UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística e Computação Científica

ABNER DE MATTOS BRITO

Concerning Formal Concept Analysis on Complete Residuated Lattices

Sobre Análise de Conceitos Formais em Reticulados Residuados Completos

Campinas

2019

Abner de Mattos Brito

Concerning Formal Concept Analysis on Complete Residuated Lattices

Sobre Análise de Conceitos Formais em Reticulados Residuados Completos

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática Aplicada.

Dissertation presented to the Institute of Mathematics, Statistics and Scientific Computing of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Applied Mathematics.

Supervisor: Laécio Carvalho de Barros

Co-supervisor: Marcelo Esteban Coniglio

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Abner de Mattos Brito e orientada pelo Prof. Dr. Laécio Carvalho de Barros.

Campinas

2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

B777c
Brito, Abner de Mattos, 1991-
Concerning formal concept analysis on complete residuated lattices / Abner de Mattos Brito. – Campinas, SP : [s.n.], 2019.

Orientador: Laécio Carvalho de Barros.
Coorientador: Marcelo Esteban Coniglio.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Teoria dos reticulados. 2. Lógica simbólica e matemática. 3. Lógica fuzzy. 4. Teoria dos conjuntos. 5. Conjuntos fuzzy. I. Barros, Laécio Carvalho de, 1954-. II. Coniglio, Marcelo Esteban, 1963-. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Sobre análise de conceitos formais em residuados reticulados completos
Palavras-chave em inglês:
Lattice theory
Symbolic and mathematical logic
Fuzzy logic
Set theory
Fuzzy sets
Área de concentração: Matemática Aplicada
Titulação: Mestre em Matemática Aplicada
Banca examinadora:
Laécio Carvalho de Barros [Orientador]
Walter Alexandre Carnielli
Renata Wassermann
Data de defesa: 28-03-2019
Programa de Pós-Graduação: Matemática Aplicada

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-0920-6801
- Currículo Lattes do autor: http://lattes.cnpq.br/4462666377529458

Dissertação de Mestrado defendida em 28 de março de 2019 e aprovada pela banca examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). LAÉCIO CARVALHO DE BARROS

Prof(a). Dr(a). WALTER ALEXANDRE CARNIELLI

Prof(a). Dr(a). RENATA WASSERMANN

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

To a certain stray dog I never got to know,
but without which this work would not be.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

In the development of this work, parts of the material here presented — particularly Section 3.1, together with other portions of text here exposed — have been already published (BRITO et al., 2018) .

I am very grateful to Laécio Barros and Esteban Coniglio, who have been my supervisors in the development of this work, and who have shared with me not only their mathematical knowledge but their experience in the form of wise advice. Furthermore, I would like to thank Estevão Laureano who has provided many valuable insights that allowed the development of this work.

I am indebted to Fábio Bertato, who has encouraged me to enter the Masters in Applied Mathematics programme at Unicamp together with Masters in Philosophy programme, which I am currently developing at the Centre for Logic, Epistemology and the History of Science (CLE-Unicamp), where Fábio is my co-supervisor. He is the one who has proposed the development of a work related to Formal Concept Analysis. Over the years Fábio has provided me with guidance not only with in academic matters but in other subjects of life.

There are many people whom I met through my studies and who have helped me in one way or another, but I would especially like to thank some friends, namely Laryssa Abdalla, Melissa Carvalho and Joyce Clímaco (IMECC), as well as João Toniolo and Nicola Salvatore (CLE). Although we didn't spend so much time together as I would like over these last two years, I consider each of you a friend. It is always great to hear that you are well in spite of all difficulties that life brings upon us. I hope that I can aid each of you in difficult times.

I am verily thankful to Marcelo Cabral, Priscila Akemi, Altamiro Menezes and Ricardo Silva for your friendship and spiritual guidance. The Lord has used each of you to relief my burden in times of distress, and to make my joy complete when life is merry. I hope I can ever repay you for what you did for me, but I am sure that in the right time the Lord will.

There are so many people from CEU who have been by my side through this time that I am afraid I will forget some names, but I will give it a try: Alexandre, "tia" Alice, Aline Muzio, Amadeo, Bia, Bruna, Felipe Macedo, Fernando (aka Ferpa), Laura, Maira, Marília, Marina, Murilo, Natã, Neusa, Quesia, Romulo, Simone. Thank you all for being at my side and remembering me in your prayers.

Among all these friends, I specially thank Gesiel and Mateus. Now we share a roof, but you shared your lives with me prior to that. Together we have thouhgt and contemplated, walked and sat down, traveled and stayed, talked and made silence, laughed and cried, listened and prayed. As Lewis writes in *The Weight of Glory*, you are signs of Grace to my senses for in you “Christ *vere latitat* [...] Glory Himself is truly hidden." Much less important than this, thanks Mateus for helping me to prove a result I was stuck on.

Thank you so much Letícia for being at my side over the last few months. You have supported me, allowed me to be myself, taken care of me when my health was poor, cheered me up when I didn’t have heart to proceed. It feels like every meal at your side tastes better. Thank you for praying with/for me. The Lord knows our hearts and He knows everything else. Surely He will guide us according to His will in all things.

I thank my father Edney, my mother Cláudia and my brother Willyam, who have always supported me in so many ways: financially, in conversations and through your prayers. Without you I wouldn’t be who I am or where I am. Thank you so much for your unconditional love.

Finally, I praise my Lord Christ Jesus, together with the Father and the Holy Spirit, One True God. Unto Him be all the glory now and evermore!

Resumo

Análise de Conceitos Formais (FCA) é uma teoria matemática elegante, intimamente ligada à teoria de reticulados completos. Conforme exibido no presente trabalho, a teoria pode ser generalizada, permitindo-nos trabalhar com contextos formais nos quais objetos e atributos estão relacionados entre si apenas parcialmente e, de fato, o "grau"de tal relação pode ser extraído de reticulados residuados completos arbitrários.

O presente texto propõe-se a ser uma introdução concisa, intuitiva e autocontida à FCA Fuzzy. Para tanto, o leitor é introduzido às ideias de ordem, reticulados (completos) e operadores de fecho em tais reticulados. Estas — juntamente com noções de teoria ingênua dos conjuntos — formam a base necessária para a compreensão de FCA clássica. Posteriormente introduzimos a teoria de conjuntos fuzzy que, juntamente com a teoria de conectivos fuzzy (em particular t-normas e seus resíduos), provê as ferramentas necessárias para se estender a teoria clássica. Acreditamos que os exemplos apresentados ao longo do texto compõem um guia esclarecedor para a intuição.

Palavras-chave: Análise de conceitos formais. Reticulado de conceitos. Análise de conceitos formais fuzzy. Conceitos formais fuzzy. Reticulado de conceitos fuzzy.

Abstract

Formal Concept Analysis (FCA) is an elegant mathematical theory closely related to the theory of complete lattices. As shown in this work the theory can be generalized, allowing us to work with formal contexts in which objects and attributes are only partially related to each other and, in fact, the "degree" of such a relationship may be drawn from arbitrary complete residuated lattices.

The present text aims to be a concise, intuitive and self-contained introduction to fuzzy FCA. In order to reach this objective, the reader is introduced to the ideas of orders, (complete) lattices and closure operators on such lattices. That — together with notions of naïve set theory — is the necessary basis for an understanding of classical FCA. Afterwards, we introduce the theory of fuzzy sets which, together with fuzzy connectives (particularly t-norms and their residua), provides us with the tools necessary to extend the classical theory. It is our belief that the examples provided throughout the text have been illuminating as a guide for intuition.

Keywords: Formal concept analysis. Concept lattice. Fuzzy formal concept analysis. Fuzzy formal concepts. Fuzzy concept lattice.

List of Figures

Figure 1 – Isomorphic Lattices . . . 31
Figure 2 – Concept Lattice of Animals . . . 42
Figure 3 – Complexity of the Algorithms . . . 46
Figure 4 – Membership to the Fuzzy Set of Young People . . . 51
Figure 5 – Operations with Fuzzy Subsets . . . 53
Figure 6 – Triangular norms . . . 58
Figure 7 – Triangular conorms . . . 60
Figure 8 – Fuzzy negations . . . 62
Figure 9 – Fuzzy implications . . . 65
Figure 10 – Sequences . . . 69
Figure 11 – Output (left) and Lattice Diagram (right) of Alg. 5 applied to Table 3. 94
Figure 12 – Medical Diagnosis: Gödel's implication . . . 97
Figure 13 – Number of concepts with multivalued Łukasiewicz implications . . . 98
Figure 15 – Medical Diagnosis: 5-valued Łukasiewicz implication . . . 100

List of Tables

Table 1 – A formal context of animals 38
Table 2 – Truth Tables of Classical Connectives 57
Table 3 – Revisiting example 4.1.1. 93
Table 4 – Clinical signs observed for diagnosis of pneumonia 95
Table 5 – Fuzzy relation of clinical signs and diagnoses 96
Table 6 – Relation of clinical signs and diagnoses, with values in L_5 97

List of abbreviations and acronyms

iff	If and only if

List of symbols

χ_A	Characteristic function of the set A
μ_A	Membership function of the fuzzy set A — formally μ_A *is* the set A
2^Y	Collection of functions on Y mapped to $\{0, 1\}$
L^Y	Collection of functions on Y mapped to L
$\mathcal{P}(Y)$	Collection of the subsets of Y — bijective with 2^Y
$\mathcal{F}(Y)$	Collection of the fuzzy subsets of Y — formally the same as L^Y
$A \times B$	Cartesian product of A and B
$A_1 \times ... \times A_n$	Cartesian product of the sets $A_1, ..., A_n$
$A \subseteq B$	A is a subset of B, possibly equal to B
$A \subset B$	A is a proper subset of B, i. e., $A \subseteq B$ but $A \neq B$
$\bigwedge S$	Infimum (greatest lower bound) of the set S, also denoted by $\inf S$
$\bigvee S$	Supremum (least upper bound) of the set S, also denoted by $\sup S$
$[a, b]$	Closed interval of the elements $a \leqslant x \leqslant b$
$]a, b[$	Open interval of the elements $a < x < b$
$[a, b[$	Half-open interval of the elements $a \leqslant x < b$
$]a, b]$	Half-open interval of the elements $a < x \leqslant b$
$\langle A, B \rangle$	Ordered pair with first coordinate A and second coordinate B
$\langle A, B, C \rangle$	Ordered triple with first coordinate A, second coordinate B and third coordinate C
$\langle A_\alpha \rangle_{\alpha \in J}$	A sequence indexed by J
(a, b)	Ordered pair, where we explicitly treat a, b as elements of given sets A, B
$(a_\alpha)_{\alpha \in J}$	A sequence indexed by J, where we explicitly treat a_α as an element A_α
$x \mapsto y$	An unnamed map is presented, with the rule that maps each x to a y

$\phi \longrightarrow \psi$	Classical (material) implication of ϕ and ψ
$x \Rightarrow y$	Fuzzy implication of x and y
$\neg\phi$	Classical negation of ϕ
$\nu(x)$	Fuzzy negation of x
$\phi \wedge \psi$	Classical conjunction ("and") of ϕ and ψ
$x \vartriangle y$	Triangular norm (or t-norm, or fuzzy conjunction) of x and y
$\phi \vee \psi$	Classical disjunction ("or") of ϕ and ψ
$x \triangledown y$	Triangular conorm (or t-conorm, or fuzzy disjunction) of x and y
$\langle \mathcal{O}, \mathcal{A}, I \rangle$	Formal context with object set $\mathcal{O}$, attribute set $\mathcal{A}$ and relation I, usually denoted by $\mathbb{C}$
$\langle \mathcal{O}, \mathcal{A}, I_{\mathrm{f}} \rangle$	Fuzzy formal context with object set $\mathcal{O}$, attribute set $\mathcal{A}$ and fuzzy relation I, usually denoted by $\mathbb{C}_{\mathrm{f}}$
$\mathfrak{B}(\mathbb{C})$	Collection of formal concepts on the formal context $\mathbb{C}$
$\mathfrak{B}_{\mathrm{f}}(\mathbb{C}_{\mathrm{f}})$	Collection of fuzzy formal concepts on the fuzzy formal context $\mathbb{C}_{\mathrm{f}}$
$\mathscr{C}(Y)$	Collection of closed (perhaps fuzzy) subsets of Y with respect to a given closure operator

List of Algorithms

Algorithm 1 – Standard Computation of Intents 43
Algorithm 2 – Next Intent 45

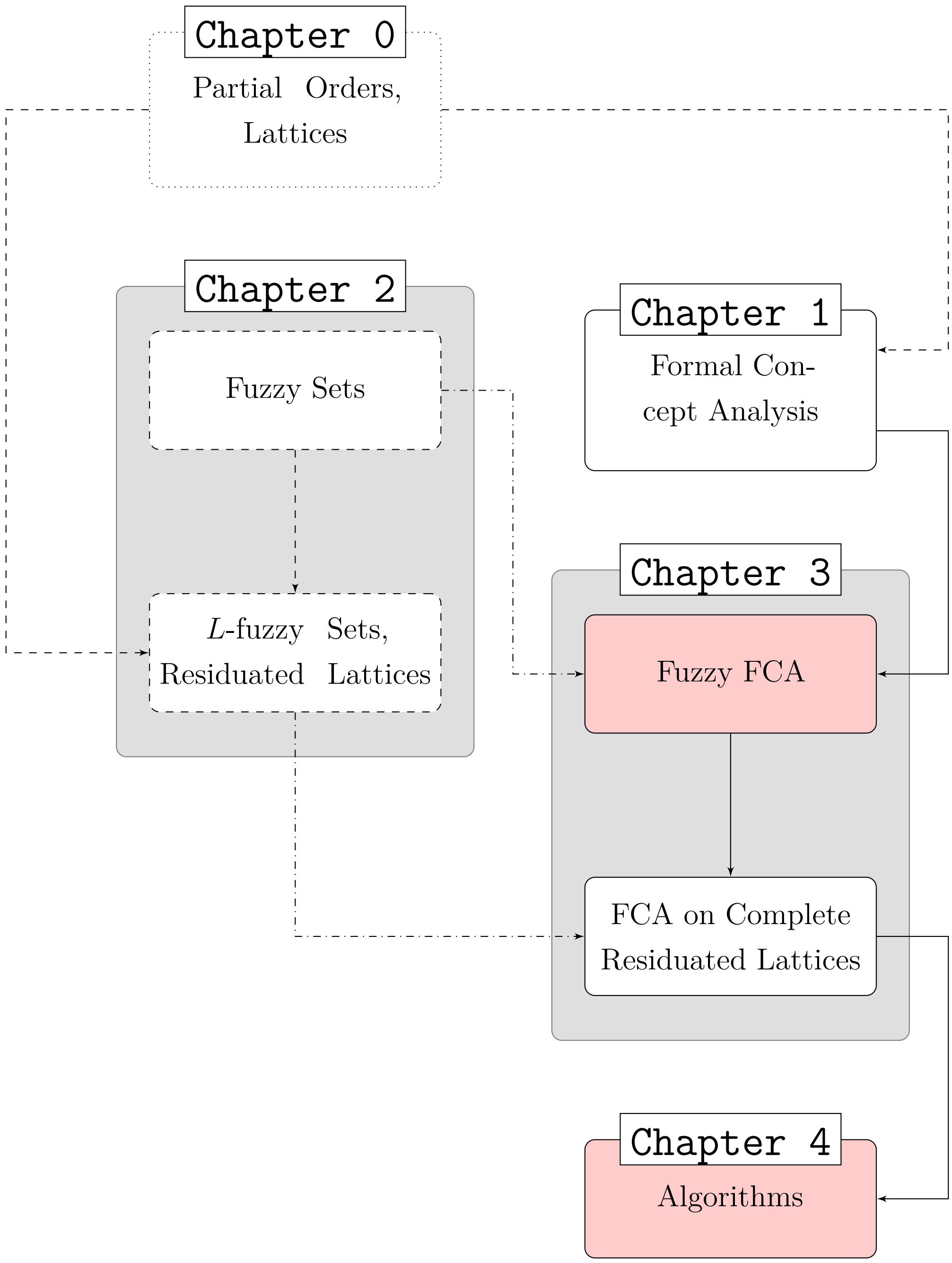
Algorithm 3 – UpperNeighbours 92
Algorithm 4 – GenerateFrom 93
Algorithm 5 – Lattice 93

Contents

	Topics Flowchart	**18**
	Preface	**19**
0	**ORDER THEORY AND LATTICES**	**21**
0.1	**Orders**	**21**
0.2	**Lattices**	**24**
1	**FORMAL CONCEPT ANALYSIS**	**37**
1.1	**Definitions and properties**	**37**
1.2	**Computing the concept lattice**	**43**
2	**FUZZY SETS AND FUZZY CONNECTIVES**	**49**
2.1	**Fuzzy sets**	**49**
2.2	**Fuzzy relations**	**53**
2.3	**Extending classical logic**	**56**
2.3.1	Logical connectives	56
2.3.2	Predicates and quantification	65
2.4	**Continuity and semicontinuity**	**67**
2.5	**Residuated Lattices**	**75**
3	**FUZZY FORMAL CONCEPT ANALYSIS**	**79**
3.1	**Formal concept analysis in the fuzzy setting**	**79**
3.2	**Formal concept analysis on complete residuated lattices**	**85**
3.2.1	The approach of Burusco and Fuentes-Gonzáles	85
3.2.2	The approach of Pollandt and Bělohlávek	86
4	**COMPUTING THE FUZZY CONCEPT LATTICE**	**89**
4.1	**Computational algorithms**	**89**
4.2	**An application of fuzzy FCA**	**94**
4.2.1	Posing the problem	94
4.2.2	Computing the fuzzy concept lattice	96
4.2.3	Interpreting the diagrams	99
4.2.4	Final considerations concerning applications	101
	Conclusion	**103**

BIBLIOGRAPHY 105

Subjects flowchart



Introduction

According to (JOSEPH, 2002), a *concept* is an abstraction produced by the intellect (e.g. the concept of *tree* abstracts what is common to several trees observed). A concept may then be communicated by a *term*, i.e., a *symbol* — a sound, a sequence of characters, an image etc. with a meaning imposed to it by convention — which expresses the concept (e.g., an image of a skull which stands for *danger*, or the sequence of characters T-R-E-E which stands for the notion of *tree*).

Also according to JOSEPH, a term has both *extension* and *intension*, the former corresponding to the set of all objects to which the term corresponds (e.g., the set of all trees), and the later, to "the sum of the essential characteristics that the term implies." Moreover, "As a term increases in intension, it decreases in extension." Given two concepts C_1 and C_2, if the extension of C_1 is a part of the extension of C_2 — i.e., if every object corresponding to C_1 also corresponds to C_2 — we say that C_1 is a *subconcept* of C_2.

Formal Concept Analysis (FCA), a mathematization of the philosophical idea of *concept* (GANTER; WILLE, 1999) , based on the idea — slightly different from JOSEPH's — that the *concept* that has extension and intension. In this text, we present an extension of FCA by grounding the theory on fuzzy sets. On the fuzzy setting computational algorithms are presented, which are applied to explore an example of medical diagnosis.

Because there is a hierarchy of concepts, in Chapter 0 basic notions of order theory are presented. We define orders, partially ordered sets, lattices covering relations and describe how a diagram may be drawn out of a lattice (or poset) in such a manner that the lattice (resp. poset) can be read back from the diagram. Furthermore, some basic notions concerning closure operators and Galois connections are presented.

In Chapter 1 the classical theory of FCA is introduced. Given a set of objects $\mathcal{O}$, a set of attributes $\mathcal{A}$ and a binary relation I on $\mathcal{O} \times \mathcal{A}$ — together they constitute a *formal context* —, we define two maps *, $^\wedge$, which are used to define *formal concepts*, which are composed of an *extent* and an *intent*. It turns out that these maps constitute a Galois connection between $\langle \mathcal{O}, \subseteq \rangle$ and $\langle \mathcal{A}, \subseteq \rangle$, and so $^{\wedge *}$ and $^{*\wedge}$ are closure operators. As a consequence, the collection of formal concepts of a given formal context is a complete lattice. An algorithm for computing every formal concept of a formal context is presented, and the idea of another algorithm, for computing the concepts together with the lattice structure is mentioned.

In order to extend FCA, fuzzy sets are introduced in Chapter 2, as well as fuzzy set operations (union, intersection and complementation). By means of fuzzy sets, we extend the classical connectives for conjunction ($\wedge$), disjunction ($\vee$), implication ($\longrightarrow$)

and negation ($\neg$) respectively to t-norms ($\triangle$), t-conorms ($\triangledown$), fuzzy implications ($\Rightarrow$) and fuzzy negations (ν). In particular, we consider $\Rightarrow$ as the residuum of a t-norm. We prove some properties concerning lower semicontinuous t-norms. Some basic notions of residuated lattices are also presented.

Chapter 3 deals with the extension of FCA to fuzzy sets. In Section 3.1 show that, if we use lower semicontinuous t-norms and their residua in order to define maps * and $^\wedge$, all the basic results of the classical theory as presented in Ch. 1 are preserved. The results have been found independently by us, but it turns out that they fall within the scope of a more general theory of FCA on residuated lattices. The idea of this more general approach is presented, alongside with another approach — pioneer in generalizing FCA to (L-)fuzzy sets, but without some useful properties.

Finally, Chapter 4 deals with the problem of computing the fuzzy concept lattice when we consider the set L of truth values as a finite subset of $[0, 1]$ — actually, the algorithm presented applies to any finite linearly ordered set. The algorithm is used to explore an example of the relation between clinical signs and pneumonia *vs.* non-pneumonia diseases in children.

We have attempted to write a self-contained text, accessible to people who understand mathematical reasoning and who have an intuitive knowledge of sets, functions, real numbers etc., as well as the notion of what is an algorithm. No rigorous, axiomatic, knowledge of such notions is required. Nonetheless, it is certainly helpful that the reader understands what are the supremum and infimum of (limited nonempty) subsets of $\mathbb{R}$, and why these may not exist for (limited nonempty) subsets of $\mathbb{Q}$. Also useful — particularly for some proofs in the auxiliary Section 2.4 is basic knowledge related to sequences.

0 Order Theory and Lattices

The idea of *concept* shall be (briefly) presented later in this text, but we mention from the start that there is a *hierarchy* of concepts. For example, the concept of *animal* is broader than that of *human* in the sense that the former includes all individuals of the later — and in fact other, non-human animals. On the other hand, although the concept of *dog* is narrower than that of *animal* it is neither broader nor narrower than the concept of *human*. In terms of order theory, we say that this is "hierarchy" is a nonlinear order. Orders will be extensively used in this text, and for this reason we present some basic notions of order theory.

The content of the present chapter has been drawn mostly from (GRÄTZER, 1971) and (BIRKHOFF, 1948) , to which we refer the interested reader. Some of the content of this chapter (and others) corresponds to set theory[1], and can be found in any good textbook on the field, such as (HRBÁČEK; JECH, 1999) .

0.1 Orders

Definition 0.1.1. A *n-place relation* on a Cartesian product of sets $A_1 \times ... \times A_n$ is a subset $R \subseteq A_1 \times ... \times A_n$. If $A := A_1 = ... = A_n$ we say that R is a n-place relation *on* A. In particular, a 2-place relation is called a *binary relation*.

It is common practice to use infix notation for binary relations, i.e., $(a_1, a_2) \in R$ is expressed as $a_1 R a_2$. Given $S_2 \subseteq S_1$ and a binary relation R_1 on S_1, the *restriction* of R_1 to S_2 is the subset R_2 of R_1 defined as the intersection of R_1 and $S_2 \times S_2$, that is,

$$R_2 = \{(x, y) \in R_1 : x, y \in S_2\} \ .$$

Example 0.1.2. Given a set S, an *equivalence relation* on S is a set $R \subseteq S^2$ that is reflexive (aRa), symmetric (aRb implies bRa) and transitive (aRb and bRc imply aRc). For instance, we say that $a, b \in \mathbb{Z}$ are *congruent modulo* n, denoted by $a \equiv b \mod n$, if $a - b = kn$ for some $k \in \mathbb{Z}$. Congruence modulo n is an equivalence relation on $\mathbb{Z}$ for all $n \in \mathbb{Z}$.

A very special class of binary relations on a given set is that of order relations.

Definition 0.1.3. A *partial order*, or simply *order*, on a set X is a binary relation $\leqslant$ on X that satisfies, for all $x, y, z \in X$:

[1] We assume the ZFC axioms as presented by (HRBÁČEK; JECH, 1999).

1. $x \leqslant x$ (reflexivity)
2. if $x \leqslant y$ and $y \leqslant x$ then $x = y$ (antisymmetry)
3. if $x \leqslant y$ and $y \leqslant z$ then $x \leqslant z$ (transitivity)

An order is called *linear* or *total* if it additionally satisfies

4. $x \leqslant y$ or $y \leqslant x$ (linearity)

A set P equipped with a partial order $\leqslant$, denoted by $\langle P, \leqslant\rangle$, is called a *partially ordered set*, or *poset*. If the order is linear, we say to have a *linearly ordered set*, a *totally ordered set* or a *chain*. We may refer to $\langle P, \leqslant\rangle$ simply as P if there is no risk of confusion concerning the order $\leqslant$. Two elements $a, b \in P$ are said to be *comparable* if either $a \leqslant b$ or $b \leqslant a$. Otherwise, a and b are *incomparable*. Clearly there are incomparable elements in a poset iff it is not a chain.

The binary relation $<$ defined as

$$a < b \text{ iff } a \leqslant b \text{ and } a \neq b \tag{0.1.4}$$

is called a *strict order*. Strict orders can be defined on their own rights rather than being defined in terms of partial orders.

Proposition 0.1.5. *A binary relation $<$ on X is a strict order iff, for all $x, y, z \in X$,*

1. *$x < x$ does not hold* *(irreflexivity)*
2. *if $x < y$ then $y < x$ does not hold* *(asymmetry)*
3. *if $x < y$ and $y < z$ then $x < z$* *(transitivity)*

Proof. We first prove that strict orders are irreflexive, asymmetric and transitive.

1. If we had $x < x$ then we would also have $x \neq x$ by (0.1.4), which is a contradiction. Thus, $x < x$ does not hold.
2. Suppose for the sake of contradiction that both $x < y$ and $y < x$. Then by (0.1.4) we would have $x \leqslant y$ and $y \leqslant x$, implying (by antisymmetry of $\leqslant$) that $x = y$, which would contradict $x < y$. Thus, $x < y$ and $y < x$ cannot both hold.
3. Suppose that $x < y$ and $y < z$. Then $x \leqslant y$ and $y \leqslant z$, whence $x \leqslant z$. Furthermore, were $x = z$ we would have $x < y$ and $y < x$, contradicting asymmetry. Thus, $x < z$.

On the other hand, suppose that $<$ is an irreflexive, asymmetric and transitive relation. Define $\preceq$ by

$$x \preceq y \text{ iff } x < y \text{ or } x = y \ . \tag{0.1.6}$$

Then $x < y$ iff $x \preceq y$ and $x \neq y$[2]. Now it is clear from (0.1.4) that $\prec$ and $\leqslant$ are the same relation. □

To finish the discussion concerning strict orders, we highlight the fact that proof of Proposition 0.1.5 tells us that there is a one-one correspondence between partial orders and strict orders, in such a way that one uniquely defines the other, i.e., a partial order $\leqslant$ defines a strict order $<$ by means of (0.1.4), a strict order $<$ defines a partial order $\prec$ as in (0.1.6) and $\leqslant$, $\prec$ are the same.

The *dual relation* of $\leqslant$ is $\geqslant$ defined as $x \leqslant y$ iff $y \geqslant x$, which is easily seen to be an order[3]. The reader should see that the dual of $\geqslant$ is $\leqslant$. Correspondingly, the dual of $<$ is $>$, the strict order defined from $\geqslant$. The *dual* of a statement[4] about posets is the statement we get by replacing every occurrence of an order ($\leqslant$ or $\geqslant$) by its dual ($\geqslant$ or $\leqslant$, respectively). Thus we are able to state a nice, work saving principle.

Duality Principle. *If a statement is true in all posets, then its dual is also true in all posets.*

Indeed, consider one such statement Φ, and let $\langle P, \leqslant \rangle$ be any poset. Because Φ is true in all posets, it is also true in $\langle P, \geqslant \rangle$. Replacing each occurrence of $\geqslant$ by $\leqslant$ yields the dual Φ^{d} of Φ, and Φ^{d} is true in $\langle P, \leqslant \rangle$.

Example 0.1.7. Consider the following.

1. The sets $\mathbb{N}, \mathbb{Q}$ and $\mathbb{R}$, equipped with their respective usual orders, are each a linearly ordered set with their respective usual orders.
2. Let S be a non-empty set and let $\mathcal{P}(S)$ be the set of its subsets (informally, the "power set" of S). Then $\langle \mathcal{P}(S), \subseteq \rangle$ is a poset, where $\subseteq$ is set inclusion[5]. Moreover, if S has at least two elements then $\mathcal{P}(S)$ is nonlinear, as $\{a\}, \{b\} \in \mathcal{P}(S)$ are incomparable with respect to set inclusion whenever $a \neq b$.

[2] If $x < y$ then it is true that $x < y$ or $x = y$, whence $x \preceq y$. Conversely, if $x \preceq y$ then $x < y$ or $x = y$ whence, if additionally we have $x \neq y$, we must have $x < y$.

[3] Check for yourself that $\geqslant$ is reflexive, antisymmetric and transitive, and that if $\leqslant$ is linear so is $\geqslant$.

[4] A statement can be given a precise, rigorous meaning which forms the base of both Statement and First Order Predicate Calculi. These are used to work out the foundations of mathematics in logical terms. For more on the Statement Calculus and First Order Predicate Calculus, See (MENDELSON, 2009) .

