Sets and Systems **124** (2001), no. 3, 329–333.

[Hey30] A. Heyting, *Die formale Regeln der intuitionistischen Logik*, Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin, 1930, pp. 42–56.

[Hwa07] S. H. Hwang, *A triadic approach of Hierarchical Classes Analysis on folksonomy mining*, International Journal of Computer Science and Network Security **7** (2007), no. 8.

[JHS+06] R. Jäschke, A. Hotho, C. Schmitz, B. Ganter, and G. Stumme, *Trias - an algorithm for mining Iceberg tri-lattices*, International Conference on Data Mining (Hong Kong), IEEE Computer Society, December 2006, pp. 907–911.

[JYP88] D. S. Johnson, M. Yannakakis, and Ch. H. Papadimitriou, *On generating all maximal independent sets*, Information Processing Letters **27** (1988), no. 3, 119–123.

[Kep06] A. Keprt, *Algorithms for Binary Factor Analysis*, Ph.D. thesis, Technical University of Ostrava, 2006.

[Kim82] K. H. Kim, *Boolean matrix theory and applications*, 1 ed., CRC Press, 1982.

[KMP00] E. P. Klement, R. Mesiar, and E. Pap, *Triangular norms*, 1 ed., Springer, 2000.

[KPR06] L. Kwuida, Ch. Pech, and H. Reppe, *Generalizations of Boolean algebras. An attribute exploration.*, Mathematica Slovaca **56** (2006), no. 2, 145–165.

[Kra03] S. Krajči, *Cluster based efficient generation of fuzzy concepts*, Neural Network World (2003), no. 5, 521–530.

[Kra04] ———, *The basic theorem on generalized concept lattice*, Concept Lattices and Their Applications (R. Bělohlávek and V. Snášel, eds.), 2004, pp. 25–33.

[Kro83] P. M. Kroonenberg, *Three-mode Principal Component Analysis. Theory and applications.*, Leiden: DSWO Press, 1983.

[KS97] R. Keefe and P. Smith (eds.), *Vagueness: A reader*, MIT Press, 1997.

[KS04] A. Keprt and V. Snášel, *Binary Factor Analysis with help of Formal Concepts*, Concept Lattices and Their Applications (R. Bělohlávek and V. Snášel, eds.), 2004.

[KSOG94] S. Krolak-Schwerdt, P. Orlik, and B. Ganter, *TRIPAT: a model for analyzing three-mode binary data*, Studies in Classification, Data Analysis, and Knowledge Organization (H. H. Bock, W. Lenski, and M. M. Richter, eds.), Information systems and data analysis, vol. 4, Springer, Berlin, 1994, pp. 298–307.

[KSVĎ02] S. Krajči, V. Snášel, P. Vojtáš, and D. Ďuráková, *Conceptual structures*, Arbeitstagung Allgemeine Algebra, 2002, pp. 34–45.

[KY95] G. J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic - theory and applications*, Prentice Hall P T R, Upper Saddle River, New Jersey 07458, 1995.

[Lin] C. M. Ling, *Representation of associative functions*, Publicationes Mathematicae Debrecen **12**, 189–212.

[Lon83] J. S. Long, *Confirmatory Factor Analysis/a preface to LISREL*, Quantitative Applications in the Social Sciences, vol. 33, Sage Publication, 1983.

[ŁT30] J. Łukasiewicz and A. Tarski, *Untersuchungen über den Aussagenkalkül*, Comptes Rendus de la Siciete des Sciences et des Letters de Varsovie **23** (1930), 1–21.

[Łuk18] J. Łukasiewicz, *Farewell lecture at Warsaw Univeristy*, March 7 1918.

[Łuk22] ———, *Zagadnienia prawdy (the problems of truth)*, Księga pamiątkowa XI zjazdu lekarzy i przyrodników polskich (1922), 84–87.

[LV01] I. Leenen and I. Van Mechelen, *An evaluation of two algorithms for Hierarchical Classes Analysis*, Journal of Classification **18** (2001), 57–80.

[LVD01] I. Leenen, I. Van Mechelen, and P. De Boeck, *Models for ordinal Hierarchical Classes Analysis*, Psychometrika (2001), 389–404.

[LVDR99] I. Leenen, I. Van Mechelen, P. De Boeck, and S. Rosenberg, *INDCLAS: A three-way Hierarchical Classes model*, Psychometrika **64** (1999), 9–24.

[LW95] F. Lehmann and R. Wille, *A Triadic Approach to Formal Concept Analysis*, International Conference on Conceptual Structures (G. Ellis, R. Levinson, W. Rich, and J. F. Sowa, eds.), Lecture Notes in Computer Science, vol. 954, Springer, 1995, pp. 32–43.

[MA75] E. H. Mamdani and S. Assilian, *An experiment in linguistic synthesis with a fuzzy logic controller*, International Journal of Man-Machine Studies **7** (1975), no. 1, 1–13.