[5] Because an order on a set S is a subset of S^2, whenever we consider set inclusion as an order we restrict ourselves to the order $\subseteq_{\mathcal{P}(S)}$ on $\mathcal{P}(S)$, i.e.:

$$A \subseteq_{\mathcal{P}(S)} B \text{ iff } A \subseteq S \text{ and } B \subseteq S \text{ and } A \subseteq B \ .$$

But whenever we use $\subseteq_{\mathcal{P}(S)}$, the power set of S is already implicit in the context, so that we write $\subseteq$ for brevity.

3. Let $S = \{a, b, c\}$. Let $\leqslant_1 = \{(a, a), (a, c), (b, b), (b, c), (c, c)\}$[6] and let $\leqslant_2$ be the dual order of $\leqslant_1$ (i.e., replace (a, c) by (c, a) and (b, c) by (c, b)). Then both $\langle S, \leqslant_1\rangle$ $\langle S, \leqslant_2\rangle$ are posets.
4. Consider $\left\langle \mathbb{R}^2, \leqslant_{\mathrm{p}}\right\rangle$, where $(x, y) \leqslant_{\mathrm{p}} (z, w)$ iff $x \leqslant_{\mathbb{R}} z$. It is reflexive, transitive but is not antisymmetric, and thus is not a poset. In fact, $(0, 0) \leqslant_{\mathrm{p}} (0, 1)$ and $(0, 1) \leqslant_{\mathrm{p}} (0, 0)$, but of course, $(0, 0) \neq (0, 1)$.
5. "Parenthood" on the set of human beings is (vacuously[7]) antisymmetric, but is irreflexive (no person is its own parent) and non-transitive (strictly speaking, a grandparent is not a parent).

0.2 Lattices

We now proceed to define the suprema and infima of subsets of a given set. Let P be a poset and let $H \subseteq P$. An element $a \in P$ is an *upper bound* of H if $h \leqslant a$ for all $h \in H$. A *least upper bound*, or *supremum*, of H is an upper bound s of H such that $s \leqslant a$ for every upper bound a of H. Antisymmetry of $\leqslant$ implies that the supremum s of H is unique, and we shall denote it as either of the following.

$$\sup H \ , \qquad \bigvee H \ ,$$

$$\bigvee_{h \in H} h \ , \qquad \bigvee_{i=1}^{n} h_i \ ,$$

where the last notation applies only if $H = \{h_1, ..., h_n\}$.

Notice that the supremum of H may or may not be an element of H. For instance, in $\mathbb{R}$ the supremum of both $A =]0, 1[$ and $B = [0, 1]$ is 1, which belongs to B but not to A.

The definition of a *lower bound* is dual to the definition of an upper bound, and the definition of the *infimum* of a set is dual to that of the *supremum* of a set. Notice that the infimum of a set is unique by the duality principle. Dual to the notations for the

[6] This means that $a \leqslant_1 a$, $a \leqslant_1 c$, $b \leqslant_1 b$, $b \leqslant_1 c$, $c \leqslant_1 c$ and nothing else satisfies $\leqslant_1$.

[7] Given some assertion about something, if one is to say that such an assertion is false one needs to present a counterexample. For instance, the statement "3 is the greatest natural number" is false because a counterexample may be presented, such as 4, a natural number greater than 3. Now, consider the statement "every negative natural number is greater than 8". Because there are no negative natural numbers, we cannot present one such number which is not greater than 8. Thus, we accept such a ridiculous statement as true — a rather vague truth — simply because it cannot be false. That is the reason why we accept as true every assertion about elements of the empty set, simply because there are no counterexamples. Such assertions are said to be *vacuously true*.

supremum, the infimum of H is denoted as either of

$$\inf H \ , \qquad \bigwedge H \ ,$$
$$\bigwedge_{h \in H} h \ , \qquad \bigwedge_{i=1}^{n} h_i \ .$$

Given $a, b \in P$, we shall denote $\sup\{a, b\}$ and $\inf\{a, b\}$ by $a \vee b$ and $a \wedge b$ respectively. In order theory it is usual to call $\wedge$ the *meet* and $\vee$ the *join*.

Let $\langle L, \leqslant \rangle$ be a poset. If for all $a, b \in L$ both $a \vee b$ and $a \wedge b$ exist, then $\langle L, \leqslant \rangle$ is called a *lattice*. The induction principle implies that this is equivalent to the existence of both infimum and supremum of arbitrary finite subsets of of L.

Definition 0.2.1. A poset $\langle L, \leqslant \rangle$ is a *lattice* iff both $\bigvee H$ and $\bigwedge H$ exist for any nonempty finite $H \subseteq L$. If $\bigvee H$ and $\bigwedge H$ exist for arbitrary (finite or infinite) nonempty $H \subseteq L$, we say that the lattice is *complete*. This lattice is said to be *finite* or *infinite* depending on whether L is finite or infinite.

Notation. If $\bigvee L$ exists (which is always the case for complete lattices), then it is denoted by 1 and it is the greatest element of L. Dually, if L has a smallest element $\bigwedge L$, it is denoted by 0.

In order to explicit the smallest and greatest elements of a complete lattice, it may be denoted as $\langle L, \leqslant, 0, 1 \rangle$.

Notice that every element of L is (vacuously) both an upper bound and a lower bound of the empty set $\varnothing$. Thus, in the lattice L, if 0 exists then $\bigvee \varnothing = 0$, and dually if 1 exists then $\bigwedge \varnothing = 1$.

Example 0.2.2.

1. Every finite lattice is complete by definition.
2. Given a nonempty set S, $\langle \mathcal{P}(S), \subseteq \rangle$ is a complete lattice with infimum given by set intersection and supremum given by set union.
3. Any closed interval of $\mathbb{R}$ is a complete lattice under its usual order, where infimum and supremum are defined as usual.
4. $\mathbb{R}$ is not a complete lattice because $\mathbb{R}$ itself has neither supremum nor infimum. However, if we consider

   $$\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty] \ ,$$

   then we have a complete lattice.
5. Let

   $$I_{\mathbb{Q}} = \{x \in \mathbb{Q} : 0 \leqslant x \leqslant 1\}$$

be ordered by the restriction $\leqslant_{I_{\mathbb{Q}}}$ of the usual order $\leqslant_{\mathbb{Q}}$ of $\mathbb{Q}$ to $I_{\mathbb{Q}}$. Then $I_{\mathbb{Q}}$ is a lattice, but not a complete lattice. In fact, the set

$$S = \{r \in I_{\mathbb{Q}} : 2r^2 \leqslant_{I_{\mathbb{Q}}} 2\}$$

has no supremum in $I_{\mathbb{Q}}$[8]. In fact, no interval $I \subset \mathbb{Q}$ is a complete lattice — although I is always a lattice under the restriction of $\leqslant_{\mathbb{Q}}$ to I — as we can always find a subinterval of I whose supremum is "absent", i.e., is not rational. The definition of $\mathbb{R}$ via Dedekind cuts[9] is made in order to literally fill the gaps that prevent intervals of $\mathbb{Q}$ from being complete lattices.

6. The posets in item 3. of Example 0.1.7 are not lattices. In fact, $\{a, b\}$ has no infimum with respect to $\leqslant_1$ and no supremum with respect to $\leqslant_2$.

Lattices can also be defined algebraically in terms of the meet and the join of elements of L. Although we shall not present the definition here, we draw the reader's attention to the fact that

$$a \leqslant b \text{ iff } a \wedge b = a \qquad (0.2.3)$$

$$\text{iff } a \vee b = b \ . \qquad (0.2.4)$$

Also, both $\vee$ and $\wedge$, considered as binary operations, are idempotent, commutative, associative and satisfy the following *absorption identities*:

$$a \wedge (a \vee b) = a \ ,$$
$$a \vee (a \wedge b) = a \ .$$

These properties suffice in order to define a lattice $\langle L, \vee, \wedge \rangle$ in a way that is equivalent to $\langle L, \leqslant \rangle$.

Now we turn to the question of whether it is possible to find a binary subrelation $\prec$ of $\leqslant$ such that all the information in $\leqslant$ can be recovered from $\leqslant$. As we shall see, the answer is "yes", and this will allow us to represent (finite) posets and lattices graphically.

A binary relation R on a set S can always be made reflexive by adding the diagonal of S^2 to it, that is, by taking the set

$$R_{\text{ref}} = R \cup \{(s, s) \in S^2 : s \in S\} \ .$$

8 Indeed, the supremum should be $\sqrt{2}/2 \notin \mathbb{Q}$.
9 See (HRBÁČEK; JECH, 1999) .

Similarly,

$$R_{\text{trans}} = R \cup \{(x, z) \in S^2 : xRy \text{ and } yRz \text{ for some } y \in S\}$$

is a transitive binary relation. Thus, a relation $\prec$ such as that of "Parenthood" (see Example 0.1.7 item 6.), which is antisymmetric but is neither reflexive nor transitive, could be made into an order $\prec_{\text{ord}}$ by first extending it to a reflexive relation, then to a transitive relation[10].

However, we would have a great difficulty to make an antisymmetric relation out of a relation which is presents symmetries, because if there are two distinct $x, y \in S$ such that

$$xRy \text{ and } yRx \quad ,$$

we would need to exclude either (x, y) or (y, x) from R, in which case information from the original relation would be lost[11]. Fortunately orders are already antisymmetric and, because symmetries do not arise by excluding pairs from a binary relation, we do not need to concern ourselves about such a process.

As seen above, a binary relation can always be made reflexive and transitive, so we now create a sub-relation $\prec$ of $\leqslant$ by excluding all the information about reflexivity and transitivity.

Definition 0.2.5. Let $\langle P, \leqslant\rangle$ be a finite poset. We define $\prec$ for each $a, b \in P$ by

$$a \prec b \text{ iff } \begin{cases} a < b \ , \text{ and} \\ \nexists c \in P \text{ such that } a < c < b \ . \end{cases} \tag{0.2.6}$$

The relation $\prec$ is called a *covering relation* on P (with respect to $\leqslant$). If $a \prec b$ we say that a is *covered* by b, or that a is a *lower neighbour* of b. Conversely, we say that b *covers* a, or that b is an *upper neighbour* of a.

Here, of course, $a < c < b$ is short for $a < c$ and $c < b$. We know that $\prec$ can be made into an order relation $\prec_{\text{ord}}$, but we need to check that $\prec_{\text{ord}}$ really is the same as $\leqslant$.

Suppose that $a < b$. If P is finite there exists a maximal chain $a = c_0 < c_1 < ... < c_{n-1} = b$, that is, there is no $c \in P$ such that $c_i < c < c_{i+1}$ for $i = 0, ..., n-2$. In fact, of all the (finitely many) subsets of P we may take all the (finite) chains H with smallest element a and greatest element b (at least one such H exists, namely $\{a, b\}$). From the finitely many such H we choose one of larger size and let n be the number of elements of

10 We would reach the same order if the "Parenthood" relation were extended first transitively and only then reflexively, because transitive extensions do not add new information for diagonal elements. In fact, for $x = y$ we have, in the extension equation above, that $(x, z) \in R_{\text{trans}}$ iff $(y, z) \in R$.

11 After this process, we would have no straightforward criteria for deciding which elements of the reduced relation we should need to make symmetric again.

H. We can now assume that $H = \{c_0, c_1, ..., c_{n-1}\}$ is such that $a = c_0 < c_1 < ... < c_{n-1} = b$. Now by maximality of H, there is no $c \in P$ such that $c_i < c < c_{i+1}$, for $i = 0, ..., n-2$ (otherwise there would be a chain larger than H with smallest and greatest elements a and b respectively). This proves the following:

Lemma 0.2.7. *Let $\langle P, \leqslant\rangle$ be a finite poset. Then $a \leqslant b$ iff $a = b$ or there exists a finite sequence $c_0, ..., c_{n-1}$ of elements of P such that*

$$a = c_0 \prec c_1 \prec ... \prec c_{n-1} = b \ ,$$

where $\prec$ is as defined in 0.2.6.

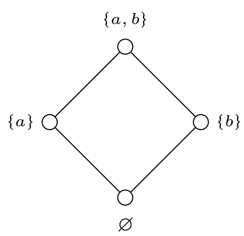
Given a poset $\langle P, \leqslant\rangle$, we say that a *covers* b (or b is *covered by* a) if $a \prec b$. For this reason, the relation $\prec$ is called the *covering relation* (or simply *covering*) corresponding to $\leqslant$. This lemma, holds true of lattices, which are a special case of posets.

Example 0.2.8. Consider the Example 0.1.7 item 3. The orders $\leqslant_1$ and $\leqslant_2$ have the following coverings respectively:

$$\prec_1 = \{(a, c), (b, c)\} \ ,$$
$$\prec_2 = \{(c, a), (c, b)\} \ .$$

The covering relation allows us to draw a diagram of a finite lattice (and of a finite poset in general). This diagram is such that there is a unique circle, called a *node*, corresponding to each element of L and, for each $a, b \in L$, $a \prec b$ iff the circle of a is below the circle of b and there is a line connecting them.

Example 0.2.9. 1. Consider the set $S = \{a, b\}$. Then $\mathcal{P}(S)$ ordered by set inclusion $\subseteq$ and has the following diagram:

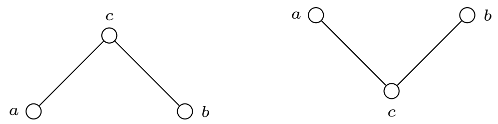


Nodes $\{a\}$ and $\{b\}$ are incomparable. $\varnothing$ is smaller than $\{a, b\}$ according to $\subseteq$ because there exists an upwards path connecting nodes $\varnothing$ and $\{a, b\}$, namely the one passing through $\{a\}$.

It has the covering

$$\prec = \{(\varnothing, \{a\}), (\varnothing, \{b\}), (\{a\}, \{a, b\}), (\{b\}, \{a, b\})\} \ .$$

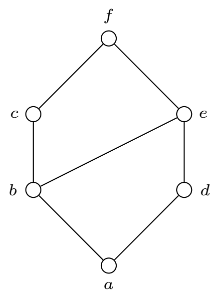
2. Consider the set $S = \{a, b, c\}$ with orders $\leqslant_1$ and $\leqslant_2$ given in Example 0.1.7 item 3. They have the following diagrams:



(a) $\langle S, \leqslant_1 \rangle$ (b) $\langle S, \leqslant_2 \rangle$

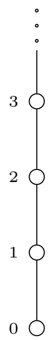
The coverings $\prec_1$ and $\prec_2$ have been given in Example 0.2.8.

3. Consider the lattice which corresponds to the following diagram:



Notice that, although node c is higher than node d, they are not comparable because it is not possible to, starting from node d, reach node c going strictly upwards on the diagram.

4. Using the same idea as in previous items, $\mathbb{N}$ with its usual order could be represented as



giving the additional information that the diagram continues upward with the elements given natural succession.

5. Neither $\mathbb{Q}$ nor $\mathbb{R}$ is well-ordered[12], and so they do not have covering relations. Nonetheless, they can be represented as usual with axes in which we point out the zero and other elements of interest.

[12] Intuitively, a set is well-ordered if, for each element, we know what is the "next" element. Because both $\mathbb{Q}$ and $\mathbb{R}$ are dense sets (i.e., given any two elements, there is another element between them), it is impossible to determine the "next" element to any rational or real number. For more on well-orders, See (HRBÁČEK; JECH, 1999) .

In many cases it happens that certain elements can be represented as the supremum of certain elements, or as their infimum. (GANTER; WILLE, 1999) present the following definitions.

Definition 0.2.10. Given a complete lattice $\langle L, \leqslant, \rangle$, consider for each $x \in L$ the elements

$$x_{\#} = \bigvee \{y \in L : y < x\} \ ,$$
$$x^{\#} = \bigwedge \{y \in L : x < y\} \ .$$

Then x is called *supremum-irreducible*, and denoted as $\bigvee$-irreducible, if $x \neq x_{\#}$. Dually, if $x \neq x^{\#}$ it is said *infimum-irreducible*, and denoted as $\bigwedge$-irreducible. The sets of all $\bigvee, \bigwedge$-irreducible elements of L are denoted $J(L)$ and $M(L)$[13], respectively, or simply J, M if there is no risk of confusion concerning L .

Definition 0.2.11. Sets $X, Y \subseteq L$ are respectively called *supremum-dense* and *infimum-dense* iff for all $z \in L$ we have respectively

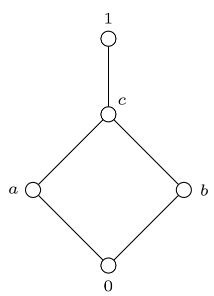
$$z = \bigvee \{x \in X : x \leqslant z\} \ ,$$
$$z = \bigwedge \{y \in Y : z \leqslant y\} \ .$$

It is clear that L is both a supremum-dense and an infimum-dense set. Nonetheless, in the finite case it is possible to find smaller sets which satisfy these properties, as stated in Prop. 0.2.12 proved by (GANTER; WILLE, 1999).

Proposition 0.2.12. *Let $\langle L, \leqslant \rangle$ be a finite lattice. An element $x \in L$ is $\bigvee$-irreducible (resp. $\bigwedge$-irreducible) iff it has exactly one lower neighbour (resp. upper neighbour).*

The sets J, M are respectively supremum-irreducible and infimum-irreducible and, if $X, Y \subseteq L$ are respectively a supremum-irreducible and an infimum-irreducible set, then $J \subseteq X$ and $M \subseteq Y$. Hence, J, M are minimal supremum-irreducible and infimum-irreducible sets, respectively.

Example 0.2.13. Consider the following lattice.



[13] As in *join* and *meet*.

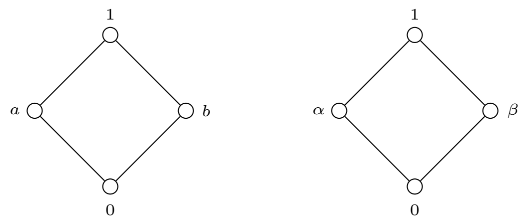


Figure 1 – Isomorphic Lattices

Nodes $a, b, 1$ have exactly one lower neighbour, and nodes a, b, c have exactly one upper neighbour. Hence,

$$J = \{a, b, 1\} \ , \qquad M = \{a, b, c\} \ .$$

Furthermore,

$$\begin{aligned}
0 &= \bigvee\{j \in J : j \leqslant 0\} = \bigvee \varnothing \ , & 0 &= \bigwedge\{m \in M : 0 \leqslant m\} = \bigwedge M \ , \\
a &= \bigvee\{j \in J : j \leqslant a\} = \bigvee\{a\} \ , & a &= \bigwedge\{m \in M : a \leqslant m\} = \bigwedge\{a, c\} \ , \\
b &= \bigvee\{j \in J : j \leqslant b\} = \bigvee\{b\} \ , & b &= \bigwedge\{m \in M : b \leqslant m\} = \bigwedge\{b, c\} \ , \\
c &= \bigvee\{j \in J : j \leqslant c\} = \bigvee\{a, b\} \ , & c &= \bigwedge\{m \in M : c \leqslant m\} = \bigwedge\{c\} \ , \\
1 &= \bigvee\{j \in J : j \leqslant 1\} = \bigvee J \ , & 1 &= \bigwedge\{m \in M : 1 \leqslant m\} = \bigwedge \varnothing \ .
\end{aligned}$$

Notice that it is also possible that $J(L) = M(L)$, depending on the structure of L[14].

There are certain circumstances under which two lattices may be considered to be "the same", i.e., they may have different elements, but have the same structure. That is what happens, for instance, with the lattices corresponding to the diagrams in Fig. 1.

Definition 0.2.14. Let $\mathcal{L}_1 = \langle L_1, \leqslant_1 \rangle$, $\mathcal{L}_2 = \langle L_2, \leqslant_2 \rangle$ be lattices. We say that $\mathcal{L}_1$ is *isomorphic* to $\mathcal{L}_2$ iff there exists a bijective function $f : L_1 \to L_2$ such that, for all $x, y \in L_1$,

$$x \leqslant_1 y \text{ iff } f(x) \leqslant_2 f(y) \ .$$

In this case, f is called an *isomorphism*.

Clearly $\mathcal{L}_1$ is isomorphic to $\mathcal{L}_2$ iff $\mathcal{L}_2$ is isomorphic to $\mathcal{L}_1$[15], so that we often say that $\mathcal{L}_1$ and $\mathcal{L}_2$ are isomorphic.

Example 0.2.15. 1. The lattices $\langle L_1, \leqslant_1 \rangle$, $\langle L_2, \leqslant_2 \rangle$, with

$$\begin{aligned}
L_1 &= \{a, b\} \ , & \leqslant_1 &= \{(a, a), (a, b), (b, b)\} \\
L_2 &= \{\alpha, \beta\} \ , & \leqslant_2 &= \{(\alpha, \alpha), (\alpha, \beta), (\beta, \beta)\}
\end{aligned}$$

[14] Consider Fig. 1.
[15] With isomorphism f^{-1}.

are isomorphic with isomorphism $f : L_1 \to L_2$ given by $f(a) = \alpha$, $f(b) = \beta$.

2. Let $S = \{\xi, \zeta\}$. Then the lattice $\langle \mathcal{P}(S), \subseteq \rangle$ is isomorphic to, say, the left lattice of Fig. 1 with wither of the following isomorphisms defined on $\mathcal{P}(S)$.

$$\begin{aligned} f(\varnothing) &= 0 \ , & g(\varnothing) &= 0 \ , \\ f(\{\xi\}) &= a \ , & g(\{\xi\}) &= b \ , \\ f(\{\zeta\}) &= b \ , & g(\{\zeta\}) &= a \ , \\ f(S) &= 1 \ , & f(S) &= 1 \ . \end{aligned}$$

3. Consider the following result[16].

   Let $(P, <)$, $(Q, \prec)$ be countable dense linearly ordered sets without endpoints[17]. Then $(P, <)$ and $(Q, \prec)$ are isomorphic.

   This implies that $\langle \mathbb{Q}, \leqslant_{\mathbb{Q}} \rangle$ is isomorphic to $\langle I_{\mathbb{Q}}, \leqslant_{I_{\mathbb{Q}}} \rangle$, where

$$I_{\mathbb{Q}} = \{x \in \mathbb{Q} : 0 < x < 1\} \ ,$$

   and $\leqslant_{I_{\mathbb{Q}}}$ is the restriction of $\leqslant_{\mathbb{Q}}$ to $I_{\mathbb{Q}}$.

When working with Formal Concept Analysis in later chapters of this text, special attention shall be given to the set inclusion order $\subseteq$. For the moment, we shall use a slightly different approach for working with members of $\mathcal{P}(Y)$.

A set-theoretic approach to number theory defines $2 := \{0, 1\}$[18]. Let us denote by 2^Y the set of all functions $Y \to \{0, 1\}$. Given a function $\chi \in 2^Y$, we can define a set $A_\chi \subseteq Y$ as

$$A_\chi = \{y \in Y : \chi(y) = 1\} \ .$$

Conversely, given $A \subseteq Y$, there is a unique $\chi_{_A} \in 2^Y$ such that

$$\chi_{_A}(y) = \begin{cases} 1, & \text{if } y \in A \\ 0, & \text{if } y \notin A \end{cases}$$

Furthermore, it can be proven that

$$\chi_{A_\chi} = A \ , \qquad A_{\chi_A} = A \ ,$$

[16] The proof can be found in (HRBÁČEK; JECH, 1999, Ch. 4).

[17] A set S is *countable* if there exists a bijection between it and $\mathbb{N}$; it is *dense* if it has at least two elements and for all $a, b \in S$, $a < b$ implies $a < c < b$ for some $c \in S$; and it has no endpoints if there are no $x, y \in S$ such that $x < a < y$ for all $a \in S$.

[18] In fact, writing $\overline{n}$ for the set-theoretic number n, we have

$$\overline{0} := \varnothing \ , \qquad \overline{n+1} := \{\overline{0}, \ldots, \overline{n}\} \ .$$

Thus, $\overline{n}$ has n members according to our intuition of natural numbers. For all purposes of this text we do not need to distinguish between the $\overline{n}$ and n, hence we shall always write n. For more on set-theoretic arithmetic, see (HRBÁČEK; JECH, 1999) .

so that we have a natural bijection between 2^Y and $\mathcal{P}(Y)$. Given $A \subseteq Y$, we call χ_A its *characteristic function.*

It is clear that, given $A, B \subseteq Y$, we have

$$A \subseteq B \text{ iff } \chi_A(y) \leqslant \chi_B(y) \text{ for all } y \in Y \ .$$

This can be used in order to extend the notion of subsetness as follows. In the remainder of this text, B^A denotes the set of all functions $A \to B$.

Proposition 0.2.16. *Let Y be a nonempty set and $\langle L, \leqslant_L \rangle$ a poset. Then the binary relation $\leqslant$ on L^Y defined for each $f, g \in L^Y$ as*

$$f \leqslant g \textit{ iff } f(y) \leqslant_L g(y) \textit{ for all } y \in Y$$

is an order.

If additionally $\langle L, \leqslant_L \rangle$ is a (complete) lattice, so is $\langle L^Y, \leqslant \rangle$.

Proof. We first prove that $\leqslant$ is an order on L^Y. Let $f, g, h \in L^Y$. Then

1. *(Reflexivity).* $f(y) \leqslant_L f(y)$ for each $y \in Y$, and so $f \leqslant f$.
2. *(Antisymmetry).* Suppose that $f \leqslant g$ and $g \leqslant f$. Then for each $y \in Y$, $f(y) \leqslant_L g(y)$ and $g(y) \leqslant_L f(y)$, so that $f(y) = g(y)$. Hence, $f = g$.
3. *(Transitivity).* Suppose that $f \leqslant g$ and $g \leqslant h$. Let $y \in Y$. Then $f(y) \leqslant_L g(y)$ and $g(y) \leqslant_L h(y)$, so that $f(y) \leqslant_L h(y)$. Hence, $f \leqslant h$.

Now, let us suppose that $\langle L, \leqslant_L \rangle$ is a (complete) lattice. Then, given a finite (arbitrary) set $\mathcal{F} \subseteq L^Y$, define $\overline{F}, \underline{F} \in L^Y$ respectively, for each $y \in Y$, by

$$\overline{F}(y) = \bigvee_{F \in \mathcal{F}} F(y) \ , \qquad \underline{F}(y) = \bigwedge_{F \in \mathcal{F}} F(y) \ .$$

Let f, g be respectively an upper bound and a lower bound of $\mathcal{F}$. Let $y \in Y$ be fixed. Then, for each $F \in \mathcal{F}$ we have

$$f(y) \geqslant F(y) \ , \qquad g(y) \leqslant F(y) \ ,$$

whence

$$f(y) \geqslant \bigvee_{F \in \mathcal{F}} F(y) = \overline{F}(y) \ , \qquad g(y) \leqslant \bigwedge_{F \in \mathcal{F}} F(y) = \underline{F}(y) \ ,$$

so that

$$\overline{F} = \bigvee \mathcal{F} \ , \qquad \underline{F} = \bigwedge \mathcal{F} \ .$$

□

In particular, with this order on 2^Y, $\langle \mathcal{P}(Y), \subseteq \rangle$ and $\langle 2^Y, \leqslant \rangle$ are isomorphic with isomorphism $f(A) = \chi_A$.

There is an interesting procedure for finding complete lattices ordered by set inclusion which uses the idea of closure operators defined below. Definition 0.2.17 and Theorem 0.2.20 below generalize results[19] by E. H. Moore, which are presented (in a form closer to contemporary mathematical practice) by BIRKHOFF (1948, p. 49).

Definition 0.2.17. Let Y be a nonempty set, $\langle P, \leqslant_P \rangle$ a poset and $\leqslant$ the order on P^Y we defined in Proposition 0.2.16. A *closure operator* on $\langle P^Y, \leqslant \rangle$ is a map $Cl : P^Y \to P^Y$ such that, for all $X, Z \in P^Y$:

1. $X \leqslant Cl(X)$ (extensivity)
2. $Cl(Cl(X)) = Cl(X)$ (idempotency)
3. If $X \leqslant Z$ then $Cl(X) \leqslant Cl(Z)$ (monotonicity)

If the closure operator on a context is clear we may denote $Cl(X)$ as $\overline{X}$. If $X = \overline{X}$ we say that X is *closed* (with respect to the closure operator).

Notice that, because of extensivity, X is closed iff $\overline{X} \leqslant X$.

Example 0.2.18. An example of closure operator is mapping a subset of a topological space S to its topological closure.

1. The closure of $]0,1[$ on $\mathbb{R}$ is $[0,1]$.
2. Consider $\mathbb{R}^2$ with its usual metric. For each $r > 0$ the closure of

   $$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < r\}$$

   is

   $$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant r\} \ .$$

Example 0.2.19. Let us consider a classical deductive system $\langle \mathcal{F}, \models \rangle$, in which $\mathcal{F}$ is a set of formulas[20] and, given two formulas $\phi, \psi \in \mathcal{F}$, let $\phi \models \psi$ mean that ψ is deductible from ϕ. If $\Gamma, \Delta \subseteq \mathcal{F}$, we write $\Gamma \models \Delta$ meaning that, assuming the formulas in Γ we derive each formula in Δ. Then $\langle \mathcal{F}, \models \rangle$ is a poset in which the deductive closure of a set of formulas Γ is the set of all the theorems deductible from it.

[19] By extending such results from $\langle \mathcal{P}(Y), \subseteq \rangle$, or equivalently from $\langle 2^Y, \leqslant_{2^Y} \rangle$, to $\langle L^Y, \leqslant_{L^Y} \rangle$.

[20] Think of a formula as a truth statement. For example, "$2 + 2 = 4$," or "$3 \leqslant 5$," or "under normal temperature and pressure conditions water is liquid."

As an example, the deductive closure of the set of logical axioms of the Statement Calculus[21] (SC) is the set of theorems of SC.

Theorem 0.2.20. *Let Y be a nonempty set, L a complete lattice and L^Y be ordered as in Proposition 0.2.16. The family $\mathscr{C}(L^Y)$ of closed elements of L^Y (according to any given closure operator) is a complete lattice. Given a collection $\langle A_j \rangle_{j \in J}$ of elements of $\mathscr{C}(L^Y)$, we have*

$$\otimes \langle A_j \rangle_{j \in J} = \bigwedge_{j \in J} A_j \quad , \tag{0.2.21}$$

$$\oplus \langle A_j \rangle_{j \in J} = \overline{\bigvee_{j \in J} A_j} \quad , \tag{0.2.22}$$

where $\bigwedge, \bigvee$ are the meet and join of L^Y, and $\otimes, \oplus$ are the meet and join of $\mathscr{C}(L^Y)$.

Proof. Let $\mathcal{C} = \langle A_j \rangle_{j \in J}$ be any nonempty collection of closed elements of L^Y, and let

$$B = \bigwedge_{j \in J} A_j \quad , \qquad\qquad C = \bigvee_{j \in J} A_j \quad .$$

Then monotonicity yields[22]

$$\overline{B} \leqslant \overline{A_j} = A_j$$

for each $j \in J$, and so B is a lower bound of $\mathcal{C}$. Thus,

$$\overline{B} \leqslant B \quad ,$$

i.e., $B \in \mathscr{C}(L^Y)$. Now, since

$$B = \left(\bigwedge_{j \in J} A_j \right) \in \mathscr{C}(L^Y) \quad ,$$

we conclude that

$$B = \otimes \langle A_j \rangle_{j \in J} \quad .$$

Conversely, for each $j \in J$ we have

$$A_j \leqslant \bigvee_{j \in J} A_j = C.$$

Now, if $D \in \mathscr{C}(L^Y)$ is an upper bound of $\mathcal{C}$ on $\mathscr{C}(L^Y)$, then by monotonicity of the closure we have

$$\overline{C} \leqslant \overline{D} = D \quad ,$$

---

[21] The Statement Calculus is a formal deductive system which considers formulas as constants and uses the logical connectives $\neg, \longrightarrow$ (logical negation and implication, respectively), from which connectives of conjunction ("and") and disjunction ("or") are defined. We shall explore a little of the Statement Calculus on Sec. 2.3. The interested reader is referred to (MENDELSON, 2009), (MARGARIS, 1990).

[22] Recall that X_j is closed.

hence

$$\overline{C} = \bigoplus \langle A_j \rangle_{j \in J} \ .$$

□

Example 0.2.23. In Example 0.2.19 we stated that the topological closure on a topological space is an instance of a closure operator, as well as the deductive closure on a (classical) deductive system. Therefore, it follows from Theorem 0.2.20 that both the set of closed sets on a topological space and the set of deductive closures on a (classical) deductive system constitute complete lattices.

Just before following to the next chapter, we state the following definition, in accordance with (FALMAGNE; DOIGNON, 2011) .

Definition 0.2.24. Let $\langle P, \leqslant \rangle$, $\langle Q, \preceq \rangle$ be ordered sets. A *Galois connection* between $\langle P, \leqslant \rangle$ and $\langle Q, \preceq \rangle$ is a pair (f, g) of functions $f : P \to Q$, $g : Q \to P$ such that, for all $p, p_1, p_2 \in P$ and for all $q, q_1, q_2 \in Q$,

1. $p_1 \leqslant p_2$ implies $f(p_2) \preceq f(p_1)$	1'. $q_1 \preceq q_2$ implies $g(q_2) \leqslant g(q_1)$
2. $p \leqslant (g \circ f)(p)$	2'. $q \preceq (f \circ g)(q)$

Clearly (f, g) is a Galois connection between $\langle P, \leqslant \rangle$ and $\langle Q, \preceq \rangle$ iff (g, f) is a Galois connection between $\langle Q, \preceq \rangle$ and $\langle P, \leqslant \rangle$. It can be proved[23] that $f \circ g$, $g \circ f$ are closure operators on Q and P respectively. In fact, in the remainder of this text the most important[24] closure operators we shall work with are Galois connections.

Now the reader has all the tools necessary to understand the ideas concerning orders and lattices on this text. We are ready to introduce the classical theory of Formal Concept Analysis.

[23] See (FALMAGNE; DOIGNON, 2011) .

[24] But we shall not work exclusively with them. In fact, in Sec. 4.1 a computational algorithm is presented which applies to general closure operators on finite sets.

1 Formal Concept Analysis

According to (JOSEPH, 2002), a *concept* is an abstraction produced by the intellect (e.g. the concept of *tree* abstracts what is common to several trees one has contact with). A *symbol* is a sign (e.g., a sound, a sequence of letters, an image) with a meaning imposed to it by convention (e.g., an image of a skull which stands for *danger*, or the sequence of characters T-R-E-E which stands for the notion of *tree*). A *term* is a concept communicated by a symbol (e.g., “tree" is a term which conveys the idea of the concept *tree*). Thus, she distinguishes *concepts* and *terms* by stating that a concept is an idea (in the mind) which represents a reality, whereas a term is that idea in transit, being communicated.

Also according to JOSEPH, a term has both *extension* and *intension*, the former corresponding to the set of all objects to which the term corresponds (e.g., the set of all trees), and the later, to “the sum of the essential characteristics that the term implies." Moreover, “As a term increases in intension, it decreases in extension."

Formal Concept Analysis (FCA) is to be regarded as a mathematization of the philosophical idea of *concept* (GANTER; WILLE, 1999) , based on the slightly different idea that it is the *concept* that has extension and intension.

1.1 Definitions and properties

The definitions and theorems presented in this section follow those presented by (GANTER; WILLE, 1999) with a different notation and slightly different proofs.

Definition 1.1.1. A *formal context* is an ordered triple $\mathbb{C} := \langle \mathcal{O}, \mathcal{A}, I \rangle$, in which $\mathcal{O}$ and $\mathcal{A}$ are non-empty sets, and $I \subseteq \mathcal{O} \times \mathcal{A}$ is a binary relation.

The elements of $\mathcal{O}$ are called *objects*, and the elements of $\mathcal{A}$ *attributes*. We say that, in the context $\mathbb{C}$, an object o has an attribute a if and only if $(o, a) \in I$. We shall write oIa for $(o, a) \in I$.

Example 1.1.2. One may represent a finite formal context in an easy manner by using a table in which rows are indexed by objects and columns by attributes. A cell in row o and column a is marked if and only if oIa. As an example, Table 1 represents a formal context of animals.

Notice that in this example there is no distinction between swans and geese. We should add more attributes if we wanted to distinguish between them. On the other hand, because in this context the attribute for having wings gives us no new information

$\mathcal{O}$	$\mathcal{A}$						
	Vertebrate	Lay eggs	Carnivorous	Has wings	Flies	Quadruped	Crawls
Eagle	X	X	X	X	X		
Snake	X	X	X				X
Goose	X	X		X	X		
Swan	X	X		X	X		
Lion	X		X			X	

Table 1 – A formal context of animals

(every winged animal in this context flies) we could rule this attribute out. However it would be an important attribute if we had for instance the object "Chicken."

This example illustrates the fact that formal contexts are often narrower than the real world. Nonetheless, it is a way of representing information that can be made very powerful as we shall see further ahead.

Definition 1.1.3. Let $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ be a formal context. We define two maps,

$$^* : 2^{\mathcal{O}} \to 2^{\mathcal{A}} \ , \qquad\qquad ^\wedge : 2^{\mathcal{A}} \to 2^{\mathcal{O}} \ ,$$

given for each $O \subseteq \mathcal{O}$ and $A \subseteq \mathcal{A}$ by the relations

$$O^* := \{a \in \mathcal{A} : oIa \text{ for all } o \in O\} \tag{1.1.4}$$

$$A^\wedge := \{o \in \mathcal{O} : oIa \text{ for all } a \in A\} \ , \tag{1.1.5}$$

where write O^* for $^*(O)$ and $A^\wedge$ for $^\wedge(A)$.

According to this definition, O^* is the set of all attributes common to every object in O, and $A^\wedge$ is the set of all objects having every attribute in A.

We can now define the central object of FCA:

Definition 1.1.6. Let $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ be a formal context. A *formal concept* (often referred to as *concept*) of $\mathbb{C}$ is an ordered pair $C = \langle O, A \rangle$ such that $O \subseteq \mathcal{O}$, $A \subseteq \mathcal{A}$, $O^* = A$ and $A^\wedge = O$. The sets O and A are called the *extent* and *intent* of the concept C respectively.

In other words, $C = \langle O, A \rangle$ is a concept if the following conditions hold:

1. A is precisely the set of all attributes common to every object of O; and
2. no object of $\mathcal{O} \backslash O$ has every attribute of A.

Example 1.1.7. The following are formal concepts of the context presented in Table 1:

$$C_{\text{Lion}} = \langle \{\text{Lion}\}, \{\text{Quadruped, Carnivorous, Vertebrate}\} \rangle \ , \tag{1.1.8}$$

$$C_{\text{Carnivorous}} = \langle \{\text{Eagle, Snake, Lion}\}, \{\text{Carnivorous, Vertebrate}\} \rangle \ . \tag{1.1.9}$$

Notice that there is an inverse relation between the numbers of elements in the intent and the extent of a concept. If we increase the number of elements in the extent (we added "Eagle" and "Snake" to it), the number of elements in the intent is reduced (eagles are not quadrupeds and neither are snakes)[1]. In fact, the following useful properties hold:

Theorem 1.1.10. *Let* $O, O_1, O_2 \subseteq \mathcal{O}$ *and* $A, A_1, A_2 \subseteq \mathcal{A}$. *Then*

1. If $O_1 \subseteq O_2$ *then* $O_2^* \subseteq O_1^*$	*1'. If* $A_1 \subseteq A_2$ *then* $A_2^\wedge \subseteq A_1^\wedge$
2. $O \subseteq O^{*\wedge}$	*2'.* $A \subseteq A^{\wedge *}$
3. $O^* = O^{*\wedge *}$	*3'.* $A^\wedge = A^{\wedge * \wedge}$

4. $O \subseteq A^\wedge$ *iff* $A \subseteq O^*$ *iff* $O \times A \subseteq I$

Proof. We shall prove items 1., 2., 3. and 4. Items with a prime can be proved analogously.

1. Let $a \in O_2^*$. Then oIa for all $o \in O_2$. In particular, oIa for all $o \in O_1$. Thus, $a \in O_1^*$.
2. Let $o \in O$. Then oIa for all $a \in O^*$, by definition of O^*. Thus, by definition of $O^{*\wedge}$, we have $o \in O^{*\wedge}$.
3. From item 2'. with $A = O^*$ we already know that $O^* \subseteq O^{*\wedge *}$. Let $a \in O^{*\wedge *}$. Then

   (**i**) oIa for all $o \in O^{*\wedge}$,

   by definition of $O^{*\wedge *}$. Now let $\tilde{o} \in O^{*\wedge}$ be fixed. For all $\tilde{a} \in \mathcal{A}$,

   (**ii**) if $\tilde{o}I\tilde{a}$ then $\tilde{a} \in O^*$,

   by definition of $O^{*\wedge}$. From (**i**) we have $\tilde{o}Ia$. Thus, using (**ii**) we conclude that $a \in O^*$.
4. Suppose $O \subseteq A^\wedge$. By item 1., $A^{\wedge *} \subseteq O^*$. Using item 2'. and transitivity of $\subseteq$, we have $A \subseteq O^*$.
   Now suppose $A \subseteq O^*$. By items 1'. and 2., we have $O \subseteq O^{*\wedge} \subseteq A^\wedge$.
   Assuming $A \subseteq O^*$, let $o \in O$ and $a \in A$. By definition of O^*, $oI\tilde{a}$ for all $\tilde{a} \in O^*$. By hypothesis, $a \in A \subseteq O^*$. Thus, oIa. Since $o \in O$ and $a \in A$ are arbitrary, $O \times A \subseteq I$. Hence, $A \subseteq O^* \Rightarrow O \times A \subseteq I$.
   Finally, suppose $O \times A \subseteq I$. Let $o \in O$. By hypothesis, for all $a \in A$ we have oIa. By definition of O^*, if oIa then $a \in O^*$. Thus if $a \in A$ then $a \in O^*$. This completes the proof.

□

Items 1., 1'., 2., 2'. of Theorem 1.1.10 yield the following.

[1] Notice how this, together with items 2., 2'. of Theorem 1.1.10, resembles JOSEPH's statement that "As a term increases in intension, it decreases in extension" (2002).

Corollary 1.1.11. *Let $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ be a formal context. Then $(^*, ^\wedge)$ is a Galois connection*[2] *between $\langle \mathcal{O}, \subseteq \rangle$ and $\langle \mathcal{A}, \subseteq \rangle$, hence $^{*\wedge}$ and $^{\wedge *}$ are both closure operators*[3].

From properties 3. and 3'. of Theorem 1.1.10 we see that, given $O \subseteq \mathcal{O}$ and $A \subseteq \mathcal{A}$, $\langle O^{*\wedge}, O^* \rangle$ and $\langle A^\wedge, A^{\wedge *} \rangle$ are concepts. On the other hand if $C = \langle O, A \rangle$ is a concept then by definition

$$C = \langle O, O^* \rangle = \langle A^\wedge, A \rangle \ .$$

This proves the following.

Lemma 1.1.12. *Let $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ be a formal context. Then C is a formal concept of $\mathbb{C}$ iff there exist $O \in \mathcal{O}$ and $A \in \mathcal{A}$ such that*

$$C = \langle O, O^* \rangle = \langle A^\wedge, A \rangle = \langle O, A \rangle \ .$$

In particular, for all $O \subseteq \mathcal{O}$, $A \subseteq \mathcal{A}$, the following are formal concepts:

$$\langle O^{*\wedge}, O^* \rangle \ , \qquad\qquad \langle A^\wedge, A^{\wedge *} \rangle \ .$$

Lemma 1.1.12 gives us a procedure for finding concepts[4]. For example, if we want to find the concept of "Carnivorous" we presented in Example 1.1.7 — that is, the concept with the smallest intent such that "Carnivorous" is an attribute —, start with the set {Carnivorous} and then apply respectively the maps $^\wedge$ and * to it:

$$\begin{aligned} \{\text{Carnivorous}\}^{\wedge *} &= \{\text{Eagle}, \text{Snake}, \text{Lion}\}^* \\ &= \{\text{Vertebrate}, \text{Carnivorous}\} \end{aligned} \tag{1.1.13}$$

From (1.1.13), using Theorem 1.1.10 we have the concept

$$C_{\text{Carnivorous}} = \langle \{\text{Eagle}, \text{Snake}, \text{Lion}\}, \{\text{Vertebrate}, \text{Carnivorous}\} \rangle \tag{1.1.14}$$

as presented in (1.1.9).

Now in Table 1, consider the concept of *eagle*:

$$C_{\text{Eagle}} = \langle \{\text{Eagle}\}, \{\text{Flies, Has Wings, Lay eggs, Carnivorous, Vertebrate}\} \rangle \ . \tag{1.1.15}$$

Comparing (1.1.15) and (1.1.14), we see that as we intersect the intents of the concepts of *lion* and *eagle* we get the intent of another concept: that of *carnivorous*. In fact, an arbitrary intersection of intents is an intent which is, of course, smaller than the original one. This suggests the existence of an order between intents — and, in fact, between attributes.

---

[2] See Def. 0.2.24.
[3] See Def. 0.2.17.
[4] Computational algorithms for this process shall be explored in Sec. 1.2.

Definition 1.1.16. Let $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ be a formal context, and denote by $\mathfrak{B}(\mathcal{C})$ the set of all formal concepts on $\mathbb{C}$. Define an order $\leqslant_{\mathbb{C}}$ on $\mathfrak{B}(\mathbb{C})$ for each $\langle O_1, A_1 \rangle$, $\langle O_2, A_2 \rangle$ by

$$\begin{aligned} \langle O_1, A_1 \rangle \leqslant_{\mathbb{C}} \langle O_2, A_2 \rangle \text{ iff } & O_1 \subseteq O_2 \ , \\ \text{iff } & A_2 \subseteq A_1 \ . \end{aligned} \tag{1.1.17}$$

Given two concepts $C_1, C_2 \in \mathfrak{B}(\mathbb{C})$, we say that C_1 is a *subconcept* of C_2 (or that C_2 is a *superconcept* of C_1) iff $C_1 \leqslant_{\mathbb{C}} C_2$.

Interestingly, $\langle \mathfrak{B}(\mathbb{C}), \leqslant_{\mathbb{C}} \rangle$ is not simply a poset. It can be shown that, with the order introduced in Def. 1.1.16, the set of all formal concepts of any context constitutes a complete lattice.

Theorem 1.1.18 (The Basic Theorem on Concept Lattices)**.** *Let $\mathbb{C}$ be a formal context. Then $\mathcal{L}_{\mathbb{C}} := \langle \mathfrak{B}(\mathbb{C}), \leqslant_{\mathbb{C}} \rangle$ is a complete lattice, called the* concept lattice *of $\mathbb{C}$.*

If K is an index set and $C_\kappa = \langle O_\kappa, A_\kappa \rangle \in \mathfrak{B}(\mathbb{C})$ for each $\kappa \in K$ then

$$\inf_{\kappa \in K} C_\kappa = \left\langle \bigcap_{\kappa \in K} O_\kappa, \left(\bigcup_{\kappa \in K} A_\kappa \right)^{\wedge *} \right\rangle \ , \tag{1.1.19}$$

$$\sup_{\kappa \in K} C_\kappa = \left\langle \left(\bigcup_{\kappa \in K} O_\kappa \right)^{* \wedge}, \bigcap_{\kappa \in K} A_\kappa \right\rangle \ . \tag{1.1.20}$$

Furthermore, a complete lattice $\langle L, \leqslant_L \rangle$ is isomorphic to $\mathfrak{B}(\mathbb{C})$ iff there are mappings $\omega : \mathcal{O} \to L$ and $\alpha : \mathcal{A} \to L$ such that $\omega(\mathcal{O})$ is supremum-dense in L, $\alpha(\mathcal{A})$ is infimum-dense in L and, for all $o \in \mathcal{O}, a \in \mathcal{A}$,

$$oIa \text{ iff } \omega(o) \leqslant \alpha(a) \ .$$

Proof. We prove that $\langle \mathfrak{B}(\mathbb{C}), \leqslant_{\mathbb{C}} \rangle$ is a complete lattice, and that (1.1.19) and (1.1.20) hold. The remainder of the proof — i.e., the part concerning isomorphisms and the maps ω, α — is left undone in this text, and the proof can be found in (GANTER; WILLE, 1999) .

Let us consider the concepts C_κ indexed by K. As we have already noticed in Cor. 1.1.11, the maps ${}^{*\wedge}$ and ${}^{\wedge *}$ are closure operators on $\mathcal{O}, \mathcal{A}$ respectively, whose closed elements are respectively extents and intents.

Thus, Theorem 0.2.20 implies[5] that the collections $\mathscr{C}(\mathcal{O}), \mathscr{C}(\mathcal{A})$ of closed subsets of $\mathcal{O}, \mathcal{A}$ with respect to ${}^{*\wedge}$, ${}^{\wedge *}$ are complete lattices[6] and, in particular,

$$\bigwedge_{\kappa \in K} O_\kappa = \bigcap_{\kappa \in K} O_\kappa \ , \qquad \bigvee_{\kappa \in K} O_\kappa = \left(\bigcup_{\kappa \in K} O_\kappa \right)^{* \wedge} \ , \qquad (\text{on } \langle \mathscr{C}(\mathcal{O}), \subseteq \rangle)$$

$$\bigwedge_{\kappa \in K} A_\kappa = \bigcap_{\kappa \in K} A_\kappa \ , \qquad \bigvee_{\kappa \in K} A_\kappa = \left(\bigcup_{\kappa \in K} A_\kappa \right)^{\wedge *} \ . \qquad (\text{on } \langle \mathscr{C}(\mathcal{A}), \subseteq \rangle)$$

[5] Here $L = \{0, 1\}$, hence by isomorphism Theorem 0.2.20 applies to $Y = \mathcal{O}, \mathcal{A}$ ordered by set inclusion.
[6] Ordered by $\subseteq$, according to Theorem 1.1.10, with infimum $\bigcap$ and supremum $\bigcup$.

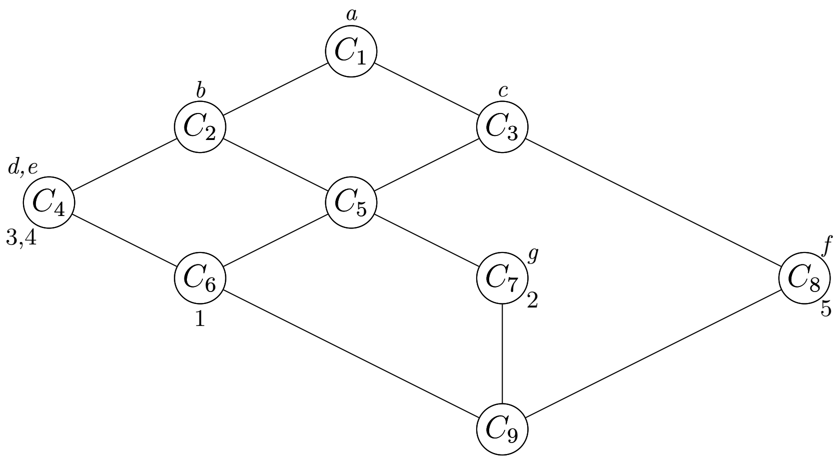


Figure 2 – Concept Lattice of Animals

The equations in the first row imply[7] that $\mathfrak{B}(\mathbb{C})$ is a complete lattice, as well as the formulae for extents of $\inf C_\kappa$ and $\sup C_\kappa$.

Now, the restriction of * to $\mathscr{C}(\mathcal{O})$ establishes an isomorphism between $\langle \mathscr{C}(\mathcal{O}), \subseteq \rangle$ and $\langle \mathscr{C}(\mathcal{A}), \supseteq \rangle$[8], so that the second row of equations above impose respectively formulae for supremum and infimum of intents on $\langle \mathfrak{B}(\mathbb{C}), \geqslant \rangle$, so that by duality we get the intents of (1.1.19) and (1.1.20).

□

Theorem 1.1.18 allows us to use lattice theory for finding out many properties that come from a formal context. In particular, a finite concept lattice has an easy visual representation (see Example 1.1.21 below). In order to interpret the concept lattice from the diagram, one may write, for each concept on the diagram, the elements of its intent and extent. However, from the order $\leqslant$ of the concept lattice a tidier manner of presenting the diagram can be devised: for a given concept, instead of writing every element of its extent (resp. intent), we write only those objects (resp. attributes) that did not appear below (resp. above) in the concept lattice. This is possible because of the Basic Theorem.

Example 1.1.21. The concept lattice of the context presented in Table 1 is shown in Fig. 2. Each concept is represented by a circle. Here animals are represented by numbers 1-5 in the order they appear in Table 1. Attributes are represented by letters a-g, also in the order they appear in the table. The extent (intent) of a given concept C has an object (attribute) iff that object (attribute) appears near a concept $\tilde{C}$ such that there is a descending (ascending) path from C to $\tilde{C}$.

[7] By definition of the order on $\mathfrak{B}(\mathbb{C})$.

[8] By Lemma 1.1.12, elements of $\mathscr{C}(\mathcal{O})$ have the form $A^\wedge$, and elements of $\mathscr{C}(\mathcal{A})$ have the form O^*. To see that $*$ is one-to-one, let A_1, A_2 be such that $A_1^{*\wedge} = A_2^{*\wedge}$. Then $A_1^{*\wedge *} = A_2^{*\wedge *}$, whence $A_1^\wedge = A_2^\wedge$ by Theorem 1.1.10 item 3'. Also, * is onto, because given $A \in \mathscr{C}(\mathcal{A})$, $A = O^*$ for some $O \in \mathcal{O}$, whence $A = (O^{*\wedge})^*$ by Theorem 1.1.10 item 3. Finally, * is order-preserving by item 1. of the same theorem.

Take, for instance, the concept C_3. It has ascending paths to concepts C_1 and C_3, and descending paths to concepts C_8, C_5, C_6, C_7 and C_9. On the other hand, C_3 has no (strictly ascending or descending) paths to C_2 or C_4. Thus, $C_3 = \langle\{1,2,5\},\{a,c\}\rangle$.

Concept lattices are visual tools that allow us to find out relations on attributes and objects (for example, any animal with property e also has property b). However, limitations may arise on the theory, as, for example, the definition of formal context allows us only to work only with precise relations. For example, a chicken can fly for short distances, but this information could not be expressed on a formal context as defined earlier. In the next section we generalize these ideas to allow fuzzy objects, attributes and relations.

1.2 Computing the concept lattice

For purposes of computing the concept lattice of a given formal context $\mathbb{C} = \langle\mathcal{O},\mathcal{A},I\rangle$, let us consider that $\mathcal{O},\mathcal{A}$ are finite. Recall that according to Lemma 1.1.12, given any $A \subseteq \mathcal{A}$ the set $A^{\wedge *}$ is the intent of a concept[9], and from an intent the extent can be uniquely determined[10]. Thus, we shall consider that, in order to compute the formal concepts of $\mathbb{C}$, it is sufficient to evaluate closures of sets of attributes. Thus, we shall write $\overline{A}$ for $A^{\wedge *}$.

The idea of Lemma 1.1.12 that every intent is the closure of a set of attributes provides us with an algorithm for computing all the intents of $\mathbb{C}$ and thus, all of its concepts.

Algorithm 1 – Standard Computation of Intents

input : Formal context $\mathbb{C}$.
output : The set Intents of all intents of $\mathbb{C}$.

```
Intents ← ∅;
foreach A ∈ P(A) do
    if Ā ∉ Intents then
        append Ā to Intents;
    end
end
```

[9] Analogously, given $O \subseteq \mathcal{O}$, the set $O^{*\wedge}$ is the extent of a concept. In fact, all the ideas presented in this section for subsets of $\mathcal{A}$ and intents are true (with analogous proofs) if we consider subsets of $\mathcal{O}$ and extents instead, and then interchange the operators $\wedge$ and $*$.

[10] In reality we have more than that: in the process of computing $A^{\wedge *}$ we do compute $A^{\wedge}$ which can be stored in the process. Nonetheless in the algorithms presented in this section, in order to find all the concepts we only need $A^{\wedge}$ to compute $A^{\wedge *}$ (and, of course, we store $A^{\wedge}$), and so we shall consider that only $A^{\wedge *}$ needs to be computed.

For running the loop started in row 2 we assume a linear order on $\mathcal{P}(\mathcal{A})$. Algorithm 1 has a critical computational issue: it has exponential computational complexity. More precisely, if $\mathcal{A}$ has n elements this algorithm has to consider 2^n subsets A of $\mathcal{A}$ and evaluate $\overline{A}$ for each of them[11].

We are provided in (GANTER; WILLE, 1999) with an algorithm which is not necessarily of exponential complexity. This algorithm is based on the idea of finding a lectic order on $\mathcal{P}(A)$ which is a linear strict order[12] on this set, and then restricting this order to a linear strict order on the set $\mathfrak{B}(\mathbb{C})$ of formal concepts of $\mathbb{C}$. In what follows, we assume that $\mathcal{A} = \{1, 2, ..., n\}$ is linearly ordered by the usual order of natural numbers[13].

Definition 1.2.1. Let $A, B \subseteq \mathcal{A}$. Then A is said *lectically smaller* than B, and this is denoted as $A < B$, iff

$$\text{There exists } i \in B \backslash A \text{ such that } A \cap \{1, ..., i-1\} = B \cap \{1, ..., i-1\} \ ,$$

where

$$\{1, ..., 0\} := \varnothing \ .$$

It can be shown that the lectic order is a linear strict order on $\mathcal{P}(\mathcal{A})$.

Example 1.2.2. The lectic order is not related with the number of members of a set, but with the smallest element distinguishing them.

1. If $A \subset B$, then $A < B$. In fact, let i be the smallest element of $B \backslash A$. Then

$$A \cap \{1, ..., i-1\} = B \cap \{1, ..., i-1\} \ .$$

   In particular, if $\varnothing \subset A \subset \mathcal{A}$ then $\varnothing < A < \mathcal{A}$.

2. If $A = \{1, 2\}$ and $B = \{3\}$, then $B < A$. In fact, the smallest element $i = 1$ of $\mathcal{A}$ is an element of $A \backslash B$, and

$$B \cap \{1, ..., i-1\} = B \cap \varnothing = \varnothing = A \cap \varnothing = A \cap \{1, ..., i-1\} \ .$$

Due to monotonicity of the closure operator, from item 1. of Example 1.2.2 we see that $\overline{\varnothing}$ is the lectically smallest intent of $\mathbb{C}$. If a procedure can be found so that, given an intent, the next intent according to the lectic order can be found, we shall have all that is necessary for finding all the intents of $\mathbb{C}$ up to $\mathcal{A}$.

11 We can save several computations here if instead of evaluating $A^{\wedge *}$ in row 3 we verified in that row whether A^* has already been computed and only then (i. e., only if we have a newly found extent) we compute $A^{\wedge *}$ and append it to the set of intents. Nonetheless, the number of evaluations of A^* would remain 2^n.

12 See Prop. 0.1.5.

13 If $\mathcal{A} = \{a_1, a_2, ...a_n\}$, we only need to order $\mathcal{A}$ according to the indices of its elements.

Define, for each $A \subseteq \mathcal{A}$ and each $i \in \mathcal{A}$ the following:

$$A \oplus i := \overline{(A \cap \{1, ..., i-1\}) \cup \{i\}} \ .$$

Because $A \oplus i$ is the closure of a set of objects it is an intent. Furthermore, if $i \notin A$, then

$$A \cap \{1, ..., i-1\} \subset (A \cap \{1, ..., i-1\}) \cup \{i\} \ ,$$

and so

$$A < (A \cap \{1, ..., i-1\}) \cup \{i\} \leqslant A \oplus i \ ,$$

implying that $A < A \oplus i$.