[Mac96] H. MacColl, *The calculus of equivalent statements*, London Mathematical Society, 1896, pp. 156–183.

[Mar94] R. J. Marks II, *Fuzzy logic technology and applications*, IEEE, 1994, Institute of Electrical and Electronics Engineers. Technical Activities Board, 1994.

[Men42] K. Menger, *Statistical Metrics*, Proceedings of the National Academy of Sciences of the United States of America **28** (1942), no. 12, 535–537.

[Mie11] P. Miettinen, *Boolean tensor factorization*, ICDM 2011, 11th IEEE International Conference on Data Mining (Vancouver, Canada) (D. Cook, J. Pei, W. Wang, O. Zaïane, and X. Wu, eds.), IEEE Computer Society, CPS, 2011, pp. 447–456.

[MME90] M. R. Mickey, P. Mundle, and L. Engelman, *Boolean Factor Analysis*, BMDP Manual, vol. 2, ch. 8M, University of California Press, 1990.

[MMG+08] P. Miettinen, T. Mielikäinen, A. Gionis, G. Das, and H. Mannila, *The discrete basis problem*, IEEE Transactions on Knowledge and Data Engineering **20** (2008), no. 10, 1348–1362.

[Obi12] S. A. Obiedkov, *Modeling preferences over attribute sets in Formal Concept Analysis*, International Conference on Formal Concept Analysis (F. Domenach, D. I. Ignatov, and J. Poelmans, eds.), Lecture Notes in Computer Science, vol. 7278, Springer, 2012, pp. 227–243.

[OK10] P. Osicka and J. Konecny, *General approach to Triadic Concept Analysis*, Concept Lattices and Their Applications (M. Kryszkiewicz and S. A. Obiedkov, eds.), CEUR Workshop Proceedings, CEUR-WS.org, 2010, pp. 116–126.

[Out10] J. Outrata, *Boolean Factor Analysis for data preprocessing in machine learning*, International Conference on Machine Learning and Applications (S. Draghici, T. M. Khoshgoftaar, V. Palade, W. Pedrycz, M. A. Wani, and X. Zhu, eds.), IEEE Computer Society, 2010, pp. 899–902.

[Pav79] J. Pavelka, *On fuzzy logic I, II, III*, Zeitschrift für mathematische Logik und Grundlagen der Mathematik **25** (1979), 45–52, 119–134, 447–464.

[Pei85] Ch. S. Peirce, *On the algebra of logic: A contribution to the philosophy of notation*, American Journal of Mathematics **7** (1885), no. 2, 180–196.

[Pei31] ———, *Collected papers*, Harvard Univ. Press, Cambridge, 1931.

[Pol97] S. Pollandt, *Fuzzy Begriffe*, Springer Verlag, Berlin Heidelberg New York, 1997.

[RG] S. Rosenberg and M. A. Gara, *The multiplicity of personal identity*, Review of Personality and Social Psychology **6**, 87–113.

[RK10] A. Revenko and S. O. Kuznetsov, *Attribute exploration of properties of functions on ordered sets*, Concept Lattices and Their Applications (M. Kryszkiewicz and S. A. Obiedkov, eds.), CEUR Workshop Proceedings, CEUR-WS.org, 2010, pp. 313–324.

[RM88] K. V. S. V. N. Raju and A. K. Majumdar, *Fuzzy functional dependencies and lossless join decomposition of fuzzy relational database systems*, ACM Transactions on Database Systems **13** (1988), no. 2, 129–166.

[Ros88] S. Rosenberg, *Self and others: Studies in social personality and autobiography*, Advances in Experimental Social Psychology **21** (1988), 57–95.

[Rus23] B. Russell, *Vagueness*, Australasian Journal of Philosophy **1** (1923), 84–92.

[Sac01] Ch. Sacarea, *Towards a theory of contextual topology*, Ph.D. thesis, TH Darmstadt, Aachen, 2001.

[SBG04] A. Smilde, R. Bro, and P. Geladi, *Multi-way analysis: Applications in the chemical sciencebroge*, Wiley, 2004.

[SBNP11] Ch. Sacarea, S. Boldeanu, L. Neamtiu, and C. Pastrav, *Investigating adverse drug reactions using conceptual landscapes*, IEEE International Conference on Intelligent Computer Communication and Processing, 2011, pp. 41–48.

[SS60] B. Schweizer and A. Sklar, *Statistical metric spaces*, Pacific Journal of Mathematics **10** (1960), 313–334.

[SSU03] A. I. Schein, L. K. Saul, and L. H. Ungar, *A generalized linear model for principal component analysis of binary data*, International Workshop on Artificial Intelligence and Statistics, 2003.