Is it possible to determine an $i \in \mathcal{A}$ so that $A \oplus i$ is *the smallest* intent greater than A? Ganter and Wille define in (GANTER; WILLE, 1999) a binary relation $<_i$ by

$$A <_i B \text{ iff } i \in B \backslash A \text{ and } A \cap \{1, ..., i-1\} = B \cap \{1, ..., i-1\} \ .$$

Notice that $A < B$ iff there exists a unique $i \in \mathcal{A}$ such that $A <_i B$. Then, the authors present us with the following.

Theorem 1.2.3. *The smallest intent greater than $A \subset \mathcal{A}$ with respect to the lectic order is*

$$A \oplus i \ ,$$

where i is the greatest element of $\mathcal{A}$ such that $A <_i A \oplus i$.

The result is the following algorithm.

Algorithm 2 – Next Intent

input : Formal context $\mathbb{C}$.
output : The set Intents of all intents of $\mathbb{C}$.

```
1  B = ∅̄;
2  append B to Intents;
3  while B ≠ 𝒜 do
4      A = B;
5      for i = Max(𝒜\A):1: −1 do
6          B = A ⊕ i;
7          if A <_i B then Break;
8      end
9      append B to Intents;
10 end
```

The notation in row 5 means that the variable i assumes values from $\max(\mathcal{A} \backslash A)$ down to 1 decreasing 1 unit each loop. Algorithm 2 tends to be much faster than the

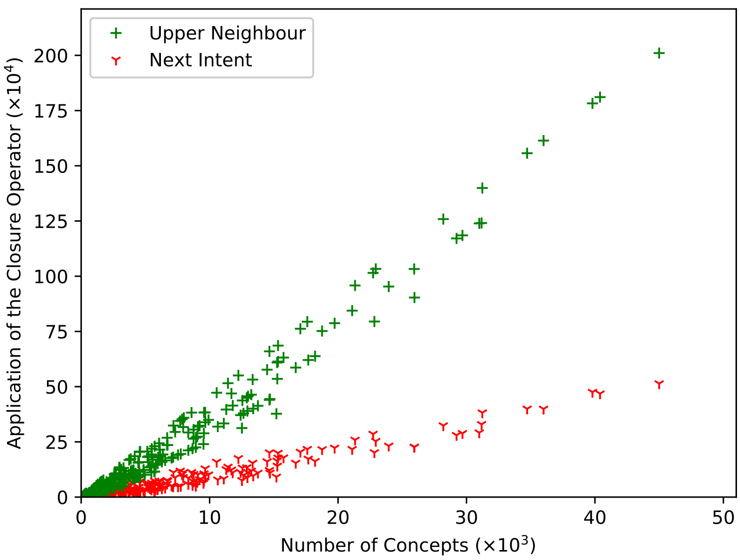


Figure 3 – Complexity of the Algorithms

standard algorithm (Alg. 1) because the number of applications of the closure operator is in general not much larger than the number of concepts. For some computational experiments, see Fig. 3.

Example 1.2.4. The complexity of the algorithm depends on the order of the rows of the formal context. In fact, consider the following formal context:

$\mathcal{O}$	$\mathcal{A}$		
	1	2	3
a		X	X
b		X	
c	X		

By applying Alg. 2, we have the following sequence of events.

Step	Evaluation					Comparison			Action
1			$\overline{\varnothing}$	=	$\varnothing$				Store $\varnothing$
2	$\varnothing$	$\oplus$	3	=	$\{2,3\}$	$\varnothing$	$\not<_3$	$\{2,3\}$	Try $\oplus 2$
3	$\varnothing$	$\oplus$	2	=	$\{2\}$	$\varnothing$	$<_2$	$\{2\}$	Store $\{2\}$
4	$\{2\}$	$\oplus$	3	=	$\{2,3\}$	$\{2\}$	$<_3$	$\{2,3\}$	Store $\{2,3\}$
5	$\{2,3\}$	$\oplus$	1	=	$\{1\}$	$\{2,3\}$	$<_1$	$\{1\}$	Store $\{1\}$
6	$\{1\}$	$\oplus$	3	=	$\{1,2,3\}$	$\{1\}$	$\not<_3$	$\{1,2,3\}$	Try $\oplus 2$
7	$\{1\}$	$\oplus$	2	=	$\{1,2,3\}$	$\{1\}$	$<_2$	$\{1,2,3\}$	Store $\{1,2,3\}$

If we interchange columns 1 and 3 we get the following context.

$\mathcal{O}$	$\mathcal{A}$		
	$1'$	$2'$	$3'$
a	X	X	
b		X	
c			X

This gives us the following sequence of steps.

Step	Evaluation					Comparison			Action
1			$\overline{\varnothing}$	$=$	$\varnothing$				Store $\varnothing$
2	$\varnothing$	$\oplus$	$3'$	$=$	$\{3'\}$	$\varnothing$	$<_{3'}$	$\{3'\}$	Store $\{3'\}$
3	$\{3'\}$	$\oplus$	$2'$	$=$	$\{2'\}$	$\{3'\}$	$<_{2'}$	$\{2'\}$	Store $\{2'\}$
4	$\{2'\}$	$\oplus$	$3'$	$=$	$\{1',2',3'\}$	$\{2'\}$	$\not<_{3'}$	$\{1',2',3'\}$	Try $\oplus 1'$
5	$\{2'\}$	$\oplus$	$1'$	$=$	$\{1',2'\}$	$\{2'\}$	$<_{1'}$	$\{1',2'\}$	Store $\{1',2'\}$
6	$\{1',2'\}$	$\oplus$	$3'$	$=$	$\{1',2',3'\}$	$\{1',2'\}$	$<_{3'}$	$\{1',2',3'\}$	Store $\{1',2',3'\}$

Now if we recall that

$$1' = 3 \ , \qquad 2' = 2 \ , \qquad 3' = 1 \ ,$$

we see that we get the same results, but with in different sequences and in different numbers of steps. Nonetheless, both executions apply the closure operator a smaller number of times than the standard algorithm, for which necessarily $2^3 = 8$ applications are made.

Algorithm 2 presents itself as a (comparably) fast algorithm for finding every formal concept of a given formal concept. However, it lacks a procedure for evaluating the structure of the concept lattice. After computing the concepts, another algorithm must be used for computing their lattice structure.

LINDIG proposed (2000) another algorithm for computing the concepts together with its lattice structure. LINDIG's Upper Neighbour algorithm shall not be presented here. Rather, we shall present an extended version of it later in this text[14]. We do nonetheless present the basic idea of the algorithm.

Recall that according to Theorem 0.2.20, given a closure operator, the collection of closed sets with respect to it constitutes a complete lattice. In particular, the operator $^{\wedge *}$ is a closure operator, and so the set of intents constitutes a complete lattice on $\mathcal{A}$, which is isomorphic to $\langle \mathfrak{B}(\mathbb{C}), \geqslant \rangle$[15].

LINDIG's Upper Neighbour[16] algorithm is based on the idea of first evaluating the smallest element of this lattice, which is the closure of the emptyset, and then recursively,

[14] See Ch. 4.
[15] That is the dual of the concept lattice. See (1.1.17).
[16] See Def. 0.2.5

for each closed set A, to compute the set of upper neighbours of A (and dually attributing A as a lower neighbour of each if its upper neighbours).

Figure 3 presents the number of applications of the closure operator of Alg. 2 (“Next Intent") and LINDIG's “Upper Neighbour" algorithm in terms of the number of concepts of the context. Our computational experiments were made with randomly generated $m \times n$ formal contexts, with $m = 5, 10, ..., 40$ and $n = 5, 10, ..., 50$.

Our goal in the remainder of this text is to present the reader with an extension of the theory of Formal Concept Analysis which allows us to work with graded values of relationships between objects and attributes, i. e., we intend to allow other possibilities than “object o has attribute a," by replacing this statement with (possibly) more flexible statements of the form “object o has attribute a to a certain degree d". This goal may be achieved by replacing the classical notion of *set* with a more flexible, fuzzy notion.

2 Fuzzy Sets and Fuzzy Connectives

Fuzzy sets extend classical set theory in order to deal with uncertainty. Whereas classical mathematics works on the assumption that we have absolute precision concerning data (we know whether or not an element belongs to a given set, or we know exactly how a function behaves within some neighbourhood of a point), fuzzy logic deals with imprecisions.

There are of course social conventions which behave classically, such as the age from which someone is allowed to drive, or to drink alcoholic beverages. Nonetheless, considerations on how such legislations vary across different countries indicates that these conventions do not correspond precisely to reality, i.e., in reality people are sufficiently mature to drink or drive at different ages, but because maturity is a subjective matter and the legislation should (in theory at least) treat individuals objectively, it establishes a number with some degree of arbitrariness[1].

In the present chapter we present the required theory of fuzzy sets and fuzzy connectives as it shall be necessary in the following chapters.

2.1 Fuzzy sets

In this section we define what we mean by a fuzzy set. In order to avoid misleading the reader with seemingly nonsensical definitions, we start by considering the classical theory of sets. Then, by extending the classical definitions, we shall define fuzzy sets. We hope that our approach will help the reader to build an intuition on fuzzy sets and, at the same time, to understand what the definitions mean. Nonetheless, we shall not present here a detailed discussion on axiomatic set theory. We refer the reader interested in the axiomatic approach to (HRBÁČEK; JECH, 1999).

Classical sets are usually conceived, at least intuitively, as collections of objects. We may speak of the collection of glasses in the cupboard, or the cutlery in the drawer. In both situations we know implicitly that we are talking about objects in a kitchen. In some cases however, ambiguity may arise unless we state explicitly what set bounds our speech. For example, depending on the context the term *number* may be interpreted as a natural number (with or without zero), an integer or a real number. Hence it is often a good starting place to state precisely the set U of all the objects of which we are speaking,

[1] As an example, the minimum legal drinking age in Austria is 16 years old, whereas in the USA it is 21 years old. For information on countries around the globe, see <https://www.who.int/substance_abuse/publications/global_alcohol_report/profiles/en/>

which is often called the *universe of discourse*[2].

Recall that subsets of U may be represented by their characteristic functions, i.e., for each $A \in \mathcal{P}(U)$ we have $\chi_A \in 2^Y$ given by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} .$$

Fuzzy sets are defined by extending $2 = \{0, 1\}$ to $I = [0, 1] \subseteq \mathbb{R}$, thus enabling us to work with uncertain degrees of membership, i.e., rather than stating something or its negation (1 and 0, respectively), we are allowed to say something in between.

Definition 2.1.1. Let U be a (classical) set. A *fuzzy subset* F of U is a function[3]

$$\mu_F : U \to I \ .$$

The set of all fuzzy subsets of U is denoted by I^U.

Henceforth we shall frequently refer to fuzzy subsets as *fuzzy sets*, as long as there is no risk of confusion concerning the set U. It is a common practice to state that F is *characterized* by the function μ_F, which is called the *membership function* of F. In order to unify this practice and our definition, henceforth we shall consider a fuzzy set and its membership function to be the same entity, and at times we may say that a fuzzy set is characterized by its membership function. We may also denote the set of fuzzy subsets of U by $\mathcal{F}(U)$[4].

A fuzzy set μ such that $\mu(u) \in \{0, 1\}$ for every $u \in U$ is called a *crisp* set. If the set $U = \{u_1, ..., u_n\}$ is finite and A is a fuzzy subset of U we may write

$$A = \frac{\mu_A(u_1)}{u_1} + ... + \frac{\mu_A(u_n)}{u_n} \ .$$

Example 2.1.2. What does it mean to be young? Individuals a and b, aged 20 and 80 respectively, may have very different opinions about individual c, aged 40, being young.

2 By speaking about a "universe of discourse" we do not state that there is a "universal set", i. e., a set which contains all sets as elements of itself as one such set would be subject to Russell‘s paradox (HRBÁČEK; JECH, 1999) . On the contrary, the universe of discourse is a set conceived in order to avoid paradoxes.

3 Some clarification may be useful here. Many authors (such as (ZADEH, 1965) and (BARROS; BASSANEZI; LODWICK, 2017)) define F as being *characterized* by a function μ_F. We do understand that this is a good idea in terms of intuition, and we shall use this idea later for practical purposes. Nonetheless, this idea presents a philosophical difficulty: if F is *characterized* by μ_F, then what *is* F itself? We then decide to define a fuzzy set *to be* the function μ_F. Additionally, this definition is in accordance with the definition of L-fuzzy sets stated later in this chapter.

4 This is merely a notation. Although for classical sets there is a distinction between 2^U and $\mathcal{P}(U)$ and they are related by a bijection, in the fuzzy case I^U and $\mathcal{F}(U)$ are one and the same set.

Nonetheless they are all likely to agree that Mr. a is young, whereas Mr. b is not. Having this in mind, we may define μ_Y, the membership to the set of young people, as follows[5]:

Let $U = \{n \in \mathbb{N} : n \leqslant 120\}$. Define[6]

$$\mu_Y(x) = \begin{cases} 1, & \text{if } 0 \leqslant x \leqslant 20 \\ \dfrac{60-x}{40}, & \text{if } 20 < x \leqslant 60 \\ 0, & \text{if } 60 < x \leqslant 120 \end{cases} .$$

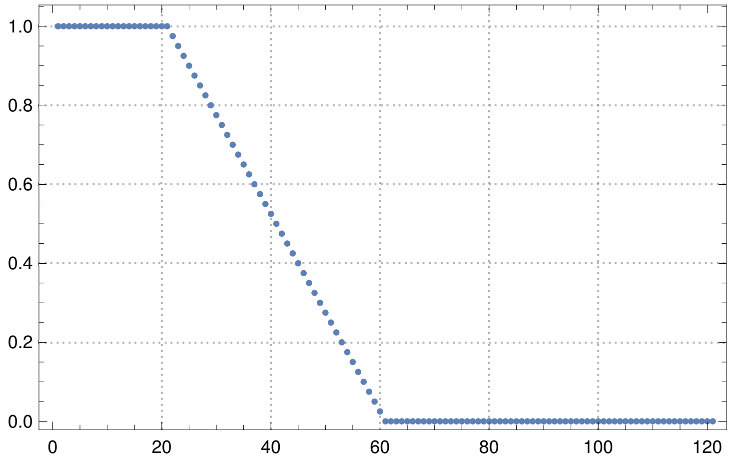


Figure 4 – Membership to the Fuzzy Set of Young People

Given two classical sets $A, B \subseteq U$, their intersection $(A \cap B)$ and union $(A \cup B)$ are expressed by the following characteristic functions:

$$\chi_{A \cap B}(x) = \begin{cases} 1, & \text{if } x \in A \text{ and } x \in B \\ 0, & \text{if } x \notin A \text{ or } x \notin B \text{ (or both)} \end{cases} ,$$

$$\chi_{A \cup B}(x) = \begin{cases} 0, & \text{if } x \notin A \text{ and } x \notin B \\ 1, & \text{if } x \in A \text{ or } x \in B \text{ (or both)} \end{cases} .$$

The complement of A (that is, the set $U \backslash A$) denoted by A', has the following characteristic function:

[5] We are aware that age alone is no sufficient criteria for defining someone as young or old. As sang by the children in the Mexican TV show *El Chavo del Ocho*, "There are young people in their eighties and there are old people who are 16 years old." Nonetheless we do choose to work only with age for simplicity.

[6] Notice that the points at $x = 20$ and $x = 60$ give us constraints. If we want to make the membership function (when extended to $[0, 120]$) of class C^n, each of these points give us two additional constraints, meaning that it could be achieved by fitting a $2n-1$ degree polynomial to the $2n$ constraints. However, high degree polynomials constrained to a limited interval may become wavy. If we used a function of the form ${}^{1}/_{1+e^{(r+sx)}}$ we would have a smooth function that approaches but never reaches 0 nor 1.

$$\chi_{A'}(x) = \begin{cases} 0, & \text{if } x \in A \\ 1, & \text{if } x \notin A \end{cases} .$$

Notice that for any $x \in U$ we have the following[7]:

$$\begin{aligned} \chi_{A \cap B}(x) &= \chi_A(x) \wedge \chi_B(x) \ , \\ \chi_{A \cup B}(x) &= \chi_A(x) \vee \chi_B(x) \ , \\ \chi_{A'}(x) &= 1 - \chi_A(x) \ . \end{aligned}$$

Now one sees that extending the classical definitions is an easy matter. The definitions presented below are called the *standard* fuzzy set operations.

Definition 2.1.3. Let A, B be fuzzy subsets of U. Then the fuzzy subsets $A \cap B$ and $A \cup B$, called the *intersection* and *union* of A and B respectively, and the set A', called the *complement* of A, have the following membership functions:

$$\begin{aligned} \mu_{A \cap B}(x) &= \mu_A(x) \wedge \mu_B(x) \ , \\ \mu_{A \cup B}(x) &= \mu_A(x) \vee \mu_B(x) \ , \\ \mu_{A'}(x) &= 1 - \mu_A(x) \ . \end{aligned}$$

When working with fuzzy sets, properties distinct from those of classical sets often arise.

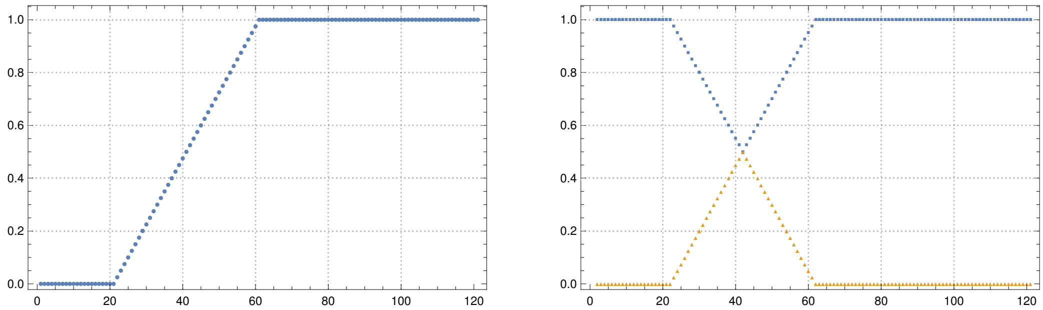
Example 2.1.4. From the set Y of young people defined in Example 2.1.2, we shall define the set O of old people as the complement Y' of Y (Fig. 5a), that is:

$$\mu_O(x) = 1 - \mu_Y(x) = \begin{cases} 0, & \text{if } \ 0 \leqslant x \leqslant 20 \\ \dfrac{x-20}{40}, & \text{if } 20 < x \leqslant 60 \\ 1, & \text{if } 60 < x \leqslant 120 \end{cases}$$

Thus, we get the following results for $Y \cup O$ and $Y \cap O$ (Fig. 5b):

$$\mu_{Y \cup O}(x) = \begin{cases} 1, & \text{if } \ 0 \leqslant x \leqslant 20 \\ \dfrac{60-x}{40}, & \text{if } 20 < x \leqslant 40 \\ \dfrac{x-20}{40}, & \text{if } 40 < x \leqslant 60 \\ 1, & \text{if } 60 < x \leqslant 120 \end{cases} \qquad \mu_{Y \cap O}(x) = \begin{cases} 1, & \text{if } \ 0 \leqslant x \leqslant 20 \\ \dfrac{60-x}{40}, & \text{if } 20 < x \leqslant 40 \\ \dfrac{x-20}{40}, & \text{if } 40 < x \leqslant 60 \\ 1, & \text{if } 60 < x \leqslant 120 \end{cases}$$

[7] Recall that $\wedge$ and $\vee$ are the meet (infimum) and join (supremum), defined in Sec. 0.2, on the lattice $2 = \{0, 1\}$ with the usual order of natural numbers.



(a) Fuzzy Set O of Old People (b) $Y \cup O$ (top) and $Y \cap O$ (bottom)

Figure 5 – Operations with Fuzzy Subsets

Notice that differently from the classical case, the union and intersection of a fuzzy set with its complement may not be the full set U or the empty set. For example, $\mu_{Y \cup O}(40) = 0.5 = \mu_{Y \cap O}(40)$.

As we have seen in Sec. 0.2[8], given $A, B \subseteq U$,

$$A \subseteq B \text{ iff } \chi_A \leqslant \chi_B \ ,$$

where $\chi_A \leqslant \chi_B$ expresses $\chi_A(u) \leqslant \chi_B(u)$ for all $u \in U$. We extend this criteria of subsetness to fuzzy sets.

Definition 2.1.5. Let A, B be fuzzy sets of U. We say that A is a *fuzzy subset* of B, denoted by $A \subseteq B$, iff

$$\mu_A \leqslant \mu_B \ ,$$

that is, iff for all $x \in U$, $\mu_A(x) \leqslant \mu_B(x)$[9].

Considering that we want to have, for every fuzzy subset A of U, the relations $\varnothing \subseteq A \subseteq U$, it follows that the membership functions of U and $\varnothing$ are pointwise given by $\mu_U(x) = 1$ and $\mu_\varnothing(x) = 0$, respectively.

2.2 Fuzzy relations

Recall that a classical n-place relation on a product $V_1 \times ... \times V_n$ is defined as a set $R \subseteq V_1 \times ... V_n$ (Def. 0.1.1). The notion of relation is present in the definition of a formal context (Def. 1.1.1), where I is a binary relation on $\mathcal{A} \times \mathcal{O}$. If we are to present a fuzzy theory of FCA, then it is clear that we shall need fuzzy relations.

8 See Prop. 0.2.16.

9 Recall from Prop. 0.2.16 that with this order $\langle I^U, \leqslant \rangle$ is a complete lattice in which, for every (classical) subset M of I^U we have

$$\left(\bigvee_{\mu \in M} \mu \right)(y) = \bigvee_{\mu \in M} \mu(y) \ , \qquad \left(\bigwedge_{\mu \in M} \mu \right)(y) = \bigwedge_{\mu \in M} \mu(y) \ ,$$

where the right-hand side of each equation corresponds to inf, sup in $[0, 1]$.

Definition 2.2.1. Let $U_1, ..., U_n$ be (classical) sets. A *n-place fuzzy relation* (or simply *fuzzy relation*) on $U = U_1 \times ... \times U_n$ is a fuzzy subset of U.

Example 2.2.2. Consider the idea of people belonging to different generations. It is no problem to think that a parent and a child belong to different generations, but we may say that in a sense two siblings born 15 years apart do not belong to the same generation. Their ages are (almost certainly) closer than a parent and a child's. Still, they are further apart than cousins born only 2 or 3 years apart One way to describe different generations is as follows.

Let A be a set of people and $f : A \to Y$ be a function that maps a person a to the year $f(a)$ when a was born. Define the fuzzy subset G of $A \times A$ by

$$\mu_G(a_1, a_2) = \bigvee \left\{0, 1 - \frac{|f(a_1) - f(a_2)|}{25}\right\} \ .$$

Then G is a fuzzy (binary) relation which accounts for membership of two people in A to the same generation. Here we consider that two people are considered to belong to the same generation if they are born in the same year, whereas they are of different generations if they are born 25 (or more) years apart

Recall that we defined an order (Def. 0.1.3) as a reflexive, transitive, antisymmetric relation. These properties (as well as other classical properties) can be extended to fuzzy sets.

Definition 2.2.3. Let U be a (classical) set and A be a binary fuzzy relation on U, i.e., a fuzzy subset of $U \times U$. We say that A is:

1. *reflexive* if $\mu_A(x, x) = 1$ for all $x \in U$;
2. *symmetric* if $\mu_A(x, y) = \mu_A(y, x)$ for all $x, y \in U$;
3. *transitive* if $[\mu_A(x, y) \wedge \mu_A(y, z)] \leqslant \mu_A(x, z)$ for all $x, y, z \in U$;
4. *antisymmetric* if $0 < \mu_A(x, y)$ and $0 < \mu_A(y, x)$ imply $x = y$ for all $x, y \in U$.

Example 2.2.4. Consider the fuzzy relation G of generations in Example 2.2.2. G is reflexive and symmetric. It is neither transitive nor antisymmetric.

Proof. 1. (Reflexivity) Every person a is born in the same year as itself, and so

$$\mu_G(a, a) = \bigvee \left\{0, 1 - \frac{|f(a) - f(a)|}{25}\right\} = \bigvee \{0, 1 - 0\} = \bigvee \{0, 1\} = 1 \ .$$

2. (Symmetry) For all $a_1, a_2 \in A$ we have

$$\mu_G(a_1, a_2) = \bigvee \left\{0, 1 - \frac{|f(a_1) - f(a_2)|}{25}\right\} = \bigvee \left\{0, 1 - \frac{|f(a_2) - f(a_1)|}{25}\right\} = \mu_G(a_2, a_1) \ .$$

3. (No transitivity) Let c_1, c_2 be siblings born in 1990 and 1985 respectively. Let p be c_1 and c_2's parent, born in 1965. Then

$$\mu_G(c_1, c_2) = \frac{4}{5} \qquad \mu_G(c_2, p) = \frac{1}{5} \qquad \mu_G(c_1, p) = 0$$

$$\mu_G(c_1, c_2) \wedge \mu_G(c_2, p) = \frac{1}{5} \qquad \mu_G(c_1, p) = 0$$

$$\boxed{[\mu_G(c_1, c_2) \wedge \mu_G(c_2, p)] > \mu_G(c_1, p)}$$

4. (No antisymmetry) Consider c_1 and c_2 of item 3.

□

Clearly only diagonal binary fuzzy relations[10] can be simultaneously symmetric and antisymmetric. In fact, if A is both symmetric and antisymmetric, let $x, y \in U$ be such that $0 < \mu_A(x, y)$. By symmetry, $0 < \mu_A(y, x)$, so that antisymmetry gives $x = y$.

Although a study of fuzzy orders an interesting topic[11], for our purposes crisp orders on sets of fuzzy functions suffice.

It is worth mentioning that an important family of subsets of a Cartesian product (thus a family of relations) is that of Cartesian products of subsets. For instance, item 4. of Theorem 1.1.10 informs us that, given $O \subseteq \mathcal{O}$, $A \subseteq \mathcal{A}$, the objects in O share all attributes in A — expressed as $O \subseteq A^{\wedge}$ — iff $O \times A \subseteq I$.

Now, if $S_i \subseteq V_i$ for each $i = 1, ..., n$, the relation $S = S_1 \times ... \times S_n$ on $V_1 \times ... \times V_n$ satisfies, for each n-tuple $(x_1, ..., x_n) \in V_1 \times ... \times V_n$, the equation

$$\chi_S(x_1, ..., x_n) = \bigwedge_{i=1}^{n} \chi_{S_i}(x_i) \quad ,$$

which motivates the following definition.

Definition 2.2.5. Let $U_1, ..., U_n$ be (classical) sets and let A_i be a fuzzy subset of U_i for each $i = 1, ..., n$. The *fuzzy Cartesian product* $A = A_1 \times ... \times A_n$ of the A_i is a fuzzy relation with membership function given by:

$$\mu_A(x_1, ..., x_n) = \bigwedge_{i=1}^{n} \mu_{A_i}(x_i) \quad .$$

Example 2.2.6. Let $C = \{c_1, c_2, c_3\}$ be a set of cars. Let the fuzzy subsets E, F of U, describing respectively *expensive cars* and *fast cars*, be given by the following:

$$\mu_E(c_1) = 0.1 \ , \qquad \mu_E(c_2) = 0.4 \ , \qquad \mu_E(c_3) = 0.9 \ ,$$
$$\mu_F(c_1) = 0.3 \ , \qquad \mu_F(c_2) = 0.5 \ , \qquad \mu_F(c_3) = 0.6 \ .$$

[10] That is, fuzzy relations on U^2 such that if $x \neq y$ then $\mu_A(x, y) = 0$.
[11] See (ŠEŠELJA; TEPAVČEVIĆ, 2007) .

Then the fuzzy Cartesian product $E \times F$ is as follows:

$$\begin{array}{lll} \mu_{E\times F}(c_1, c_1) = 0.1 \ , & \mu_{E\times F}(c_1, c_2) = 0.1 \ , & \mu_{E\times F}(c_1, c_3) = 0.1 \ , \\ \mu_{E\times F}(c_2, c_1) = 0.3 \ , & \mu_{E\times F}(c_2, c_2) = 0.4 \ , & \mu_{E\times F}(c_2, c_3) = 0.4 \ , \\ \mu_{E\times F}(c_3, c_1) = 0.3 \ , & \mu_{E\times F}(c_3, c_2) = 0.5 \ , & \mu_{E\times F}(c_3, c_3) = 0.6 \ . \end{array}$$

This Cartesian product satisfies none of the properties in Def. 2.2.3.

2.3 Extending classical logic

This section is divided in three parts. In the first part we introduce the *Statement Calculus* (SC), which deals with *statements* (sentences that are either true or false) and *connectives* that "paste" statements together creating new statements. Then we extend the classical connectives to fuzzy connectives. In the second part we consider what is the meaning of using the universal and existential quantifiers (∀ and ∃, respectively), and then we present their fuzzy correspondents. We refer the interested reader to (MARGARIS, 1990) and (MENDELSON, 2009) . These references further develop SC as well as the so called *First Order Predicate Calculus*, which builds up on SC and includes the universal and existential quantifiers. For the fuzzy counterparts of the logical connectives the reader is referred to (BARROS; BASSANEZI; LODWICK, 2017) and (KLEMENT et al., 2000) . Finally, we state definitions of continuity and semi-continuity of functions, and provide some basic results concerning continuous and semi-continuous t-norms and implications. These results will be crucial in the next chapter in order to prove theorems concerning an extended version of FCA.

2.3.1 Logical connectives

According to MARGARIS, "A *statement* is a declarative sentence that is either true or false (but not both)" (1990, p. 1). For example, the sentence

All swans are white

is a statement, but the sentence

When did you arrive here?

is not.

Statements can be connected with one another in ways that produce new statements. Given two statements P and Q, the ways in which statements are connected in classical logic are the following:

P	Q	$P \wedge Q$
0	0	0
1	0	0
0	1	0
1	1	1

(a) Conjunction

P	Q	$P \vee Q$
0	0	0
1	0	1
0	1	1
1	1	1

(b) Disjunction

P	Q	$P \longrightarrow Q$
1	1	1
1	0	0
0	1	1
0	0	1

(c) Implication

P	$\neg P$
0	1
1	0
0	0

(d) Negation

Table 2 – Truth Tables of Classical Connectives

1. **Conjunction:** is the statement "P and Q", which is true iff both P and Q are true. It is denoted by $P \wedge Q$.
2. **Disjunction:** is the statement "P or Q", which is true iff at least one of the statements P and Q is true. It is denoted by $P \vee Q$.
3. **Implication:** is the statement "If P then Q", which is devised to preserve truth: if P is true, so is Q. It is denoted by $P \longrightarrow Q$.

And finally, "connecting" one single statement P, we have:

4. **Negation:** is the statement "Not P", which is true iff P is false. It is denoted by $\neg P$.

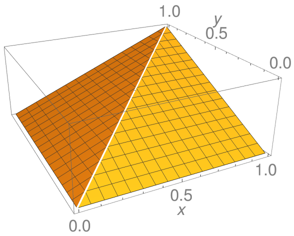
The attentive reader will have noticed that the symbols for conjunction and disjunction, which are usual in formal (mathematical) logic, are the same as those we used earlier for denoting the meet and join, related to lattices. This is no mere coincidence. Once we assume the values 1 for truth and 0 for falsehood, conjunction and disjunction have the *truth tables* presented in Tables 2a and 2b.

The last two rows of Table 2c may seem difficult to understand. A statement of the form $P \longrightarrow Q$ with P false is said to be *vacuously true*[12] (MARGARIS, 1990, p. 45). This principle lies behind the mathematical practice of assuming the truth of any statement concerning elements of the empty set: "if $a \in \varnothing$ then $R(a)$ for any property R", simply because $a \in \varnothing$ is always false. These rows entail traditional principle known as *ex falso quodlibet* (Latin for "anything follows from falsehood") or the *principle of explosion*: if something is false ($\neg P$), then it implies anything ($P \longrightarrow Q$), i.e.,

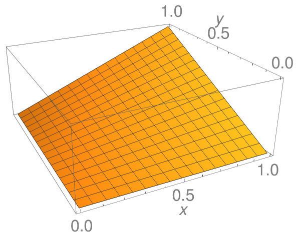
$$\neg P \longrightarrow (P \longrightarrow Q) \ .$$

We now extend these connectives. Triangular norms, which extends conjunction, and triangular conorms, which extends disjunction, are defined in accordance with (KLEMENT et al., 2000) . The examples of triangular norms and triangular conorms presented in this section can be found in the same reference.

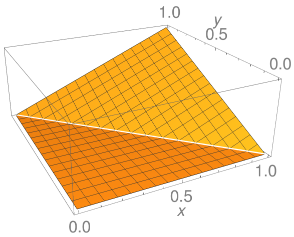
[12] This idea was already presented in Footnote 7 in Sec. 0.1.



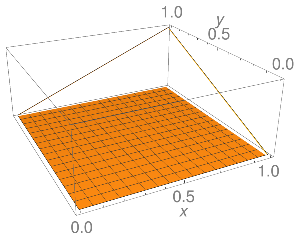
(a) Minimum



(b) Product



(c) Łukasiewicz



(d) Drastic product

Figure 6 – Triangular norms

Definition 2.3.1. A *triangular norm* (or *t-norm*) is a map $\triangle\colon [0,1]^2 \to [0,1]$ satisfying, for all $x, y, z \in [0,1]$:

1. $x \triangle y = y \triangle x$ (commutativity)
2. $x \triangle (y \triangle z) = (x \triangle y) \triangle z$ (associativity)
3. If $y \leqslant z$ then $x \triangle y \leqslant x \triangle z$ (monotonicity)
4. $x \triangle 1 = x$ (boundary condition)

Example 2.3.2. The maps $\triangle_{\mathrm{M}}, \triangle_{\mathrm{L}}, \triangle_{\mathrm{P}}, \triangle_{\mathrm{D}}\colon [0,1]^2 \to [0,1]$ defined below are t-norms. See Figure 6 for graphical representations of each of them.

1. $x \triangle_{\mathrm{M}} y = \min\{x, y\}$ (minimum)
2. $x \triangle_{\mathrm{P}} y = xy$ (product)
3. $x \triangle_{\mathrm{L}} y = \max\{x + y - 1, 0\}$ (Łukasiewicz t-norm)
4.

$$x \triangle_{\mathrm{D}} y = \begin{cases} 0, & \text{if } (x,y) \in [0,1[^2 \\ x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \end{cases} \qquad \text{(drastic product)}$$

Every t-norm is defined on the whole boundary of $[0,1]^2$. In fact, for each $x \in [0,1]$ we have

$$1 \triangle x = x \triangle 1 = x \ ,$$

and $0 \leqslant x \vartriangle 0 = 0 \vartriangle x \leqslant 0 \vartriangle 1 = 0$, whence

$$0 \vartriangle x = 0 = x \vartriangle 0 \ .$$

Furthermore, if $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$ then

$$x_1 \vartriangle y_1 \leqslant x_1 \vartriangle y_2 = y_2 \vartriangle x_1 \leqslant y_2 \vartriangle x_2 = x_2 \vartriangle y_2 \ .$$

Definition 2.3.3. Let $\vartriangle_1, \vartriangle_2$ be t-norms. If for all $x, y \in [0,1]$ we have $\vartriangle_1 \leqslant \vartriangle_2$ then we say that $\vartriangle_1$ is *weaker* than $\vartriangle_2$ (or equivalently $\vartriangle_2$ is *stronger* than $\vartriangle_1$), and we write $\vartriangle_1 \leqslant \vartriangle_2$.

If moreover $\vartriangle_1 \neq \vartriangle_2$, we write $\vartriangle_1 < \vartriangle_2$.

Proposition 2.3.4. *The t-norms of Example 2.3.2 have the following order:*

$$\vartriangle_{\mathrm{D}} < \vartriangle_{\mathrm{L}} < \vartriangle_{\mathrm{P}} < \vartriangle_{\mathrm{M}} \ .$$

Moreover, $\vartriangle_{\mathrm{D}}$ and $\vartriangle_{\mathrm{M}}$ are respectively the weakest and strongest t-norms.

Proof. Since all t-norms are equal on the boundary of $[0,1]^2$ and they only assume values on $[0,1]$ it is immediate that $\vartriangle_{\mathrm{D}}$ is the weakest t-norm. To show that $\vartriangle_{\mathrm{M}}$ is the strongest, consider a t-norm $\vartriangle$. Monotonicity yields

$$x \vartriangle y \leqslant x \vartriangle 1 = x \qquad\qquad x \vartriangle y \leqslant 1 \vartriangle y = y \ ,$$

so that

$$x \vartriangle y \leqslant \min\{x, y\} = x \vartriangle_{\mathrm{M}} y \ .$$

Thus, $\vartriangle_{\mathrm{M}}$ is the strongest t-norm.

Now, for all $x, y \in]0,1[$ we have

$$x(1-y) < (1-y) \text{ iff } x - xy < 1 - y \text{ iff } x + y - 1 < xy$$

and $0 < xy$. Hence we have

$$x \vartriangle_{\mathrm{L}} y < x \vartriangle_{\mathrm{P}} y \ .$$

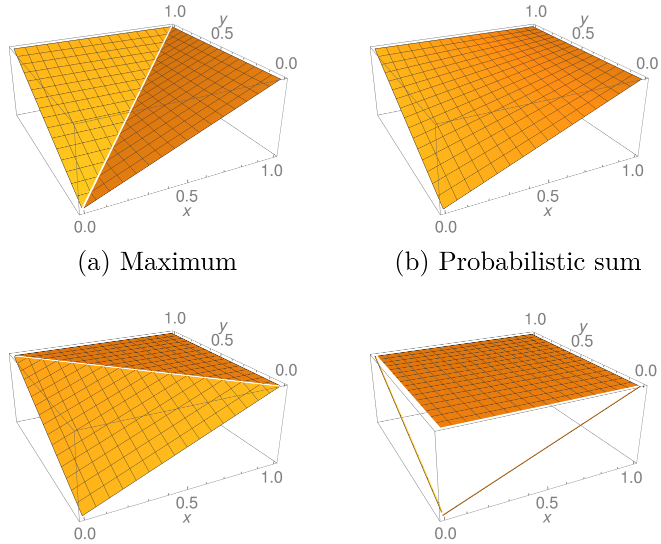
Therefore,

$$\vartriangle_{\mathrm{D}} < \vartriangle_{\mathrm{L}} < \vartriangle_{\mathrm{P}} < \vartriangle_{\mathrm{M}} \ .$$

□

We proceed with the extension of disjunction (maximum).

Definition 2.3.5. A *triangular conorm* (or *t-conorm*) is a map $\triangledown : [0,1]^2 \to [0,1]$ satisfying, for all $x, y, z \in [0,1]$:



(a) Maximum (b) Probabilistic sum

(c) Łukasiewicz (d) Drastic sum

Figure 7 – Triangular conorms

1. $x \triangledown y = y \triangledown x$ (commutativity)
2. $x \triangledown (y \triangledown z) = (x \triangledown y) \triangledown z$ (associativity)
3. $x \triangledown y \leqslant x \triangledown z$, if $y \leqslant z$ (monotonicity)
4. $x \triangledown 0 = x$ (boundary condition)

Example 2.3.6. The maps $\triangledown_{\mathrm{M}}, \triangledown_{\mathrm{L}}, \triangledown_{\mathrm{P}}, \triangledown_{\mathrm{D}} : [0,1]^2 \to [0,1]$ defined below are t-conorms. See Figure 7 for graphical representations of each of them.

1. $x \triangledown_{\mathrm{M}} y = \max\{x, y\}$ (maximum)
2. $x \triangledown_{\mathrm{P}} y = x + y - xy$ (probabilistic sum)
3. $x \triangledown_{\mathrm{L}} y = \min\{x + y, 1\}$ (Łukasiewicz t-conorm)
4. $$x \triangledown_{\mathrm{D}} y = \begin{cases} 1, & \text{if} \quad (x,y) \in]0,1]^2 \\ x, & \text{if} \quad y = 0 \\ y, & \text{if} \quad x = 0 \end{cases}$$ (drastic sum)

Notice that for each $j \in \{\mathrm{M}, \mathrm{P}, \mathrm{L}, \mathrm{D}\}$, we have

$$x \triangledown_j y = 1 - \left[(1-x) \vartriangle_j (1-y)\right] \tag{2.3.7}$$

for all $x, y \in [0,1]$. In fact, it can be shown that a map $\triangledown : [0,1]^2 \to [0,1]$ is a t-conorm iff there exists a t-norm $\vartriangle$ such that for all $x, y \in [0,1]$

$$x \triangledown y = 1 - [(1-x) \vartriangle (1-y)] \ . \tag{2.3.8}$$

This property, together with the boundary conditions for t-norms, tells us that every t-conorm assumes the same values on the boundary of $[0,1]$. Moreover, if we define order relations for t-conorms as we have defined them for t-norms, that is

$$\triangledown_1 \leqslant \triangledown_2 \text{ iff } x \triangledown_1 y \leqslant x \triangledown_2 y \text{ for all } x, y \in [0,1] \ ,$$

and if moreover $\triangledown_1 \neq \triangledown_2$ then

$$\triangledown_1 < \triangledown_2 \ ,$$

the following result holds.

Proposition 2.3.9. *The t-conorms of Example 2.3.6 have the following order:*

$$\triangledown_{\mathrm{M}} < \triangledown_{\mathrm{P}} < \triangledown_{\mathrm{L}} < \triangledown_{\mathrm{D}} \ .$$

Moreover, $\triangledown_{\mathrm{M}}$ and $\triangledown_{\mathrm{D}}$ are respectively the weakest and strongest t-conorms.

Proof. Let $\triangledown_1, \triangledown_2$ be t-conorms and let $\triangle_1, \triangle_2$ respectively be their dual t-norms, that satisfy (2.3.8). Suppose that $\triangle_1 \leqslant \triangle_2$. Then for all $x, y \in [0,1]$

$$\begin{aligned} 1 - x \triangledown_1 y &= (1-x) \triangle_1 (1-y) \\ &\leqslant (1-x) \triangle_2 (1-y) \\ &= 1 - x \triangledown_2 y \ , \end{aligned}$$

so that $\triangledown_2 \leqslant \triangledown_1$. The result follows from this, together with Proposition 2.3.4 and (2.3.7). □

For more details on t-norms and t-conorms, see (KLEMENT et al., 2000) .

Definition 2.3.10. A *fuzzy negation* (or simply *negation*) is a map $\nu : [0,1] \to [0,1]$ such that for all $x, y \in [0,1]$:

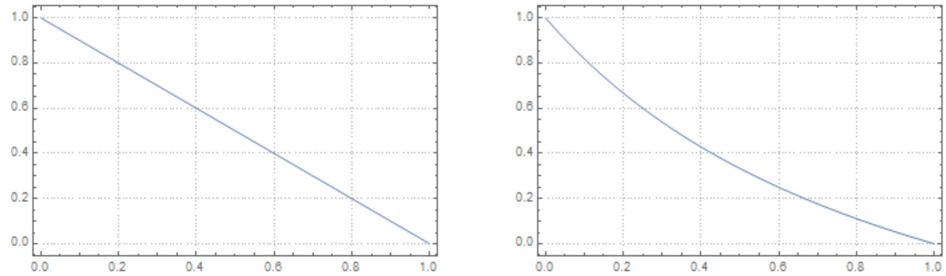
1. $\nu(0) = 1$ and $\nu(1) = 0$ (boundary conditions)
2. $\nu(\nu(x)) = x$ (involution)
3. If $x \leqslant y$ then $\nu(y) \leqslant \nu(x)$ (monotonicity)

Example 2.3.11. The following maps are negations (see Figure 8):

1. $\nu_1(x) = 1 - x$
2. $\nu_2(x) = \dfrac{1-x}{1+x}$

Notice that (2.3.8) expresses that there exists a one-one correspondence between t-norms and t-conorms that satisfy, for all $x, y \in [0,1]$, the identities

$$\nu_1 (x \triangle y) = \nu_1(x) \triangledown \nu_1(y) \ ,$$
$$\nu_1 (x \triangledown y) = \nu_1(x) \triangle \nu_1(y) \ ,$$



(a) $\nu_1(x) = 1 - x$ (b) $\nu_2(x) = \dfrac{1-x}{1+x}$

Figure 8 – Fuzzy negations

which extend the De Morgan laws.

Finally we define an extended implication.

Definition 2.3.12. A *fuzzy implication* is a map $\Rightarrow : [0,1]^2 \to [0,1]$ such that for all $x, y, z \in [0,1]$:

1. $(0 \Rightarrow 0) = 1$, $(1 \Rightarrow 0) = 0$, (boundary conditions)
   $(0 \Rightarrow 1) = 1$ and $(1 \Rightarrow 1) = 1$
2. If $y \leqslant x$ then $(x \Rightarrow z) \leqslant (y \Rightarrow z)$ (monotonicity in the first component)
3. If $x \leqslant y$ then $(z \Rightarrow x) \leqslant (z \Rightarrow y)$ (monotonicity in the second component)

In classical logic it is usual to define some connectives in terms of others. For example, the following three relations hold:

1. $P \longrightarrow Q = (\neg P) \vee Q$
2. $P \longrightarrow Q = \max\{x \in \{0,1\} : P \wedge x \leqslant Q\}$
3. $P \longrightarrow Q = (\neg P) \vee (P \wedge Q)$.

To see that equalities of items 1. and 3. hold, see the following truth table.

P	Q	$\neg P$	$(\neg P) \vee Q$	$P \wedge Q$	$(\neg P) \vee (P \wedge Q)$	$P \longrightarrow Q$
1	1	0	**1**	1	**1**	**1**
1	0	0	**0**	0	**0**	**0**
0	1	1	**1**	0	**1**	**1**
0	0	1	**1**	0	**1**	**1**

Now let us check that the equality of item 2. holds. If $P \leqslant Q$ it is clear that $P \wedge x \leqslant Q$ for $x \in \{0,1\}$. Thus, if $Q = 1$ or if $P = 0 = Q$ we have

$$\max\{x \in \{0,1\} : P \wedge x \leqslant Q\} = 1 \ .$$

On the other hand, if $P = 1$ and $Q = 0$ then $P \wedge 0 = 0 \leqslant Q$, but $P \wedge 1 = 1 \nleqslant Q$. Thus,

$$\max\{x \in \{0,1\} : P \wedge x \leqslant Q\} = 0 \ .$$

Hence, for all $P, Q \in \{0,1\}$ we have

$$\max\{x \in \{0,1\} : P \wedge x \leqslant Q\} = P \longrightarrow Q \ .$$

If on the right side of the first two equalities we replace classical connectives by their corresponding fuzzy connectives, with supremum rather than maximum, we get fuzzy implications. The third of these equalities, however, does not always give rise to a fuzzy implication as the following example shows.

Example 2.3.13. Let $\nu(x) = 1 - x$ and consider the minimum t-norm $\wedge$ and maximum t-conorm $\vee$. Let $P_1 = 0.9, P_2 = 1$ and $Q = 1$. Then

$$\begin{aligned} \nu(P_1) \vee (P_1 \wedge Q) &= \nu(0.9) \vee (0.9 \wedge 1) \\ &= 0.1 \vee 0.9 \\ &= 0.9 \ ; \end{aligned}$$

$$\begin{aligned} \nu(P_2) \vee (P_2 \wedge Q) &= \nu(1) \vee (1 \wedge 1) \\ &= 0 \vee 1 \\ &= 1 \ . \end{aligned}$$

Definition 2.3.14. A *S-implication* is an implication $\Rightarrow_S$ defined by

$$(x \Rightarrow_S y) = \nu(x) \,\triangledown\, y \ ,$$

where ν is a fuzzy negation and $\triangledown$ is a t-conorm.

A *R-implication* (or *residual implication*, or simply *residuum*) is an implication $\Rightarrow_R$ defined by

$$(x \Rightarrow_R y) = \sup\{z \in [0,1] : x \,\triangle\, z \leqslant y\} \ ,$$

where $\triangle$ is a t-norm. We say that $\Rightarrow$ is the *residuum* of $\triangle$.

Notice that for the residuum $\Rightarrow_R$ of $\triangle$, if $x \leqslant y$ then $x \,\triangle\, 1 = x \leqslant y$, and so $(x \Rightarrow_R y) = 1$.

Example 2.3.15. In the following examples, the first two items correspond to S-implications. The other items are R-implications.Figure 9 illustrates the fact that, increasing y from the diagonal $x = y$, the value of a residuum always equals 1, but this must not be true for S-implications.

1. The map $\Rightarrow_{\mathrm{kd}}$ defined as

$$(x \Rightarrow_{\mathrm{kd}} y) = (1-x) \triangledown_{M} y$$

is a S-implication, called the *Kleene-Dienes* implication, with the negation $\nu(x) = 1-x$ and the maximum t-conorm.

2. If we consider the negation ν_2 of Example 2.3.11 and the Łukasiewicz t-conorm, we define the S-implication

$$(x \Rightarrow_{S} y) = \left(\frac{1-x}{1+x} + y\right) \wedge 1$$

3. The map $\Rightarrow_{\mathrm{g}}$, defined as

$$(x \Rightarrow_{\mathrm{g}} y) = \begin{cases} 1, & \text{if } x \leqslant y \\ y, & \text{if } x > y \end{cases}$$

and called *Gödel implication* is the residuum of the minimum t-norm.

4. The map $\Rightarrow_{\mathrm{gn}}$, defined as

$$(x \Rightarrow_{\mathrm{gn}} y) = \begin{cases} 1, & \text{if } x \leqslant y \\ y/x, & \text{if } x > y \end{cases}$$

and called *Goguen implication* is the residuum of the product t-norm.

5. The map $\Rightarrow_{\mathrm{ł}}$ defined as

$$(x \Rightarrow_{\mathrm{ł}} y) = (1-x+y) \wedge 1$$

is the residuum of the Łukasiewicz t-norm, called the *Łukasiewicz implication.*

6. The map $\Rightarrow_{\mathrm{d}}$ defined as

$$(x \Rightarrow_{\mathrm{d}} y) = \begin{cases} 1, & \text{if } x < 1 \\ y, & \text{if } x = 1 \end{cases}$$

is the residuum of the drastic product t-norm.

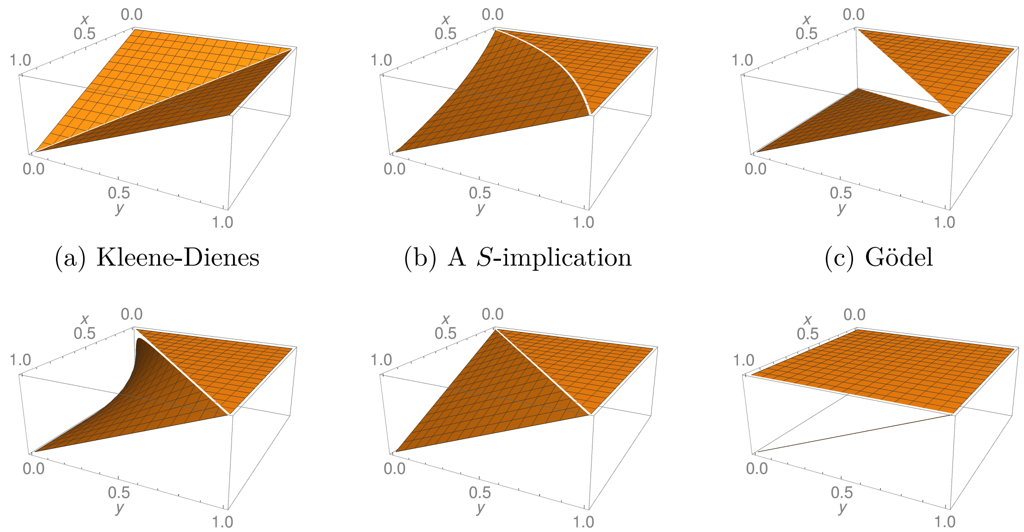
Now that we have extended the classical connectives, we can also extend the notion of Cartesian product, introduced in Definition 2.2.5.

Definition 2.3.16. Let U_1, U_2 be (classical) sets, and let A_1, A_2 be a fuzzy subsets of U_1, U_2 respectively. Let $\triangle$ be a t-norm. Then the *Cartesian product of* A_1, A_2 *induced by* $\triangle$ is the fuzzy set $A_1 \times_{\triangle} A_2$ that has membership function

$$\mu_{A_1 \times_{\triangle} A_2}(x_1, x_2) = \mu_{A_1}(x_1) \,\triangle\, \mu_{A_2}(x_2) \ .$$

If $A_1, ..., A_n, A_{n+1}$ are fuzzy subsets of the classical sets $U_1, ..., U_n, U_{n+1}$ respectively, and $A_1 \times_{\triangle} ... \times_{\triangle} A_n$ has already been defined, then

$$A_1 \times_{\triangle} ... \times_{\triangle} A_n \times_{\triangle} A_{n+1} := (A_1 \times_{\triangle} ... \times_{\triangle} A_n) \times_{\triangle} A_{n+1} \ .$$



(a) Kleene-Dienes (b) A S-implication (c) Gödel

(d) Goguen (e) Łukasiewicz (f) Drastic product

Figure 9 – Fuzzy implications

Having extended the classical connectives to a fuzzy setting, we make a brief discussion on how to "extend" the quantifiers ($\forall$ and $\exists$), which shall be necessary when we develop fuzzy FCA. For a deeper investigation of $\exists$ and $\forall$, and the so called First Order Predicate Calculus, See (MARGARIS, 1990) and (MENDELSON, 2009).

2.3.2 Predicates and quantification

According to ANGIONI (2006, pp. 17–18),

> By *predication* it is understood the statement that (i) has the form "S is P" or some equivalent form reducible to that, and (ii) intends to report given facts in the world [...] The basic structure of predication, as proposed by Aristotle, is constituted of a minimum of three elements: two terms (one of which is the *subject* and the other the *predicate*) and the copulative operator.[13]

The statements

The snow is white

and

$$2 + 2 = 4$$

[13] "Por *predicação*, entende-se o enunciado que (i) possui a forma 'S é P' ou alguma forma equivalente e redutível àquela, (ii) pretende repostar-se a fatos dados no mundo [...] A estrutura básica da predicação, tal como proposta por Aristóteles, constitui-se de três elementos mínimos: dois termos (sendo um deles o *sujeito* e o outro o *predicado*) e o operador copulativo."

are examples of predication (in the second case, the sentence can be read as "2+2 is equal to 4").

In symbolic (mathematical) logic it is usual to consider the whole predicative sentence as a predicate. In this case, each term is an element of a given set, and a predicate is the characteristic function of a relation on the sets under consideration. When speaking about predicates, we shall denote the characteristic function of n-place relation P applied to the point $(x_1, ..., x_n)$ by $P(x_1, ..., x_n)$. For instance, consider the sets

$$W := \{\text{"Things made of water"}\} \ ,$$
$$C := \{\text{"Colours"}\} \ .$$

Let

$$P := \{(x, y) \in W \times C : x \text{ has colour } y\} \ .$$

Then the predicate

The snow is white

is true as $P(\text{snow}, \text{white})$ holds.

At times we may want to state that non-specified elements of the set satisfy certain predicate. For instance, we may want to say that everything that is made of water has a colour, i.e.

$$\forall x \exists y P(x, y) \ .$$

The operators for universality ("everything", "each", expressed symbolically as $\forall$) and existence ("some", "at least one", written as $\exists$) are called *quantifiers* ("universal quantifier" and "existential quantifier," respectively).

Although $P(x, y)$ is not a statement, since the elements x and y have not been specified, it does become a statement when each of its variables is either replaced by an element of the corresponding set (e.g., $P(\text{snow}, \text{white})$) or is under the scope of a quantifier (e.g., $\exists y P(\text{snow}, y)$, i.e., "snow has a colour"), and so all the theory of the statement calculus apply to it.

As we have already seen, a fuzzy relation is a fuzzy subset of the Cartesian product of the sets under consideration. Thus, we shall express fuzzy predicates by membership functions. But how do we extend the use of quantifiers?

Remember that we use real numbers, 0 and 1, to express truth and falsehood. Now a statement $\forall x Q(x)$ is true iff $Q(a)$ holds for each a in the set. If $Q(a_0)$ is false for at least one $a_0 \in U$, then the statement $\forall x Q(x)$ is false. Thus, classically we have

$$\forall x Q(x) = \min_{a \in U} Q(a) \ .$$

A similar argument leads us to the classical expression

$$\exists x Q(x) = \max_{a \in U} Q(a) \ .$$

Thus, the way in which we extend the universal and existential quantifiers is by taking, respectively, the infimum and supremum of the membership function over its domain, as inf extends min and sup extends max.

As we shall see in Ch. 3, what we need to preserve the theorems of FCA presented in Ch. 1 are the residua of lower semicontinuous t-norms. The next section is dedicated to define continuity, semicontinuity and prove some results that will be used in Ch. 3.

2.4 Continuity and semicontinuity

In our development of a fuzzy extension of FCA we shall need some ideas concerning continuity and semicontinuity of t-norms and implications. In this section we shall define continuity, semicontinuity and find how semicontinuity of a t-norm relates to its residuum. Our definitions of semicontinuity are in accordance with (BOURBAKI, 1966).

Definition 2.4.1. Let $p \in \mathbb{N}$. The *distance* between two points $x = (x_1, \ldots, x_p)$, $y = (y_1, \ldots, y_p) \in \mathbb{R}^p$ is defined as

$$d_p(x, y) = \sqrt{\sum_{i=1}^{p} (x_i - y_i)^2} \ .$$

In particular, $d_0(x, y) = 0$.

Definition 2.4.2. Given a point $x \in \mathbb{R}^p$ and $r > 0$, the *open ball of centre x and radius r* is the set $B_r(x) := \{y \in \mathbb{R}^p : d_p(x, y) < r\}$. A set $S \subseteq \mathbb{R}^p$ is *open* iff it is the (arbitrary) union of open balls. A *neighbourhood* of a set $A \subseteq \mathbb{R}^p$ is a set V such that $A \subseteq S \subseteq V$, such that S is an open set. In particular, S is a neighbourhood of A. A *neighbourhood of a point a* is a neighbourhood of $A = \{a\}$.

In other words, an open ball of centre x is the collection of those points close to x (according to the distance d_p), and a neighbourhood of a set (or a point) A corresponds to the "surroundings" of A.

Definition 2.4.3. A *sequence* on a set X is a map $\mathbb{N} \to X$. If for each $n \in \mathbb{N}$ we have $n \mapsto x_n$ we write (x_n) or $(x_n)_{n \in \mathbb{N}}$.

Now, suppose that X is an ordered set. If for all $n \in \mathbb{N}$ we have $x_n \leqslant x_{n+1}$ we say that the sequence is *increasing*. If it additionally happens for all $n \in \mathbb{N}$ that $x_n \neq x_{n+1}$ — i.e., $x_n < x_{n+1}$ — the we say that the sequence is *strictly increasing*.

Decreasing and strictly decreasing sequences are defined dually.

Definition 2.4.4. We say that a sequence (x_n) on $\mathbb{R}^p$ *converges* to x if for every open ball B containing x, there exists $n_0 \in \mathbb{N}$ such that $x_n \in B$ for all $n \geqslant n_0$, and we write $x_n \to x$ or $\lim\limits_{n\to\infty} x_n = x$.

If $p = 1$ so that $\mathbb{R}^p$ is (linearly) ordered and (x_n) is decreasing (resp. increasing) such that $x_n \to x$, we may write $x_n \searrow x$ (resp. $x_n \nearrow x$).