[Stu96] G. Stumme, *Attribute exploration with background implications and exceptions*, Data Analysis and Information Systems. Statistical and Conceptual approaches. Proc. GfKl'95. Studies in Classification, Data Analysis, and Knowledge Organization 7 (Heidelberg) (H.-H. Bock and W. Polasek, eds.), Springer, 1996, pp. 457–469.

[Stu99] ______, *Acquiring expert knowledge for the design of conceptual information systems*, Knowledge Acquisition, Modeling, and Management (D. Fensel and R. Studer, eds.), Lecture Notes in Computer Science, vol. 1621, Springer, 1999, pp. 275–290.

[SW91] N. Spangenberg and K. E. Wolff, *Comparison between Biplot Analysis and Formal Concept Analysis of repetory grids*, Classification, data analysis and knowledge organization, Berlin-Heidelberg: Springer, 1991, pp. 104–111.

[TMGM06] N. Tatti, T. Mielikäinen, A. Gionis, and H. Mannila, *What is the dimension of your binary data?*, IEEE International Conference on Data Mining, IEEE, dec 2006, pp. 603–612.

[TT06] F. Tang and H. Tao, *Binary Principal Component Analysis*, The British Machine Vision Conference, 2006, doi:10.5244/C.20.39, pp. 39.1–39.10.

[Tuc66] L. R. Tucker, *Some mathematical notes on Three-mode Factor Analysis*, Psychometrika **31** (1966), 279–311, 10.1007/BF02289464.

[VD89] I. Van Mechelen and P. De Boeck, *Implicit taxonomy in psychiatric diagnosis: A case study*, Journal of Social and Clinical Psychology **8** (1989), 276–287.

[VLC07] I. Van Mechelen, I. Lombardi, and E. Ceulemans, *Hierarchical classes modeling of rating data*, Psychometrika **72** (2007), 475–488.

[Wil82] R. Wille, *Restructuring lattice theory: an approach based on hierarchies of concepts*, Ordered Sets (Dordrecht/Boston) (I. Rival, ed.), Reidel, 1982, pp. 445–470.

[Wil87] ———, *Bedeutungen von Begriffsverbänden*, Beiträge zur Begriffsanalyse (B. Ganter, R. Wille, and K. E. Wolff, eds.), B.I.–Wissenschaftsverlag, Mannheim, 1987, pp. 161–211.

[Wil95] ———, *The basic theorem of Triadic Concept Analysis*, Order **12** (1995), 149–158.

[Yan82] M. Yannakakis, *The complexity of the partial order dimension problem*, SIAM Journal on Matrix Analysis and Applications **3** (1982), no. 3, 351–358.

[YJ01] S. B. Yahia and A. Jaoua, *Discovering knowledge from fuzzy concept lattice*, Data mining and computational intelligence (2001), 167–190.

[Zad65] L. A. Zadeh, *Fuzzy sets*, Information and Control **8** (1965), 338–353.

[Zad75a] ———, *The concept of a linguistic variable and its application to approximate reasoning*, Information Sciences **8** (1975), 199–245,301–357.

[Zad75b] ———, *Fuzzy logic and approximate reasoning*, Synthese **30** (1975), 407–428.

[Zad96] ———, *Fuzzy sets, fuzzy logic, and fuzzy systems*, World Scientific Publishing, River Edge, NJ, USA, 1996, pp. 195–209.

[Zic91] M. Zickwolff, *Rule exploration: First order logic in Formal Concept Analysis*, Technische Hochschule Darmstadt, 1991.

[ZV06] Z. Zivkovic and J. J. Verbeek, *Transformation invariant component analysis for binary images*, IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 254–259.

Affirmation

(a) Hereby I affirm that I wrote the present thesis without any inadmissible help by a third party and without using any other means than indicated. Thoughts that were taken directly or indirectly from other sources are indicated as such. This thesis has not been presented to any other examination board in this or a similar form, neither in this nor in any other country.

(b) The present thesis has been produced since October 2009 at the Institut für Algebra, Department of Mathematics, Faculty of Science, TU Dresden under the supervision of Prof. Dr. Bernhard Ganter.

(c) There have been no prior attempts to obtain a PhD at any university.

(d) I accept the requirements for obtaining a PhD (Promotionsordnung) of the Faculty of Science of the TU Dresden, issued February 2, 2011 with the changes in effect since June 15, 2011.

Versicherung

(a) Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

(b) Die vorliegende Dissertation wurde seit Oktober 2009 am Institut für Algebra, Fachrichtung Mathematik, Fakultät Mathematik und Naturwissenschaften, Technische Universität Dresden unter der Betreuung von Prof. Dr. Bernhard Ganter angefertigt.

(c) Es wurden zuvor keine Promotionsvorhaben unternommen.

(d) Ich erkenne die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften der TU Dresden vom 2. Februar 2011, in der geänderten Fassung mit Gültigkeit vom 15. Juni 2011 an.

Dresden, October 8, 2012