Example 2.4.5. Consider the following sequences on $\mathbb{R}$.

1. The sequence (see Fig. 10a)
$$\left(\frac{1}{n+1}\right)$$
is a strictly decreasing sequence that converges to 0, i.e., ${}^1\!/\!{}_{(n+1)} \searrow 0$.
2. The sequence (see Fig. 10b)
$$(\ln(n+1))$$
is an increasing diverging[14] sequence, where ln is the natural logarithm.
3. The sequence (see Fig. 10c)
$$\left(\sum_{k=0}^{n} {}^{(-1)^k}\!/\!{}_{k+1}\right)_{n\in\mathbb{N}}$$
is neither increasing nor decreasing, yet it converges to $\ln 2$.

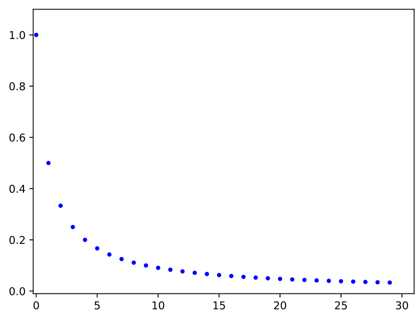
Definition 2.4.6. Let $X \subseteq \mathbb{R}^p$ and $Y \subseteq \mathbb{R}^q$ be non-empty (classical) sets. Let $f : X \to Y$ be a map. We say that f is *continuous at* $x \in X$ if, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $y \in X$,
$$d_p(x, y) < \delta \text{ implies } d_q\left(f(x), f(y)\right) < \varepsilon \ .$$

In other words, f is continuous at x if, for $y \in X$ sufficiently close to x — distance less than δ —, the points $f(x), f(y)$ are close in Y — less than ε apart. If f is continuous at every $x \in X$ it is said to be *continuous*.

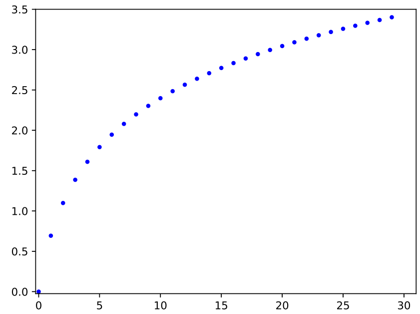
In our investigation based on t-norms, t-conorms and implications, we have $X = [0, 1]^2$ and $Y = [0, 1]$. We shall denote d_2 by d and $d_1(x, y)$ by $|x - y|$.

Example 2.4.7. The maximum t-conorm $\triangledown_{\text{M}}$ is continuous. Consequently, the minimum t-norm $\triangle_{\text{M}}$ is also continuous. In fact, let $x = (x_1, x_2) \in [0, 1]^2$, $\varepsilon > 0$ and take $\delta = \varepsilon$. Let

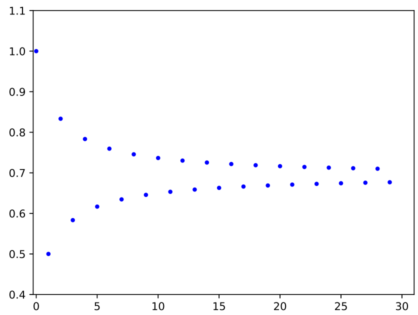
[14] It is *unbounded above*, i.e., for all $k \in \mathbb{N}$ there is an $n \in \mathbb{N}$ such that $\ln(n + 1) > k$. A sequence (x_n) is called *unbounded below* iff $(-x_n)$ is unbounded above. A sequence is *unbounded* iff it is either unbounded above or below (or both).



(a) Strictly Decreasing



(b) Logarithm



(c) Partial Sum

Figure 10 – Sequences

$y = (y_1, y_2) \in [0,1]^2$ be such that $d(x,y) < \delta$. Define $m_x := x_1 \triangledown_{\mathrm{M}} x_2$ and $m_y := y_1 \triangledown_{\mathrm{M}} y_2$. Since $m_x, m_y \geqslant 0$ we have

$$x_1 y_1 \leqslant m_x y_1 \leqslant m_x m_y \quad .$$

Similarly, $x_2 y_2 \leqslant m_x m_y$, and so adding and multiplying by -2 we have

$$-4 m_x m_y \leqslant -(2x_1 y_1 + 2x_2 y_2) \quad .$$

Thus,

$$\begin{aligned}
(|m_x - m_y|)^2 &= (m_x - m_y)^2 \\
&\leqslant 2(m_x - m_y)^2 \\
&= 2m_x^2 + 2m_y^2 - 4m_x m_y \\
&\leqslant (x_1^2 + x_2^2) + (y_1^2 + y_2^2) - (2x_1 y_1 + 2x_2 y_2) \\
&= (x_1^2 + y_1^2 - 2x_1 y_1) + (x_2^2 + y_2^2 - 2x_2 y_2) \\
&= (x_1 - y_1)^2 + (x_2 - y_2)^2 \\
&= d(x,y)^2 \\
&< \delta^2 = \varepsilon^2 \quad .
\end{aligned}$$

Since $\varepsilon > 0$, taking the square root on both sides we get

$$|m_x - m_y| < \varepsilon \ ,$$

and so $\triangledown_{\mathrm{M}}$ is continuous at x. But x is arbitrary, hence $\triangledown_{\mathrm{M}}$ is continuous.

Now, according to (2.3.7),

$$x \triangledown_{\mathrm{M}} y = 1 - [(1-x) \vartriangle_{\mathrm{M}} (1-y)] \ ,$$

that is, $\vartriangle_{\mathrm{M}}$ is a composition of continuous functions (the continuous $\triangledown_{\mathrm{M}}$ and differences), hence it is continuous.

Continuity of a t-norm does not imply continuity of its residuum — for example, the residuum of the continuous minimum t-norm is the discontinuous Gödel implication[15]. Nonetheless, we can say something about such implications. The following definition is made by (BOURBAKI, 1966).

Definition 2.4.8. Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}$ and $f : X \to Y$. We say that f is *lower semicontinuous at* $x_0 \in X$ if for h such that $h < f(x_0)$ there is a neighbourhood V of x_0 such that

$$h < f(x) \text{ for all } x \in V \ .$$

If f is upper semicontinuous at each $x_0 \in X$ we say that it is *upper semicontinuous*.

The notion of *upper semicontinuity* (either at a point, or of the whole function) is defined dually[16].

Notice that f is lower semicontinuous iff $-f$ is upper semicontinuous[17]. Furthermore, see that a map is continuous at x_0 iff it is both lower and upper semicontinuous at x_0.

BOURBAKI also proves[18] the following.

Proposition 2.4.9. *A function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is lower semicontinuous at $x_0 \in X$ iff*

$$\liminf_{n\to\infty} f(x_n) \geqslant f(x_0)$$

[15] Indeed, $\Rightarrow_{\mathrm{g}}$ is discontinuous at (x,x) for all $x \in [0,1[$. In fact, let $\varepsilon = 1 - x$. Then for all $\delta > 0$, $(x, x - \delta/2) \in B_\delta(x,x)$, but

$$\left|(x \Rightarrow_{\mathrm{g}} x) - \left(x \Rightarrow_{\mathrm{g}} \left(x - \frac{\delta}{2}\right)\right)\right| = \left|1 - \left(x - \frac{\delta}{2}\right)\right| = \left[\left(1 + \frac{\delta}{2}\right) - x\right] > 1 - x = \varepsilon \ .$$

[16] I.e., replace each instance of $<$ by $>$.

[17] Where $-f$ is the map $x \mapsto -f(x)$.

[18] In fact, Bourbaki's result is more general, but Prop. 2.4.9 follows when considering the *Fréchet filter* on $\mathbb{N}$, i.e., the complements of finite subsets of $\mathbb{N}$.

for each sequence $(x_n)_{n\in\mathbb{N}}$ *converging to* x_0*, where*

$$\liminf_{n\to\infty} k_n := \inf_{n>0}\left(\sup_{m\geqslant n} k_m\right)$$

for any sequence $(k_n)_{n\in\mathbb{N}}$*.*

Because of the duality between lower and upper semicontinuousness[19], we have the following.

Corollary 2.4.10. *A function* $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ *is upper semicontinuous at* $x_0 \in X$ *iff*

$$\limsup_{n\to\infty} f(x_n) \leqslant f(x_0)$$

for each sequence $(x_n)_{n\in\mathbb{N}}$ *converging to* x_0*, where*

$$\limsup_{n\to\infty} k_n := \sup_{n>0}\left(\inf_{m\geqslant n} k_m\right)$$

for any sequence $(k_n)_{n\in\mathbb{N}}$*.*

We now derive a useful expression concerning lower semicontinuous t-norms in Lemma 2.4.14, and for its proof we use Proposition 2.4.11.

Proposition 2.4.11. *Let* $\triangle$ *be a lower semicontinuous t-norm and* $x, y \in [0,1]$*. Let*

$$A = \{z \in [0,1] : x \triangle z \leqslant y\} \ .$$

Then $\sup A \in A$*.*

Proof. Let $z_0 = \sup A$ and (z_n) be a strictly increasing sequence on $[0,1]$ that converges to z_0. By definition of z_0 we have

$$x \triangle z_m \leqslant y \text{ for all } m > 0 \ ,$$

as $z_m < z_0$ for each $m > 0$. Now, for each $n > 0$ we have

$$\inf_{m\geqslant n}(x \triangle z_m) \leqslant x \triangle z_n \leqslant y \ ,$$

so that

$$\sup_{n>0}\left(\inf_{m\geqslant n}(x \triangle z_n)\right) \leqslant y \ .$$

By lower semicontinuity of $\triangle$,

$$x \triangle z_0 \leqslant \liminf_{n\to\infty}(x \triangle z_n) = \sup_{n>0}\left(\inf_{m\geqslant n}(x \triangle z_m)\right) \leqslant y \ .$$

Therefore, $z_0 \in A$. □

[19] In the sense that f is lower semicontinuous iff $-f$ is upper semicontinuous.

Corollary 2.4.12 and Lemma 2.4.14 are based on (HÁJEK, 1998) .

Corollary 2.4.12. *Let $\triangle$ be a lower semicontinuous t-norm and let $\Rightarrow$ be its residuum. Then for all $x, y, z \in [0,1]$ we have*

$$x \triangle z \leqslant y \text{ iff } z \leqslant (x \Rightarrow y) \quad . \tag{2.4.13}$$

Proof. Suppose that $x \triangle z \leqslant y$. Then

$$z \in \{z_0 \in [0,1] : x \triangle z_0 \leqslant y\} \quad ,$$

whence

$$z \leqslant \sup\{z_0 \in [0,1] : x \triangle z_0 \leqslant y\} = (x \Rightarrow y) \quad .$$

Conversely, if $z \leqslant (x \Rightarrow y)$ then either

1. $z < (x \Rightarrow y)$, that is

$$z < \sup\{z_0 \in [0,1] : x \triangle z_0 \leqslant y\} \quad , \text{ whence}$$
$$z \in \{z_0 \in [0,1] : x \triangle z_0 \leqslant y\} \quad ; \text{ or}$$

2. $z = (x \Rightarrow y)$, so that

$$z \in \{z \in [0,1] : x \triangle z \leqslant y\}$$

by Proposition 2.4.11.

In both cases, $x \triangle z \leqslant y$. □

The maps $\triangle$ and $\Rightarrow$ satisfying (2.4.13) are called an *adjoint pair*.

Lemma 2.4.14. *Let $\triangle : [0,1]^2 \to [0,1]$ be a lower semicontinuous t-norm and let $\Rightarrow$ be its residuum. Then, for all $x, y \in [0,1]$,*

$$x \leqslant [(x \Rightarrow y) \Rightarrow y] \quad .$$

Proof. Let $x, y \in [0,1]$. Notice that $(x \Rightarrow y) = \sup A$ in Proposition 2.4.11, and so, $x \triangle (x \Rightarrow y) \leqslant y$. Using commutativity of $\triangle$, it is clear that

$$x \in B := \{z \in [0,1] : (x \Rightarrow y) \triangle z \leqslant y\} \quad .$$

Hence,

$$x \leqslant \sup B = [(x \Rightarrow y) \Rightarrow y] \quad .$$

□

Lower semicontinuity of $\triangle$ is a necessary hypothesis for Lemma 2.4.14 as the following example shows.

Example 2.4.15. Let $\Rightarrow_{\mathrm{d}}$ be the residuum of the drastic product t-norm, as defined in item 6 of Example 2.3.15. Let $x = \dfrac{2}{3}$ and let $y = \dfrac{1}{3}$. Then

$$(x \Rightarrow_{\mathrm{d}} y) = \left(\frac{2}{3} \Rightarrow_{\mathrm{d}} \frac{1}{3}\right) = 1 \ ,$$

as $x > y$. Thus,

$$[(x \Rightarrow_{\mathrm{d}} y) \Rightarrow_{\mathrm{d}} y] = \left[1 \Rightarrow_{\mathrm{d}} \frac{1}{3}\right] = \frac{1}{3} < \frac{2}{3} = x \ ,$$

and $\Rightarrow_{\mathrm{d}}$ does not satisfy Lemma 2.4.14.

Lemma 2.4.16 and Prop. 2.4.17 are proved by (BRITO et al., 2018).

Lemma 2.4.16. *Let $X, Y \subseteq \mathbb{R}$, $f : X \to Y$ be an increasing function, (x_n) be a sequence on X such that $x_n \searrow x \in X$ and (y_n) be a sequence on Y.*

If for all $n > 0$, $f(x_n) \leqslant y_n$, then

$$f(x) \leqslant \liminf_{n\to\infty} y_n \leqslant \limsup_{n\to\infty} y_n \ .$$

Proof. Suppose that $f(x_n) \leqslant y_n$ for all $n > 0$. Since (x_n) is decreasing, for each $n > 0$ we have

$$\sup_{j \geqslant n} x_j = x_n \ ,$$

so that

$$\inf_{k>0}\left(\sup_{j\geqslant k} x_j\right) \leqslant x_n \ ,$$

so that we have

$$x = \limsup_{j\to\infty} x_j \leqslant x_n \ .$$

as (x_n)

Thus, for each $n > 0$,

$$f(x) \leqslant f(x_n) \leqslant y_n \ .$$

Now, taking the limit inferior on the right-hand side, we conclude the proof.

□

Proposition 2.4.17. *Let $\triangle$ be a lower semicontinuous t-norm, and let $\Rightarrow$ be the residuum of $\triangle$. Then for each $x_0, y_0 \in [0,1]$ the maps $x \mapsto (x \Rightarrow y_0)$ and $y \mapsto (x_0 \Rightarrow y)$ are upper semicontinuous.*

Proof. Let $x', y' \in [0,1]$. Let $(x_n), (y_n)$ be sequences on $[0,1]$ converging respectively to x' and y'. For each $n > 0$, consider the following definitions:

$$z_n^{(1)} = (x_n \Rightarrow y_0) \;; \qquad z_n^{(2)} = (x_0 \Rightarrow y_n) \;;$$
$$\tilde{x}^{(1)} = x' \;; \qquad \tilde{x}^{(2)} = x_0 \;;$$
$$\tilde{x}_n^{(1)} = \inf_{m \geqslant n} x_m \;; \qquad \tilde{x}_n^{(2)} = x_0 \;;$$
$$\tilde{y}^{(1)} = y_0 \;; \qquad \tilde{y}^{(2)} = y' \;;$$
$$\tilde{y}_n^{(1)} = y_0 \;; \qquad \tilde{y}_n^{(2)} = \sup_{m \geqslant n} y_m \;;$$
$$\tilde{z}_n^{(1)} = \left(\tilde{x}_n^{(1)} \Rightarrow \tilde{y}_n^{(1)}\right) \;; \qquad \tilde{z}_n^{(2)} = \left(\tilde{x}_n^{(2)} \Rightarrow \tilde{y}_n^{(2)}\right) \;.$$

Notice the following:

1. The sequence $(\tilde{x}_n^{(1)})$ is constructed as the infima of the decreasing sets $A_n = \{x_m : m \geqslant n\}$. Thus, $(\tilde{x}_n^{(1)})$ is increasing. Furthermore as n increases the elements left in each A_n get always closer to x'. Thus, $\tilde{x}_n^{(1)} \nearrow x'$. In general, $\tilde{x}_n^{(i)} \nearrow \tilde{x}^{(i)}$, for $i = 1, 2$.
2. By an argument analogous to that of item 1., $\tilde{y}_n^{(i)} \searrow \tilde{y}^{(i)}$, for $i = 1, 2$.
3. For each $i = 1, 2$, considering monotonicity of $\Rightarrow$ in each component, one sees that for all $n > 0$, $z_n^{(i)} \leqslant \tilde{z}_n^{(i)}$. Thus, $\limsup_{n\to\infty} z_n^{(i)} \leqslant \limsup_{n\to\infty} \tilde{z}_n^{(i)}$.
4. The sequence $(\tilde{z}_n^{(i)})$ is decreasing, for $i = 1, 2$.

In the following, $i = 1, 2$. By Proposition 2.4.11, for each $n > 0$,

$$\tilde{x}_n^{(i)} \bigtriangleup \tilde{z}_n^{(i)} \leqslant \tilde{y}_n^{(i)}$$

as $\bigtriangleup$ is lower semicontinuous. For $n_0 > 0$ fixed and by monotonicity of $\bigtriangleup$ we have, for each $n > n_0$,

$$\tilde{x}_{n_0}^{(i)} \bigtriangleup \tilde{z}_n^{(i)} \leqslant \tilde{y}_n^{(i)} \;.$$

Thus, an application of Lemma 2.4.16 yields

$$\tilde{x}_{n_0}^{(i)} \bigtriangleup \left(\limsup_{n\to\infty} \tilde{z}_n^{(i)}\right) \leqslant \liminf_{n\to\infty} \tilde{y}_n^{(i)} = \tilde{y}^{(i)} \;.$$

But $n_0 > 0$ is arbitrary, and so

$$\liminf_{n\to\infty} \left[\tilde{x}_n^{(i)} \bigtriangleup \left(\limsup_{n\to\infty} \tilde{z}_n^{(i)}\right)\right] \leqslant \tilde{y}^{(i)} \;.$$

Remember that $\bigtriangleup$ is lower semicontinuous and so, since $(\tilde{x}_n^{(i)})$ converges to $\tilde{x}^{(i)}$, we see that

$$\tilde{x}^{(i)} \bigtriangleup \left(\limsup_{n\to\infty} \tilde{z}_n^{(i)}\right) \leqslant \tilde{y}^{(i)} \;.$$

Hence we have

$$\limsup_{n\to\infty} \tilde{z}_n^{(i)} \leqslant \sup\{z \in [0,1] : \tilde{x}^{(i)} \vartriangle z \leqslant \tilde{y}^{(i)}\}$$
$$= \left(\tilde{x}^{(i)} \Rightarrow \tilde{y}^{(i)}\right) \quad .$$

Thus,

$$\limsup_{n\to\infty} z^{(i)} \leqslant \left(\tilde{x}^{(i)} \Rightarrow \tilde{y}^{(i)}\right) \quad .$$

For $i = 1$ this means $x \mapsto (x \Rightarrow y_0)$ is upper semicontinuous at x', and for $i = 2$, $y \mapsto (x_0 \Rightarrow y)$ is upper semicontinuous at y'. Since $x', y' \in [0,1]$ are arbitrary, both the maps $x \mapsto (x \Rightarrow y_0)$ and $y \mapsto (x_0 \Rightarrow y)$ are upper semicontinuous.

□

Being able to extend classical methods to fuzzy ones is already a very powerful tool. But what would happen if what we have at hand consists in a finite subset of $[0,1]$, or even if the truth-values at hand are not linearly ordered? These questions are dealt with in the next section.

2.5 Residuated Lattices

Definition 2.5.1. Let U be a (classical) set, and let $\langle L, \leqslant\rangle$ be a complete lattice, where $L \neq \varnothing$. A *L-fuzzy subset*[20] F of U is a function

$$\mu_F : U \to L \quad .$$

The set of all L-fuzzy subsets of U is denoted by L^U.

Notice that, we have special cases of L-fuzzy sets when $L = \{0,1\}$ (in which case we have characteristic functions of classical subsets of U) and when $L = [0,1]$ (which, under the usual order, corresponds to standard fuzzy sets).

Now recall from Definitions 2.3.1, 2.3.5 and 2.3.12 that t-norms, t-conorms and fuzzy implications respectively were defined solely in terms the (standard) order of $[0,1]$ and of equality. Thus, each of these definitions can be generalized here.

Definition 2.5.2. Let L be a complete lattice. Then a *triangular norm* and a *triangular conorm* on L are maps $\vartriangle, \triangledown : L^2 \to L$ respectively, which satisfy the following for all

[20] The definition of L-fuzzy sets was first introduced in 1967 by Goguen in (GOGUEN, 1967), as a generalization of the idea of fuzzy sets introduced two years earlier by Zadeh in (ZADEH, 1965).

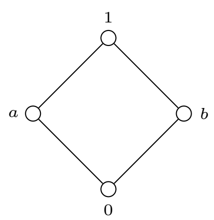
$x, y, z \in L$ the following:

1. $x \vartriangle y = y \vartriangle x$,
2. $(x \vartriangle y) \vartriangle z = x \vartriangle (y \vartriangle z)$,
3. If $y \leqslant z$ then $x \vartriangle y \leqslant x \vartriangle z$,
4. $x \vartriangle 1 = x$,

1'. $x \triangledown y = y \triangledown x$,
2'. $(x \triangledown y) \triangledown z = x \triangledown (y \vartriangle z)$,
3'. If $y \leqslant z$ then $x \triangledown y \leqslant x \triangledown z$,
4'. $x \triangledown 0 = x$.

Example 2.5.3. Clearly if $L = [0, 1]$ then t-norms and t-conorms have the meaning that has been presented earlier. Let us now consider some other examples.

1. For each $n > 0$, let
$$L_{n+1} = \{k/n : k = 0, ..., n\} \ .$$
The respective restrictions $\vartriangle_{L_{n+1}}, \triangledown_{L_{n+1}}$ of a t-norm and a t-conorm $\vartriangle, \triangledown : [0, 1]^2 \to [0, 1]$ to L_{n+1} are a t-norm and a t-conorm on L_{n+1} respectively. In fact, for each $x, y \in L_{n+1}$,
$$x \vartriangle_{L_{n+1}} y = x \vartriangle y \ , \qquad x \triangledown_{L_{n+1}} y = x \triangledown y \ .$$
2. If L is an arbitrary finite linearly ordered set with $n+1$ elements, then it is isomorphic to L_{n+1} in the sense that there exists a bijection $f : L \to L_{n+1}$ such that, for all $x, y \in L$,
$$x \leqslant_L y \text{ iff } f(x) \leqslant_{L_{n+1}} f(y) \ .$$
Thus, each t-norm and t-conorm $\vartriangle, \triangledown$ on L_{n+1} induces a t-norm and a t-conorm $\vartriangle_f, \triangledown_f$ on L, defined respectively, for all $x, y \in L$, by
$$x \vartriangle_f y = f(x) \vartriangle f(y) \ , \qquad x \triangledown_f y = f(x) \triangledown f(y) \ .$$
3. Let L be a (complete) lattice. Then the meet and join[21] operators on L are respectively a t-norm and a t-conorm on L. For instance, the (finite, hence) complete lattice defined by



[21] By *meet* and *join* of $X = \{x, y\} \subseteq L$ we mean respectively the greatest lower bound and the least upper bound of X on L. The meet and join binary operators are respectively the maps $L^2 \to L$ that map every pair (x, y) to their meet and join.

has respective meet and join given by[22]

$\wedge$	0	a	b	1
0	0	0	0	0
a	0	a	0	a
b	0	0	b	b
1	0	a	b	1

$\vee$	0	a	b	1
0	0	a	b	1
a	a	a	1	1
b	b	1	b	1
1	1	1	1	1

We could now define L-fuzzy implications by generalizing the definition of a fuzzy implication, but for our purposes it is sufficient to generalize the idea of residuum of a t-norm. Recall that Corollary 2.4.12 stated that, given a left semicontinuous t-norm $\triangle$ and its residuum they satisfy, for all $x, y, z \in [0, 1]$,

$$z \leqslant (x \Rightarrow y) \text{ iff } x \triangle z \leqslant y \ . \tag{2.5.4}$$

That's the property which is generalized.

Definition 2.5.5. A *residuated lattice* is a tuple $\langle L, \vee, \wedge, *, \Rightarrow, 0, 1\rangle$ such that:

1. $\langle L, \vee, \wedge, 0, 1\rangle$ is a lattice[23] with least element 0 and greatest element 1;
2. $*$ is a binary operation on L that is associative, commutative and such that $1 * x = x$ for all $x \in L$;
3. $*$ and $\Rightarrow$ form an adjoint pair, i. e., for all $x, y, z \in L$ we have

$$z \leqslant (x \Rightarrow y) \text{ iff } x * z \leqslant y \ . \tag{2.5.6}$$

A residuated lattice is called *complete* iff the lattice in item 1. is complete.

Notice that a t-norm always satisfies item 2., with the additional property that t-norms are monotonic increasing maps.

Example 2.5.7. It follows from (2.5.4) that if $\triangle$ is a lower semicontinuous t-norm on $L = [0, 1]$ and $\Rightarrow$ is its residuum, then $\langle L, \sup, \inf, \triangle, \Rightarrow, 0, 1\rangle$ is a residuated lattice. Now, consider the following.

1. Let L be a finite linearly ordered set with smallest and greatest elements $0, 1$ respectively, and let $\triangle$ be a t-norm on it. Define $\Rightarrow$ for all $x, y \in L$ as follows.

$$(x \Rightarrow y) = \max\{z \in L : x \triangle z \leqslant y\} \ .$$

[22] For any finite (and sufficiently small) set S, it is usual to present a binary operator $S^2 \to S$ as a table in which rows and columns are "indexed" by the elements of S and, if $(x, y) \mapsto z$, then the element of row x and column y is z.

[23] Recall that (0.2.3, 0.2.4) uniquely define the order relation in terms of the join and meet binary operators.

Then $\langle L, \vee, \wedge, \triangle, \Rightarrow, 0, 1\rangle$ is a residuated lattice, where $\vee$, $\wedge$ are respectively the join and meet operators on L.

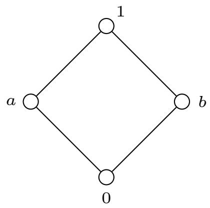
It is sufficient to prove that (2.5.6) holds. Let $x, y, z \in L$. Suppose that $z \leqslant (x \Rightarrow y)$. Since L is finite and linearly ordered, we have

$$z \in \{z_0 \in L : x \triangle z_0 \leqslant y\} \quad ,$$

whence $x \triangle z \leqslant y$. Conversely, if $x \triangle z \leqslant y$, then

$$z \leqslant \max\{z_0 \in L : x \triangle z_0 \leqslant y\} = (x \Rightarrow y) \quad .$$

2. The following lattice, t-norm and residuum constitute a residuated lattice.



$\wedge$	0	a	b	1
0	0	0	0	0
a	0	a	0	a
b	0	0	b	b
1	0	a	b	1

$\Rightarrow$	0	a	b	1
0	1	1	1	1
a	b	1	b	1
b	a	a	1	1
1	0	a	b	1

Again, we only need to prove that (2.5.6) holds. If $z = 0$ then both $z \leqslant (x \Rightarrow y)$ and $0 = x \wedge z \leqslant y$ are true for all $x, y, \in L$. If $z = 1$, then (2.5.6) reduces to $(x \Rightarrow y) = 1$ iff $x \leqslant y$, which can be seen by inspection.

Let us consider $0 < z < 1$. Then either $z = a$ or $z = b$. Suppose WLOG that $z = a$ — the proof for $z = b$ is analogous. Then both

$$x \wedge a \leqslant y \quad , \qquad\qquad a \leqslant (x \Rightarrow y)$$

are true iff at least one of the following holds.

$$x = 0 \quad , \qquad x = b \quad , \qquad y = a \quad , \qquad y = 1 \quad .$$

Therefore, (2.5.6) is always true.

We now have developed the tools necessary to extend/generalize any expression of classical (first order) logic. We are finally in position to go back to FCA and generalize it.

3 Fuzzy Formal Concept Analysis

We now introduce the reader to some developments on fuzzy extensions of FCA, here referred to as *Fuzzy Formal Concept Analysis*. In Sec. 3.1 we shall presented the ideas and results by using standard fuzzy sets (i. e., with $L = \{0, 1\}$). The contents in this section have been developed independently by the authors of this text[1] and later found to fall within the scope of a broader theory, which we shall briefly discuss in Sec. 3.2.

3.1 Formal concept analysis in the fuzzy setting

Recall that the theory of FCA as presented in Ch. 1 is based on a given formal context which consists of sets of objects and attributes, and a binary relation on them. Because a fuzzy set is defined on a classical universal set U, we continue to consider the universal sets of objects and attributes as classical sets, but we are now allowed to work with fuzzy subsets of objects and attributes[2]. Furthermore, we consider the binary relation to be fuzzy.

Definition 3.1.1. A *fuzzy formal context* is an ordered triple $\mathbb{C}_{\mathrm{f}} = \langle \mathcal{O}, \mathcal{A}, I_{\mathrm{f}} \rangle$, in which $\mathcal{O}$ and $\mathcal{A}$ are non-empty (classical) sets, and $I_{\mathrm{f}} \subseteq \mathcal{O} \times \mathcal{A}$ is a fuzzy binary relation.

The fact that we have a fuzzy relation enables us to express to which degree the object o has the attribute a, rather than the classical "either oIa or $(o, a) \notin I$".

Recall that for the crisp relation I we defined in (1.1.4) and (1.1.5) the maps $^* : 2^{\mathcal{O}} \to 2^{\mathcal{A}}$ and $^\wedge : 2^{\mathcal{A}} \to 2^{\mathcal{O}}$ by

$$O^* := \{a \in \mathcal{A} : oIa \text{ for all } o \in O\}$$
$$A^\wedge := \{o \in \mathcal{O} : oIa \text{ for all } a \in A\} \; .$$

Consider for instance the first of these equations. In order to decide whether a given $a \in \mathcal{A}$ is a member of O^*, it suffices to test whether oIa holds for each $o \in O$. We need not concern ourselves with $o \notin O$. That is exactly the mathematical practice regarding the (classical) implication. When we write down $\Phi \longrightarrow \Psi$, we mean that whenever Φ is

[1] See (BRITO et al., 2018) .

[2] When applying the theory there are several situations in which it makes no sense to work with fuzzy subsets of objects, but there are other situations in which it does make sense. For instance, a patient is not $2/3$ of a person, but a "partial" symptom (noise when breathing to a degree $2/3$) may be related to a "partial" disease (a not fully developed pneumonia, with degree $3/4$).

true, so must be Ψ, but in case Φ is false we have no interest in Ψ[3]. Hence we may say that $a \in O^*$ iff $o \in O \longrightarrow oIa$.

By using an analogous argument for $A^\wedge$ we see that the expressions for O^* and $A^\wedge$ can be translated respectively as

$$\chi_{O^*}(\tilde{a}) = \forall o \in \mathcal{O}(o \in O \longrightarrow oI\tilde{a}) \ ,$$
$$\chi_{A^\wedge}(\tilde{o}) = \forall a \in \mathcal{A}(a \in A \longrightarrow \tilde{o}Ia) \ .$$

Now recall that in Section 2.3.2 we argued that the universal quantifier can be extended by taking the infimum. As a result, we may define the following.

Definition 3.1.2. Let $\mathbb{C}_f = \langle \mathcal{O}, \mathcal{A}, I_f \rangle$ be a fuzzy formal context and let $\Rightarrow$ be a fuzzy implication. Then given fuzzy subsets $O \subseteq \mathcal{O}$ and $A \subseteq \mathcal{A}$ we define fuzzy subsets $O^* \subseteq \mathcal{A}$ and $A^\wedge \subseteq \mathcal{O}$ respectively by their membership functions:

$$\mu_{O^*}(a) = \inf_{o \in \mathcal{O}} [\mu_O(o) \Rightarrow \mu_{I_f}(o, a)] \ , \tag{3.1.3}$$
$$\mu_{A^\wedge}(o) = \inf_{a \in \mathcal{A}} [\mu_A(a) \Rightarrow \mu_{I_f}(o, a)] \ . \tag{3.1.4}$$

A *fuzzy formal concept* is a pair $\langle O, A \rangle \in \mathcal{O} \times \mathcal{A}$ such that $O^* = A$ and $A^\wedge = O$.

In order to extend results achieved working out the classical theory, we need to find out what fuzzy connectives preserve the classical properties we want to keep. For instance, remember that item 2. of Theorem 1.1.10 stated that given a formal context $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ and given $O \subseteq \mathcal{O}$, we have

$$O \subseteq O^{*\wedge} \ .$$

Considering how the fuzzy sets O^* and $A^\wedge$ are defined, we need to do some investigation concerning fuzzy implications. We start by asking ourselves: does continuity preserve the properties we want?

Example 3.1.5. Let $\mathbb{C}_f = \langle \mathcal{O}, \mathcal{A}, I_f \rangle$ be the fuzzy context defined so that $\mathcal{O} = \{o_1, o_2\}$, $\mathcal{A} = \{a_1, a_2\}$ and I_f is expressed by

	a_1	a_2
o_1	0.2	0.7
o_2	0.4	0.6

[3] We say that the classical implication is *vacuously true*, that is, an expression $\Phi \longrightarrow \Psi$ is true whenever Φ is false. It is also true, of course, when both Φ and Ψ are true, and false only in case Φ is true and Ψ is false.

Let $(x \Rightarrow y) := (1 - x) \vee y$ for each $x, y \in [0,1]$ (this is the Kleene-Dienes implication). Consider the fuzzy set $O = \dfrac{0.5}{o_1} + \dfrac{0.8}{o_2}$. Then

$$\begin{aligned}\mu_{O^*}(a_1) &= \inf_{o \in \mathcal{O}} [\mu_O(o) \Rightarrow \mu_{I_\mathrm{f}}(o, a_1)] \\ &= [(1-0.5) \vee 0.2] \bigwedge [(1-0.8) \vee 0.4] \\ &= 0.5 \wedge 0.4 = 0.4\end{aligned}$$

and

$$\begin{aligned}\mu_{O^*}(a_2) &= \inf_{o \in \mathcal{O}} [\mu_O(o) \Rightarrow \mu_{I_\mathrm{f}}(o, a_2)] \\ &= [(1-0.5) \vee 0.7] \bigwedge [(1-0.8) \vee 0.6] \\ &= 0.7 \wedge 0.6 = 0.6 \ ,\end{aligned}$$

so that $O^* = \dfrac{0.4}{a_1} + \dfrac{0.6}{a_2}$. Now, we have

$$\begin{aligned}\mu_{O^{*\wedge}}(o_2) &= \inf_{a \in \mathcal{A}} [\mu_{O^*}(a) \Rightarrow \mu_{I_\mathrm{f}}(o_2, a)] \\ &= [(1-0.4) \vee 0.4] \bigwedge [(1-0.6) \vee 0.6] \\ &= 0.6 \wedge 0.6 = 0.6 \ .\end{aligned}$$

Since $\mu_{O^{*\wedge}}(o_2) = 0.6 < 0.8 = \mu_O(o_2)$, we conclude that $O \nsubseteq O^{*\wedge}$.

The implication used in this example is continuous as it is the composition of continuous functions (maximum and difference). Thus continuity is not a sufficient condition on a fuzzy implication for us to extend Theorem 1.1.10. As we shall see, continuity is not a necessary condition as well. Since item 4. of the aforementioned theorem uses a Cartesian product we are led to think that t-norms may be involved and that we must use their residua.

However, not any residuum solves the problem we are currently tackling. For example, remember that the item 4. of Theorem 1.1.10 states that if $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I \rangle$ is a formal context, then given $O \subseteq \mathcal{O}$ and $A \subseteq \mathcal{A}$ have

$$A \subseteq O^* \text{ iff } O \times A \subseteq I \ .$$

Example 3.1.6. Let $\mathbb{C}_\mathrm{f} = \langle \{o\}, \{a\}, I_\mathrm{f} \rangle$ be a fuzzy formal context in which $\mu_{I_\mathrm{f}}(o, a) = 0.4$. Let $O = {}^{0.5}/_{o}$ and let $A = {}^{1}/_{a}$. Let $\triangle$ be the drastic product t-norm[4] and $\Rightarrow$ the residuum

[4] As introduced in Example 2.3.2,

$$x \triangle_D y = \begin{cases} 0, & \text{if } (x \vee y) < 1 \\ x, & \text{if } y = 1 \\ y, & \text{if } x = 1 \ . \end{cases}$$

of $\vartriangle$. Then

$$\begin{aligned}(\mu_O(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) &= (0.5 \Rightarrow 0.4)\\ &= \sup\{z \in [0,1] : 0.5 \vartriangle z \leqslant 0.4\}\\ &= 1\end{aligned}$$

as $0.5 \vartriangle z = 0$ for all $z \in [0,1[$. Thus, $O^* = {}^1\!/_a$, and we have $A \subseteq O^*$. Nevertheless,

$$\begin{aligned}\mu_O(o) \vartriangle \mu_A(a) &= 0.5 \vartriangle 1\\ &= 0.5\\ &> 0.4 = \mu_{I_{\mathrm{f}}}(o,a) \quad .\end{aligned}$$

Therefore, $O \times_\vartriangle A \nsubseteq I_{\mathrm{f}}$.

We see that $O \times_\vartriangle A \nsubseteq I_{\mathrm{f}}$ results from the discontinuity of $\vartriangle$. Nonetheless, it turns out that, for many of the results we wish to prove, something weaker than continuity of the t-norm is required: all we need are lower semicontinuous t-norms[5].

Theorem 3.1.7. *Let $\mathbb{C} = \langle \mathcal{O}, \mathcal{A}, I_{\mathrm{f}} \rangle$ be a fuzzy context. Let O, O_1, O_2 be fuzzy subsets of $\mathcal{O}$ and A, A_1, A_2 be fuzzy subsets of $\mathcal{A}$. Let $\vartriangle$ be a lower semicontinuous t-norm, and let $\Rightarrow$ be the residuum of $\vartriangle$. Then:*

1. If $O_1 \subseteq O_2$ then $O_2^* \subseteq O_1^*$	*1'. If $A_1 \subseteq A_2$ then $A_2^\wedge \subseteq A_1^\wedge$*
2. $O \subseteq O^{*\wedge}$	*2'. $A \subseteq A^{\wedge *}$*
3. $O^* = O^{*\wedge *}$	*3'. $A^\wedge = A^{\wedge * \wedge}$*

4. $O \subseteq A^\wedge$ iff $A \subseteq O^*$ iff $O \times_\vartriangle A \subseteq I_{\mathrm{f}}$

Proof. We shall prove items 1., 2., 3. and 4. Items with a prime can be proved analogously.

1. Suppose that $O_1 \subseteq O_2$. Let $a \in \mathcal{A}$. By hypothesis, $\mu_{O_1}(o) \leqslant \mu_{O_2}(o)$ for each $o \in \mathcal{O}$. Since $\Rightarrow$ is decreasing in its first component we have

$$(\mu_{O_2}(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) \leqslant (\mu_{O_1}(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) \quad .$$

   Taking the infimum over o on both sides, we have (by definition of O_1^* and O_2^*) $\mu_{O_2^*}(a) \leqslant \mu_{O_1^*}(a)$. But $a \in \mathcal{A}$ is arbitrary. Thus $O_2^* \subseteq O_1^*$.

2. Let $o \in \mathcal{O}$. Remember that $\Rightarrow$ is decreasing in its first component. Thus

$$\begin{aligned}\mu_O(o) &\leqslant [(\mu_O(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)] && \text{Lemma 2.4.14}\\ &\leqslant \left[\inf_{\tilde{o}\in\mathcal{O}} (\mu_O(\tilde{o}) \Rightarrow \mu_{I_{\mathrm{f}}}(\tilde{o},a)) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] && \text{Monotonicity of } \Rightarrow\\ &= [\mu_{O^*}(a) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)] \quad .\end{aligned}$$

[5] Lower semicontinuous functions were introduced in Def. 2.4.8.

Taking the infimum over a on the right side, we get $\mu_O(o) \leqslant \mu_{O^{*\wedge}}(o)$. Since $o \in \mathcal{O}$ is arbitrary, $O \subseteq O^{*\wedge}$.

3. From item 2., we have $O \subseteq O^{*\wedge}$. Thus, using 1., $O^{*\wedge *} \subseteq O^*$. On the other hand, using $A = O^*$ in 2'., we have $O^* \subseteq O^{*\wedge *}$. Hence, $O^* = O^{*\wedge *}$.

4. We first prove that $O \subseteq A^\wedge$ iff $A \subseteq O^*$. Suppose that $O \subseteq A^\wedge$. By 1., $A^{\wedge *} \subseteq O^*$. Using 2'. and transitivity of $\subseteq$, we have $A \subseteq O^*$. The proof that $A \subseteq O^*$ implies $O \subseteq A^\wedge$ is analogous.
Now we prove that $O \subseteq A^\wedge$ iff $O \times_\triangle A \subseteq I_\mathrm{f}$. In fact, $O \subseteq A^\wedge$ iff for all $o \in \mathcal{O}$ and for all $a \in \mathcal{A}$

$$\begin{aligned}\mu_O(o) &\leqslant \mu_{A^\wedge}(o)\\ &= \inf_{\tilde{a}\in\mathcal{A}} [\mu_A(\tilde{a}) \Rightarrow \mu_{I_\mathrm{f}}(o,\tilde{a}))]\\ &\leqslant [\mu_A(a) \Rightarrow \mu_{I_\mathrm{f}}(o,a)] \quad ,\end{aligned}$$

By Corollary 2.4.12 (and applying commutativity of $\triangle$) this holds iff

$$\mu_O(o) \triangle \mu_A(a) \leqslant \mu_{I_\mathrm{f}}(o,a) \quad .$$

Thus $O \subseteq A^\wedge$ iff $O \times_\triangle A \subseteq I_\mathrm{f}$.

□

Notice that, as in the classical case, $({}^*, {}^\wedge)$ is a Galois connection between $\langle \mathcal{F}(\mathcal{O}), \subseteq \rangle$ and $\langle \mathcal{F}(\mathcal{A}), \subseteq \rangle$, so that ${}^{*\wedge}$ and ${}^{\wedge *}$ are closure operators.

Another interesting matter is that not only the set $\mathfrak{B}_\mathrm{f}(\mathbb{C}_\mathrm{f})$ of fuzzy concepts on $\mathbb{C}_\mathrm{f}$ is a complete lattice (under an order similar to that of classical concepts): the formulae for finding infima and suprema on this lattice are similar to those of the classical case.

Proposition 3.1.8. *Let $\mathbb{C}_\mathrm{f} = \langle \mathcal{O}, \mathcal{A}, I_\mathrm{f} \rangle$ be a fuzzy context. Let $\triangle$ be a lower semicontinuous t-norm, and let $\Rightarrow$ be its residuum. Let J be an index set and, for each $\alpha \in J$, let $O_\alpha \subseteq \mathcal{O}$ and $A_\alpha \subseteq \mathcal{A}$. Then*

$$\bigcap_{\alpha\in J} O_\alpha^* = \left(\bigcup_{\alpha\in J} O_\alpha\right)^* \quad , \tag{3.1.9}$$

$$\bigcap_{\alpha\in J} A_\alpha^\wedge = \left(\bigcup_{\alpha\in J} A_\alpha\right)^\wedge \quad . \tag{3.1.10}$$

Proof. We prove (3.1.9). The proof of (3.1.10) is analogous. Let $a \in \mathcal{A}$. For each $\alpha_0 \in J$ we have

$$\begin{aligned}\mu_{(\cup_{\alpha\in J} O_\alpha)^*}(a) &= \inf_{o\in\mathcal{O}} \left[\mu_{\cup_{\alpha\in J} O_\alpha}(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] \\ &= \inf_{o\in\mathcal{O}} \left[\left(\sup_{\alpha\in J} \mu_{O_\alpha}(o)\right) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] \\ &\leqslant \left[\left(\sup_{\alpha\in J} \mu_{O_\alpha}(o)\right) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] \\ &\leqslant \left[\mu_{O_{\alpha_0}}(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] \quad ,\end{aligned}$$

as $\Rightarrow$ is decreasing in the first component. Applying the infimum over $o \in \mathcal{O}$ and then the infimum over $\alpha \in J$ on the right-hand side yields

$$\mu_{(\cup_{\alpha\in J} O_\alpha)^*}(a) \leqslant \mu_{\cap_{\alpha\in J} O^*_\alpha}(a) \quad .$$

Since $a \in \mathcal{A}$ is arbitrary,

$$\left(\bigcup_{\alpha\in J} O_\alpha\right)^* \subseteq \bigcap_{\alpha\in J} O^*_\alpha \quad .$$

Conversely, we want to prove

$$\bigcap_{\alpha\in J} O^*_\alpha \subseteq \left(\bigcup_{\alpha\in J} O_\alpha\right)^* \quad . \tag{3.1.11}$$

Notice that if for all $o \in \mathcal{O}$ and for all $a \in \mathcal{A}$ we have

$$\mu_{\cap_{\alpha\in J} O^*_\alpha}(a) \leqslant \left[\left(\mu_{\bigcup_{\alpha\in J} O_\alpha}(o)\right) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] \quad , \tag{3.1.12}$$

then taking the infimum over $o \in \mathcal{O}$ on the right-hand side of (3.1.12) yields (3.1.11). Hence it is sufficient to prove (3.1.12).

Suppose for the sake of contradiction that (3.1.12) does not hold, that is, for some $o \in \mathcal{O}$ and $a \in \mathcal{A}$,

$$\kappa := \inf_{\alpha\in J} \left[\inf_{\tilde{o}\in\mathcal{O}} \left(\mu_{O_\alpha}(\tilde{o}) \Rightarrow \mu_{I_{\mathrm{f}}}(\tilde{o},a)\right)\right] > \left[\left(\sup_{\alpha\in J} \mu_{O_\alpha}(o)\right) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right] \quad .$$

Define $x_0 = \sup_{\alpha\in J} \mu_{O_\alpha}(o)$. Let (α_n) be a sequence on J such that $\left(\mu_{O_{\alpha_n}}(o)\right)$ converges to x_0 and, for each $n > 0$, define $x_n = \mu_{O_{\alpha_n}}(o)$. By hypothesis, for each $n > 0$,

$$\begin{aligned}(x_0 \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) &< \inf_{\alpha\in J} \left[\inf_{\tilde{o}\in\mathcal{O}} \left(\mu_{O_\alpha}(\tilde{o}) \Rightarrow \mu_{I_{\mathrm{f}}}(\tilde{o},a)\right)\right] \qquad (= \kappa) \\ &\leqslant \inf_{\tilde{o}\in\mathcal{O}} \left(\mu_{O_{\alpha_n}}(\tilde{o}) \Rightarrow \mu_{I_{\mathrm{f}}}(\tilde{o},a)\right) \\ &\leqslant \left(\mu_{O_{\alpha_n}}(o) \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)\right) \\ &= (x_n \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) \quad .\end{aligned}$$

Thus,

$$(x_0 \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) < \kappa \leqslant \limsup_{n\to\infty} (x_n \Rightarrow \mu_{I_{\mathrm{f}}}(o,a)) \quad ,$$

and so $x \mapsto (x \Rightarrow \mu_{I_f}(o, a))$ is not upper semicontinuous, contradicting Proposition 2.4.17.

Hence, (3.1.12) holds.

□

Theorem 3.1.13. *Let $\mathbb{C}_f = \langle \mathcal{O}, \mathcal{A}, I_f \rangle$ be a fuzzy formal context. Using a lower semicontinuous t-norm and its residuum to define the maps * and $^\wedge$, define the order $\leqslant$ on the set $\mathfrak{B}_f(\mathbb{C}_f)$ of all the fuzzy formal concepts of $\mathbb{C}_f$ by*

$$\langle O_1, A_1 \rangle \leqslant \langle O_2, A_2 \rangle \text{ iff } O_1 \subseteq O_2 \qquad (\text{iff } A_2 \subseteq A_1) \ . \tag{3.1.14}$$

Then $\mathcal{L}_{\mathbb{C}_f} := \langle \mathfrak{B}_f(\mathbb{C}_f), \leqslant \rangle$ is a complete lattice, called the fuzzy concept lattice *of $\mathbb{C}_f$.*

If K is an index set and $C_\kappa = \langle O_\kappa, A_\kappa \rangle \in \mathfrak{B}(\mathbb{C})$ for each $\kappa \in K$ then

$$\inf_{\kappa \in K} C_\kappa = \left\langle \bigcap_{\kappa \in K} O_\kappa, \left(\bigcup_{\kappa \in K} A_\kappa \right)^{\wedge *} \right\rangle \ , \tag{3.1.15}$$

$$\sup_{\kappa \in K} C_\kappa = \left\langle \left(\bigcup_{\kappa \in K} O_\kappa \right)^{* \wedge}, \bigcap_{\kappa \in K} A_\kappa \right\rangle \ . \tag{3.1.16}$$

Proof. Similar to that of the Basic Theorem (Theorem 1.1.18). □

3.2 Formal concept analysis on complete residuated lattices

As we have already seen it is possible to extend classical FCA by replacing classical sets with fuzzy sets. Moreover, in Sect. 2.5 we have presented a generalized version of fuzzy sets, called L-fuzzy sets, where the unit interval is replaced by a complete lattice L.

The present section follows BĚLOHLÁVEK; VYCHODIL's "What is a fuzzy concept lattice?" in which some distinct approaches to L-fuzzy concept analysis are presented (2005).

3.2.1 The approach of Burusco and Fuentes-Gonzáles

The field of Fuzzy FCA was launched by (FUENTES-GONZÁLEZ; BURUSCO, 1994). Their approach was based on an algebraic system $\langle L, \leqslant, ', \triangledown, 0, 1 \rangle$ in which

1. $\langle L, \leqslant, 0, 1 \rangle$ is a complete lattice;
2. $\triangledown$ is a t-conorm[6] on L; and

[6] According to Def. 2.5.2.

3. $'$ is a *complementation*[7] on L.

Then, given classical sets X, Y and a L-fuzzy relation $I_L \subseteq L^{X\times Y}$[8], they define maps $* : L^X \to L^Y$ and $^\wedge : L^Y \to L^X$ by

$$A^*(y) = \bigwedge_{x\in X} (A'(x) \triangledown I(x,y)) \quad ,$$
$$B^\wedge(x) = \bigwedge_{y\in Y} (B'(y) \triangledown I(x,y)) \quad .$$

This approach leads to a complete L-fuzzy concept lattice. Nonetheless, it does not present some useful properties, namely, $^{\wedge *}$ and $^{*\wedge}$ are not closure operators[9].

3.2.2 The approach of Pollandt and Bělohlávek

In the late 1990's (POLLANDT, 1997) and (BĚLOHLÁVEK, 1999) independently developed an approach which could[10] be stated as a straightforward generalization of the approach presented in Sec. 3.1 to complete residuated lattices.

In fact, let $\langle L, \vee, \wedge, *, \Rightarrow, 0, 1\rangle$ be a complete residuated lattice[11]. An *L-context* is a triple $\langle X, Y, I\rangle$ in which $I \in L^{X\times Y}$ is an L-fuzzy binary relation.

The maps $^* : L^X \to L^Y$, $^\wedge : L^Y \Rightarrow L^X$ are defined for each $A \in L^X$, $B \in L^Y$ by

$$A^*(y) = \bigwedge_{x\in X} (A(x) \Rightarrow I(x,y)) \quad ,$$
$$B^\wedge(x) = \bigwedge_{y\in Y} (B(y) \Rightarrow I(x,y)) \quad .$$

In working with this approach, $(^*, ^\wedge)$ constitutes a Galois connection[12], so that $^{*\wedge}$, $^{\wedge *}$ are closure operators.

---

7 FUENTES-GONZÁLEZ; BURUSCO's complementation, or *negation*, is introduced by (LÓPEZ; FUENTEZ-GONZÁLES, 1991) and is similar the notion of negation introduced in Def. 2.3.10. They define a negation as a map $' : L \to L$ that satisfies, for all $x, y \in L$:

a) $(x')' = x$;

b) $(x \vee y)' = x' \wedge y'$; ($\vee$ and $\wedge$ are the meet and join of L)

c) $0' = 1$.

8 So that, according to the approach presented thus far, we could say that $\langle X, Y, L^{X\times Y}\rangle$ is a L-fuzzy formal context.

9 Notice that, according to FUENTES-GONZÁLEZ; BURUSCO's approach, implications are generalized based on the form $\neg A \vee B$. As we have seen in Example 3.1.5, one such implication may not yield monotonic maps $^{\wedge *}$, $^{*\wedge}$, hence it they are not closure operators.

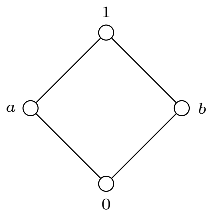
10 By reversing the thermodynamic arrow of time.

11 Recall that the notion of residuated lattices, presented in Sec. 2.5 (Def. 2.5.5), generalizes the notion o lower semicontinuous t-norms and their residua.

12 In an extended notion of Galois connections, according to (BĚLOHLÁVEK, 1999).

With this approach the Basic Theorem on Concept Lattices (Theorem 1.1.18) is given two different versions: one in which L-fuzzy concepts are ordered by L-fuzzy set inclusion[13] of their extents (BĚLOHLÁVEK, 1999) ; and another in which a so-called L-order[14] on L-fuzzy concepts is used (BĚLOHLÁVEK, 2004) .

Example 3.2.1. Recall that, according to Example 2.5.7, the following lattice diagram, t-norm and residuum constitute a residuated lattice.



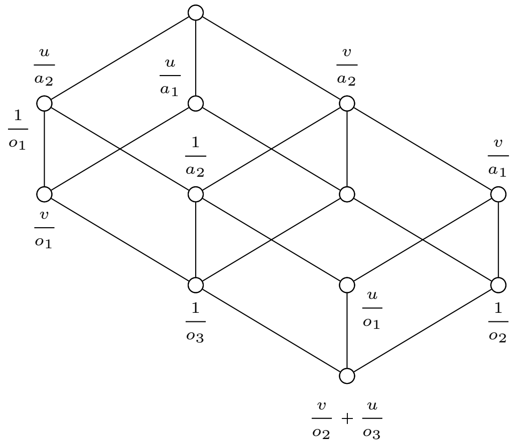
$\wedge$	0	a	b	1
0	0	0	0	0
a	0	a	0	a
b	0	0	b	b
1	0	a	b	1

$\Rightarrow$	0	a	b	1
0	1	1	1	1
a	b	1	b	1
b	a	a	1	1
1	0	a	b	1

Consider the following L-fuzzy context.

$\mathcal{O}$	$\mathcal{A}$	
	a_1	a_2
o_1	0	a
o_2	1	b
o_3	a	1

By evaluating every L-fuzzy formal concept and ordering the results, we found the following diagram.



[13] In the sense that $A \subseteq B$ in L^Y iff $A(y) \leqslant B(y)$ for all $y \in Y$.

[14] On a complete residuated lattice $\langle L, \vee, \wedge, *, \Rightarrow, 0, 1\rangle$, the notions of equality, order and subsetness are generalized to "L-notions", and after several other definitions and results, BĚLOHLÁVEK concludes that "$\langle\langle \mathfrak{B}(\mathbb{C}), \approx\rangle, \preceq\rangle$ is [a] completely lattice L-ordered set," with formulae for infimum, supremum, and conditions under which a "completely lattice L-ordered set" is isomorphic to this structure (2004, p. 11).

It is a bothersome, and even impracticable task to compute every single concept of a (L) fuzzy context. Clearly, computer algorithms must be used in this task, and that is precisely the purpose of the next and final chapter of this text.

4 Computing the Fuzzy Concept Lattice

4.1 Computational algorithms

As mentioned in Sect. 1.2, (LINDIG, 2000) introduces an algorithm for computing the concept lattice of classical FCA, which is based on the idea of finding upper neighbours of each given node of the lattice. In fact, the algorithm is not restricted to concept lattices, but it works for any finite lattice of closed sets with respect to a given closure operator[1]. Although we have not presented the classical algorithm in Sect. 1.2, we now present an extended version of that algorithm which, when restricted to truth-values 0 and 1, corresponds to the classical case. This extended algorithm was introduced by (BĚLOHLÁVEK et al., 2010).

For practical purposes, we consider a finite linearly ordered set $L = \{a_1, ..., a_n\}$, where $a_1 < ... < a_n$. If $i = 1, ..., n-1$ we write a_i^+ for a_{i+1}. Furthermore, we consider a nonempty finite set $Y = \{y_1, ..., y_m\}$ (which we may interpret as a set of attributes) and a closure operator $C : L^Y \to L^Y$, i. e., if A, B are L-fuzzy sets (members of L^Y) then

1. $A \subseteq C(A)$,
2. $C(A) \subseteq C(C(A))$,
3. $A \subseteq B$ implies $C(A) \subseteq C(B)$.

Here, as before, $A \subseteq B$ abbreviates $\forall y \in Y, A(y) \leqslant B(y)$, where $\leqslant$ is the order on L. Furthermore, because we shall work with exactly one closure operator C, we shall write $\overline{A}$ for $C(A)$.

We denote by $\mathscr{C}$ the collection of closed L-fuzzy sets of Y, i. e.,

$$\mathscr{C} = \{A \in L^Y : A = \overline{A}\} .$$

We know from Theorem 0.2.20 that $\langle \text{fix}(Y), \subseteq \rangle$ is a (complete) lattice. If $A, B \in \mathscr{C}$ and B is an upper neighbour[2] of A, we write

$$A \prec B .$$

If $A(y) < 1$ denote by $[y]_A$ the L-fuzzy set

$$[y]_A = \overline{\left[A \cup \left(\frac{A(y)^+}{y}\right)\right]} ,$$

[1] Recall basic properties of closure operator on Def. 0.2.17 and Theorem 0.2.20.

[2] That is, $A \subset B$ and there is no B' such that $A \subset B' \subseteq B$. For basic notions concerning upper neighbours and covering relations, see Def. 0.2.5 and Lem. 0.2.7.

that is, we take the L-fuzzy set A, increase the truth-value of y to the next possible value, and then take the closure of this set.

It is clear that $A \subset [y]_A$ (by monotonicity of the closure operator), but it may or may not happen that $A \prec [y]_A$.

Example 4.1.1. Consider the following fuzzy context:

	a_1	a_2	a_3
o_1	0.5	0	0
o_2	0.5	1	0
o_3	0.5	0	0.5

We now define $Y = \mathcal{A} = \{a_1, a_2, a_3\}$, $L = \{0, 0.5, 1\}$, define $^\wedge, ^*$ by means of 3-valued Łukasiewicz's implication and let the closure operator C be the composition $^{\wedge *}$. Let $A = {}^{0.5}/_{a_1} + {}^{0}/_{a_2} + {}^{0.5}/_{a_3}$. Then

$$[a_1]_A = \overline{\left(\frac{1}{a_1} + \frac{0}{a_2} + \frac{0.5}{a_3}\right)} = \frac{1}{a_1} + \frac{0.5}{a_2} + \frac{0.5}{a_3} \ ,$$

$$[a_2]_A = \overline{\left(\frac{0.5}{a_1} + \frac{0.5}{a_2} + \frac{0.5}{a_3}\right)} = \frac{1}{a_1} + \frac{0.5}{a_2} + \frac{0.5}{a_3} \ ,$$

$$[a_3]_A = \overline{\left(\frac{0.5}{a_1} + \frac{0}{a_2} + \frac{1}{a_3}\right)} = \frac{1}{a_1} + \frac{0.5}{a_2} + \frac{1}{a_3} \ ,$$

Notice that $A \subset [a_1]_A = [a_2]_A \subset [a_3]_A$, whence $[a_3]_A$ is not an upper neighbour of A.

Let $A \subset B$ and define

$$y_i = \max\{y_j \in Y : A(y_j) < B(y_j)\} \ , \tag{4.1.2}$$

and let $g_A(B) := y_i$. Belohlávek et. al. prove that if $A \prec B$ then $B = g_A(B)$, and $g_A(B)$ is called an *upper neighbour generator*, so that B is generated by y_i.

Thus far we know that an upper neighbour B of A can always be described in the form $[y]_A$, but a closed L-fuzzy set $[y]_A$ may not be an upper neighbour of A. Belohlávek et. al. then find the following criteria for verifying whether y is an upper neighbour generator.

Theorem 4.1.3. *Let*

$$\begin{aligned} \mathcal{M}_A &= \{y \in Y : A \prec [y]_A\} \\ &= \{y \in Y : A \prec B \text{ and } y = g_A(B)\} \end{aligned}$$

be the set of upper neighbour generators of A. Let $y_i \in Y$ be such that

$$y_i = g_A([y_i]_A) \ . \tag{4.1.4}$$

Then $y_i \in \mathcal{M}_A$ iff for each $y_k \in \mathcal{M}_A$ such that $k < i$ we have

$$([y_i]_A)\,(y_k) = A(y_k) \ . \tag{4.1.5}$$

Example 4.1.6. Let us consider Ex. 4.1.1 again. Recall that $A = {}^{0.5}/_{a_1} + {}^{0}/_{a_2} + {}^{0.5}/_{a_3}$. There, $[a_1]_A = [a_2]_A$, and so we have redundant information. Such a redundancy can be avoided when we consider (4.1.4). Since

$$[a_1]_A = \frac{1}{a_1} + \frac{0.5}{a_2} + \frac{0.5}{a_3} \ ,$$

we have

$$g_A([a_1]_A) = a_2 \ .$$

This, means that $a_1 \notin \mathcal{M}_A$, and we need not concern ourselves with a_1 when finding upper neighbours of A, because $[a_2]_A$ gives the same result.

What about a_3? We see that

$$[a_3]_A = \frac{1}{a_1} + \frac{0.5}{a_2} + \frac{1}{a_3} \ ,$$

As we have already seen, $A \subset [a_2]_A \subset [a_3]_A$, so that $A \nprec [a_3]_A$, and so $a_3 \notin \mathcal{M}_A$, even though $a_3 = g_A([a_3]_A)$, thus satisfying (4.1.4).

According to (4.1.5) there must exist $k < 3$ such that $y_k \in \mathcal{M}_A$ and

$$([a_3]_A)\,(y_k) > A(y_k) \ .$$

The only values $k < 3$ are $k = 1$ and $k = 2$, but we have already seen that $a_1 \notin \mathcal{M}_A$. Therefore $a_2 \in \mathcal{M}_A$, and we conclude that

$$\mathcal{M}_A = \{a_2\} \ ,$$

meaning that A has a unique upper neighbour, namely $[a_2]_A$.

Theorem 4.1.3 gives us all the information necessary for finding a minimal set $\mathcal{M}_A$ of upper neighbour generators of a given closed L-fuzzy set A and, from it, we can find the set $\mathcal{U}_A$ of upper neighbours of A as shown in Alg. 3.

Notice that if the conditional in row 7 of Alg. 3 returns *False*, then there exists y_k in the intersection and either:

1. $k < i$, so that (4.1.5) does not hold[3]; or

[3] If $k = 1$, condition (4.1.5) is vacuously satisfied. For $k = 2, ..., i-1$, y_k remains in *Min* iff $y_k \in \mathcal{M}_A$.

Algorithm 3 – UpperNeighbours

```
input  : Set A
output : The set U_A of upper neighbours of A

U_A ← ∅;
Min ← {y ∈ Y : A(y) < 1} ;                    /* Candidates to M_A */
for i = 1 : m do
    if A(y_i) < 1 then
        B ← [y_i]_A;
        Increased ← {y_k ∈ Y : k ≠ i and A(y_k) < B(y_k)};
        if Min ∩ Increased = ∅ then
            append B to U_A;
        else
            remove y from Min;
        end
    end
end
```

2. $k > i$, so that (4.1.4) does not hold.

Now, we should be able to determine a simple recursive procedure to build up the lattice as follows:

1. Compute $\overline{\varnothing}$ — the smallest closed set — and store it.
2. For a given closed set $A \neq Y$, find its set $\mathcal{U}_A$ of upper neighbours and, for each $B \in \mathcal{U}_A$:

   Store B in the set $\mathscr{C}$ of closed sets (if it has not been previously computed);

   Store A in the set $\mathcal{L}_B$ of lower neighbours of B.

This can be accomplished by combining Algorithm 3 with Algorithms 4 and 5. Variable $\mathcal{F}$ is shared by Algorithms 4 and 5, as well as $\mathcal{U}_A, \mathcal{L}_A$ for each $A \in \mathcal{F}$[4]. Furthermore, sets L, Y and the closure operator $C : L^Y \to L^Y$ are assumed to be known by all three algorithms[5].

Example 4.1.7. Let us consider once again Example 4.1.1. Recall that the fuzzy formal context was given by

[4] For implementation purposes, think of a $k \times 3$ matrix M, in which k increases as we compute new elements of $\mathcal{F}$. Then, in each row of M we have $\langle A, \mathcal{U}_A, \mathcal{L}_A, \rangle$, where $A \in \mathcal{F}$.

[5] The universal set Y must be known explicitly, whereas it is sufficient that L and C be programmed in a closure operator procedure.

Algorithm 4 – GenerateFrom

input : Set A
action : Evaluates sublattice of elements greater than A

```
if A ≠ Y then
    U_A ← UpperNeighbours(A);
    N ← U_A \ F;
    foreach B ∈ U_A do
        append A to L_B;
        if B ∈ N then
            append B to F;
        end
        foreach B ∈ N do
            GenerateFrom (B);
        end
    end
end
```

Algorithm 5 – Lattice

input :
output : $\{\langle A, \mathcal{U}_A, \mathcal{L}_A \rangle : A \in \mathcal{F}\}$

```
F ← ∅;
A ← ∅̄;
append A to F;
GenerateFrom(A);
```

	a_1	a_2	a_3
o_1	0.5	0	0
o_2	0.5	1	0
o_3	0.5	0	0.5

Table 3 – Revisiting example 4.1.1.

We consider $Y = \{a_1, a_2, a_3\}$ and $L = \{0, 0.5, 1\}$ with 3-valued Łukasiewicz implication in order to define the closure $^{\wedge *}$. Then, application of Alg. 5 to Table 3 returns the triples[6], and corresponding lattice diagram, presented in Fig. 11.

[6] Here the subscripts are indices representing the orders in which the nodes have been found. $A = (\alpha_1, \alpha_2, \alpha_3)$ expresses

$$A = \frac{\alpha_1}{a_1} + \frac{\alpha_2}{a_2} + \frac{\alpha_3}{a_3} \ .$$

A	$\mathcal{U}_A$	$\mathcal{L}_A$
$A_0 = (0.5,\ 0.0,\ 0.0)$	$\{A_1, A_2\}$	$\varnothing$
$A_1 = (0.5,\ 0.5,\ 0.0)$	$\{A_3, A_4\}$	$\{A_0\}$
$A_2 = (0.5,\ 0.0,\ 0.5)$	$\{A_4\}$	$\{A_0\}$
$A_3 = (1.0,\ 0.0,\ 0.0)$	$\{A_5\}$	$\{A_1\}$
$A_4 = (1.0,\ 0.5,\ 0.5)$	$\{A_5, A_7\}$	$\{A_1, A_2\}$
$A_5 = (1.0,\ 1.0,\ 0.5)$	$\{A_6\}$	$\{A_3, A_4\}$
$A_6 = (1.0,\ 1.0,\ 1.0)$	$\varnothing$	$\{A_5, A_7\}$
$A_7 = (1.0,\ 0.5,\ 1.0)$	$\{A_6\}$	$\{A_4\}$

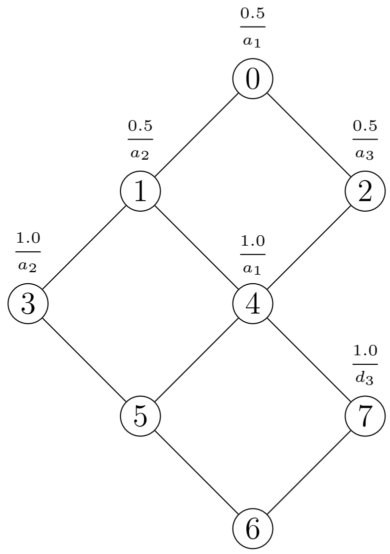


Figure 11 – Output (left) and Lattice Diagram (right) of Alg. 5 applied to Table 3.

Now, we explore an application of the theory we have presented thus far.

4.2 An application of fuzzy FCA

Let us now explore an application of Fuzzy FCA. In this section we shall consider an example of medical diagnosis of pneumonia in children, based on a work by PEREIRA et al. (2004), which comprises a study of the diagnosis of pneumonia in children by means of fuzzy relations. The approach used in deciding how to compute and interpret the fuzzy concept lattice is entirely ours.

4.2.1 Posing the problem

Pereira's study comprises 153 children who were randomly divided in two groups: an analysis sample (115 cases) and a validation sample (38 cases). These patients were independently diagnosed by two paediatricians according to the following possible diagnoses: *pneumonia*, *non-pneumonia diseases*, and *healthy*. Only cases of agreement were included in the group. The clinical signs observed were

1. *Radiological signs*, measured by alveolar and interstitial infiltrates, atelectasis, pleural effusion, pneumatoceles, airtrapping, pneumothorax;
2. *Dyspnea*, taking into account mild discomfort, lower rib in-drawing with tachypnea, intercostal in-drawing with severe tachypnea and/or presence of nasal flaring, full retraction of ribs plus cyanosis and/or poor peripheral blood perfusion;
3. *Auscultation signs*, increasing with rales, crackles, bronchial breathing;
4. *Heart rate*, ranging from normal to highly tachycardic;

	Clinical sign	Parameters
s_1	X-ray	0 – 7
s_2	Dyspnea	0 – 4
s_3	Auscultation	0 – 3
s_4	Heart rate	0 – 4
s_5	Temperature	0 – 3
s_6	Toxemia	0 – 4
s_7	Respiratory rate	0 – 4

Table 4 – Clinical signs observed for diagnosis of pneumonia

5. *Body temperature*, considered to be either normal ($T \leqslant 37^{o}$C), mild fever ($37^{o}\text{C} < T < 38.5^{o}$C), fever ($38.5^{o}\text{C} \leqslant T < 40^{o}$C) or severe fever ($40^{o}\text{C} < T$);
6. *Toxemia*, according to the presence of pallor, pallor and listlessness, irritability, drowsiness; and
7. *Respiratory rate*, according to age group, ranging from normal to highly tachypneic.

These signs, together with the range of parameters (measure scales) are presented in Table 4.

The scalings have been normalized so that if the original values of a clinical sign were natural numbers 0–n_s, after normalization we have

$$V_s = \{k/n_s : k = 0, \dots, n_s\} \quad .$$

Now, let X, Y, Z be respectively the sets of clinical signs, patients and diagnoses. The data collected from the patients consisted in a fuzzy binary relation S on $X \times Y$ and a crisp binary relation T on $Y \times Z$. Then, a fuzzy relation R on $X \times Z$ was obtained by means of the fuzzy relation composition[7]

$$R(x, z) = \bigvee_{y \in Y} [S(x, y) \wedge T(y, z)] \quad ,$$

which is expressed in Table 5.

[7] This composition of fuzzy binary relations extends the composition of classical binary relations

$$(T \circ S) = \{(x, z) \in X \times Z : \exists y \in Y \text{ such that } (x, y) \in S \text{ and } (y, z) \in T\} \quad .$$

When functions are defined as binary relations such that (x, y_1) and (x, y_2) imply $y_1 = y_2$, the composition of functions is their composition as binary relations, and $(x, y) \in f$ is expressed as $f(x) = y$.

Clinical sign	Diagnosis		
	Pneumonia	Non-pneumonia disease	Healthy
s_1: X-ray	$3/7$	0	0
s_2: Dyspnea	1	0.25	0
s_3: Auscultation	$2/3$	0	0
s_4: Heart rate	1	1	0.5
s_5: Temperature	$2/3$	1	0
s_6: Toxemia	0.75	0.5	0
s_7: Respiratory rate	1	0.75	0

Table 5 – Fuzzy relation of clinical signs and diagnoses

Table 5 can be considered as a fuzzy formal context, in which $\mathcal{O}$ is the set of clinical signs, $\mathcal{A}$ is the set of diagnoses and $I_{\mathrm{f}} = R$. Our goal is to build the fuzzy concept lattice by means of Alg. 5 — hence, of course, using Algorithms 3 and 4. Recall that Alg. 3 depends on L-fuzzy sets and we use the closure operator $^{\wedge *}$. Thus, we need to choose a set L (which by hipothesis of the algorithms used is finite and linearly ordered) and the residuum of a lower semicontinuous t-norm. We have chosen to work with two different residua: Gödel's implication and Łukasiwcicz's implication.

4.2.2 Computing the fuzzy concept lattice

First we consider Gödel's implication. Recall that

$$(x \Rightarrow_{\mathrm{g}} y) = \begin{cases} 1, & \text{if } x \leqslant y \\ y, & \text{if } x > y \end{cases} .$$

Because the only possible values this implication assumes are 1 and the value of the consequent (y), we may assume that L corresponds to the set of values on Table 5, as it already includes 1. This choice allows us to make the diagram on Fig. 12.

Now, let us consider Łukasiewicz implication

$$(x \Rightarrow_{\mathrm{l}} y) = (1 - x + y) \wedge 1 .$$

It is based on a sum and a difference, and so the safest manner to be assured that repeated applications of $\Rightarrow_{\mathrm{l}}$ will always result in a value that is a member of L is to choose a set L closed under additions and subtractions.

Because of the parameters on Table 4, which are later normalized, our first attempt was to make $L \subseteq [0,1]$ closed under additions and subtractions of $1/3$, $1/4$ and $1/7$. The least common multiple of 3, 4 and 7 is 84, so we chose

$$L = \{k/84 : k = 0, 2, ..., 84\} .$$

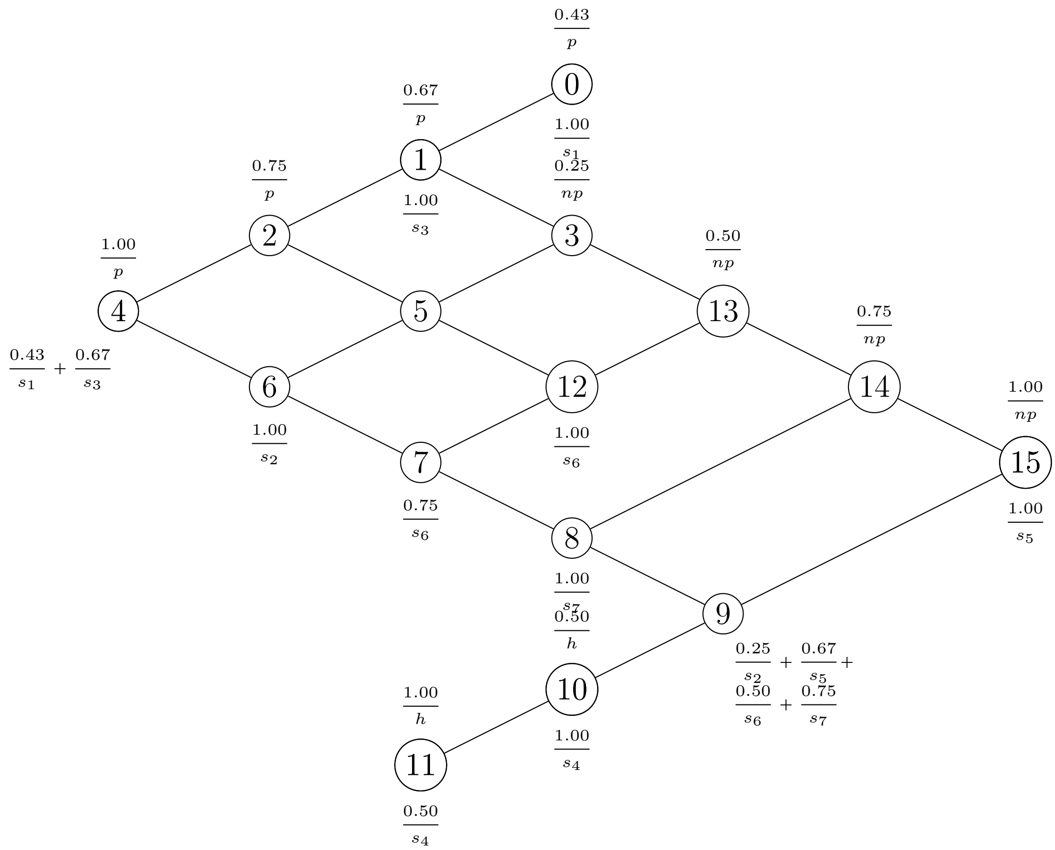


Figure 12 – Medical Diagnosis: Gödel's implication

This has been a misleading attempt, as it turns out the resulting fuzzy concept lattice has 55 160 nodes, which would be impracticable to read — not to mention the difficulty of even drawing such a diagram. Thus, some slightly different approach must be used.

A better attempt has been to choose a small natural number $n_0 > 0$, define

$$L_{n_0+1} = \{k/n_0 : k = 0, ..., n_0\} \ , \tag{4.2.1}$$

and then to redefine the formal context by taking each value v of it and mapping it to the smallest $l \in L_{n+1}$ such that $v < l$. For example, if we consider L_5 (so that $n_0 = 4$), we convert Table 5 into Table 6.

Clinical sign	Diagnosis		
	Pneumonia	Non-pneumonia disease	Healthy
s_1: X-ray	0.5	0	0
s_2: Dyspnea	1	0.25	0
s_3: Auscultation	0.75	0	0
s_4: Heart rate	1	1	0.5
s_5: Temperature	0.75	1	0
s_6: Toxemia	0.75	0.5	0
s_7: Respiratory rate	1	0.75	0

Table 6 – Relation of clinical signs and diagnoses, with values in L_5

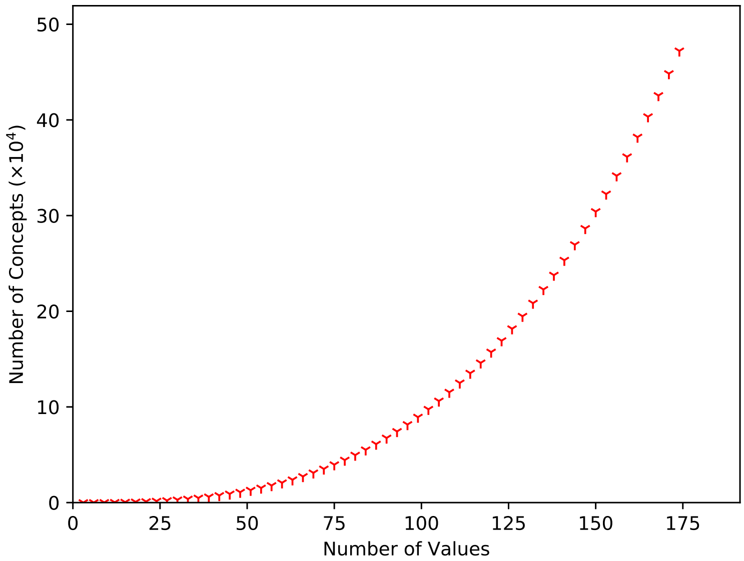


Figure 13 – Number of concepts with multivalued Łukasiewicz implications

When L_n is defined[8] as in (4.2.1) and $\Rightarrow_{\text{ł}}$ is defined on L_n, we say to be working with the n-valued Łukasiewicz implication.

We have numerically evaluated the number of concepts corresponding to Table 5 by using n-valued Łukasiewicz implications, $n = 3k$ for $k = 1, ..., 58$. The data is shown in Fig. 13. It turns out that the data can be interpolated[9] by a third degree polynomial[10]

$$y(n) = an^3 + bn^2 + cn \ ,$$

where y corresponds to the number of concepts and we have

$$a \approx 8.62 * 10^{-2} \ , \qquad b \approx 5.56 * 10^{-1} \ , \qquad c \approx 2.22 * 10^{-1} \ .$$

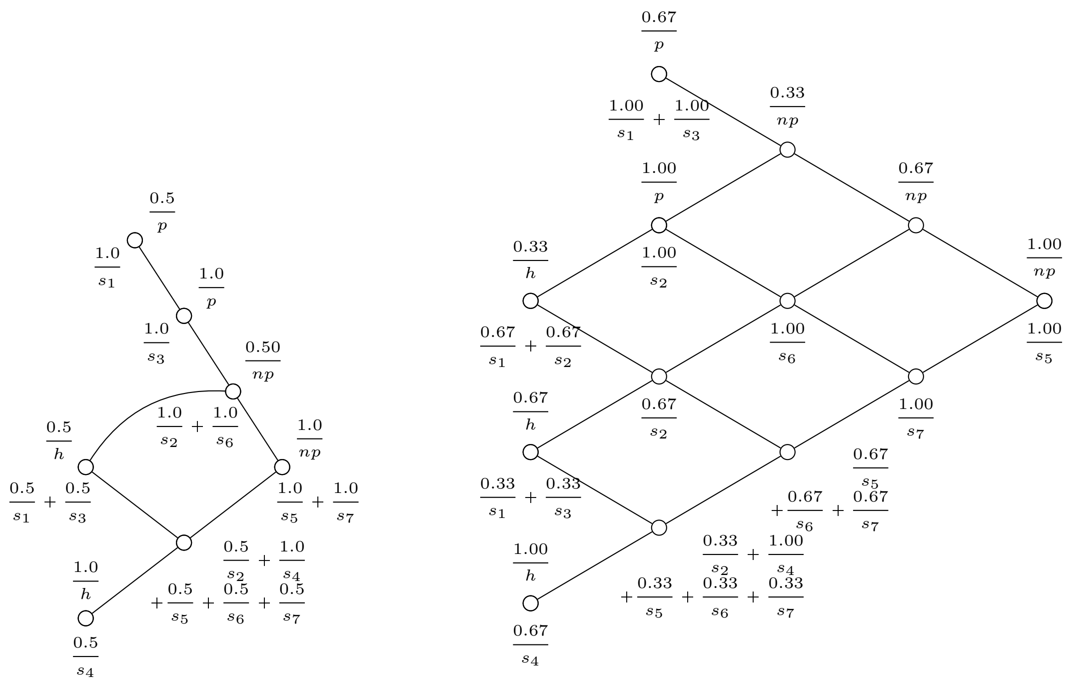
Figure 14, presents the lattices generated by $3, 4$-valued Łukasiewicz implications, whereas Fig. 15 corresponds to 5-valued implication.

8 With $n = n_0 + 1$.

9 With relative numerical errors of the order of 10^{-16}, which can be regarded as roundoff errors.

10 For finite $\mathcal{O}, \mathcal{A}$ and considering L_n, the number of concepts can never exceed n^k, where $k = \min\{|\mathcal{O}|, |\mathcal{A}|\}$ (= 3 for Table 6), $|S|$ being the number of elements (or *cardinality*) of S. Initial numerical experiments suggest that the number y of concepts can be expressed as

$$y(n) = \sum_{i=1}^{k} a_i n^i \ .$$



(a) 3-valued implication (b) 4-valued implication

Figure 14 – Medical Diagnosis: multivalued Łukasiewicz implications

4.2.3 Interpreting the diagrams

We shall focus our attention on Fig. 15. For this reason we have labelled the nodes of the diagram of 5-valued Łukasiewicz implication with a top-bottom, left-right order (which is different from the order provided by the Upper Neighbours algorithm).

Recall that in Theorem 3.1.13 we have found equations (3.1.15, 3.1.16), which give us the infimum and supremum of any collection $\langle C_\kappa \rangle_{\kappa\in K}$ of formal concepts as follows.

$$\inf_{\kappa\in K} C_\kappa = \left\langle \bigcap_{\kappa\in K} O_\kappa, \left(\bigcup_{\kappa\in K} A_\kappa \right)^{\wedge *} \right\rangle ,$$

$$\sup_{\kappa\in K} C_\kappa = \left\langle \left(\bigcup_{\kappa\in K} O_\kappa \right)^{* \wedge}, \bigcap_{\kappa\in K} A_\kappa \right\rangle .$$

This means that the attributes of $\sup_{\kappa\in K} C_\kappa$ correspond to the intersection of the attributes of the C_κ, and dually the objects in $\inf_{\kappa\in K} C_\kappa$ correspond to the intersection of the objects of the C_κ.

This has a very important consequence on how one should interpret the attributes and objects of a given node. Let us consider, for instance, node C_{14}. We see that this node is the infimum of nodes C_4 and C_9, so that one question we should ask ourselves

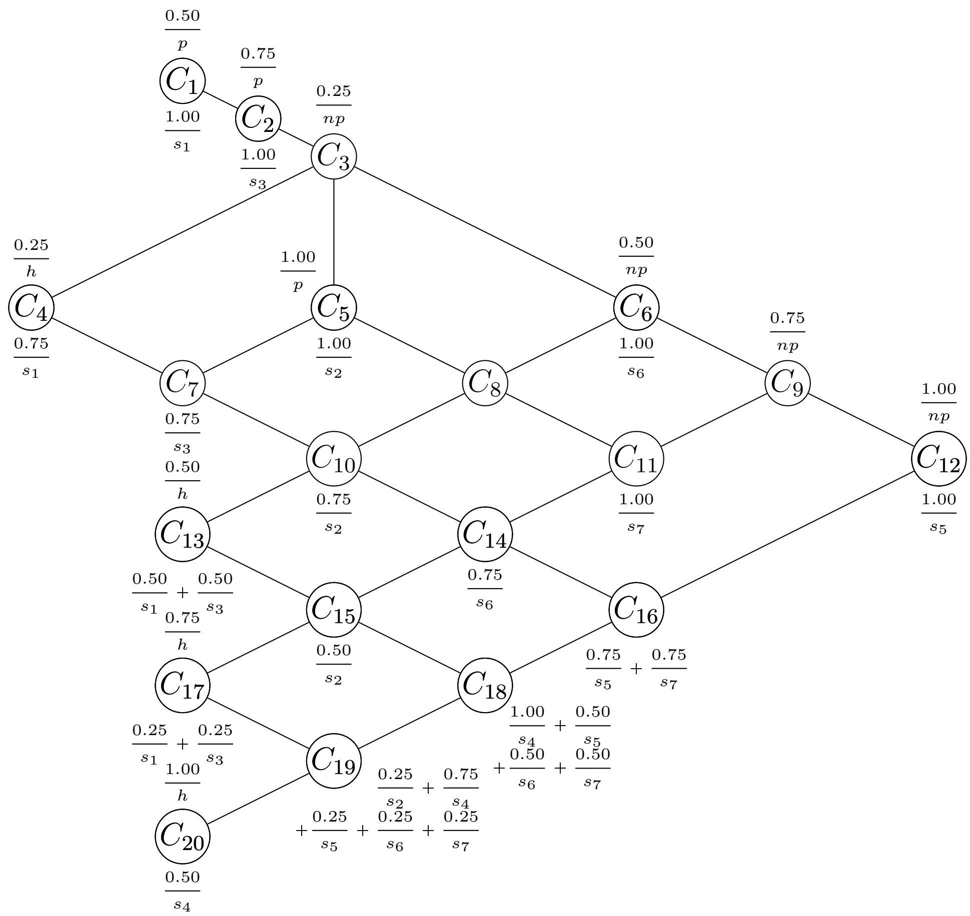


Figure 15 – Medical Diagnosis: 5-valued Łukasiewicz implication

in considering node C_{14} is: "what symptoms (objects) are common to [possibly different] persons presenting the attribute of being healthy with degree 0.25 [node C_4] and some non-pneumonia disease with degree 0.75 [node C_9]?" The answer is that every symptom, up to their respective degrees observed node C_14 — i.e., in every node below C_{14} including itself — is common to both of the conditions of being "0.25 healthy" and of presenting some "0.75 non-pneumonia disease." It might be possible however that a person with either condition would present feebler degrees of some signs.

The same reasoning could be applied to answer the dual question: "if one observes clinical sings as they are presented in node C_{14}, what does this indicate in terms of the attributes 'healthy', 'non-pneumonia disease' and 'pneumonia'?" Well, it indicates that each condition may be present up to the degree one reads in nodes above C_{14} (again, including itself).

Now, let us read some actual information from the diagram. According to C_1, only patients with (perhaps not fully developed) pneumonia present all seven clinical signs under consideration to their most severe degree. Nevertheless, but it is not usual, even among patients suffering from pneumonia, to present all these signs to such a high degree. In fact, node C_5 informs us that, although all seven signs are usually present in pneumonia patients, only s_2 and s_4 (dyspnea and tachycardia) are in their respective worst degrees. But then, that a clinical sign is not present in its most severe degree is not sufficient reason to disregard it. In fact, PEREIRA et al. conclude that

> [although] heart rate, body temperature, toxemia, and respiratory rate [signs s_4 through s_6] are important clinical signs to separate 'non-pneumonia disease' from healthy subjects, fuzzy relations complementarily inform that these clinical signs do also have a strong relationship with the diagnosis of pneumonia[11].

This is precisely what we see in C_{15}. It informs us that, although a somewhat high respiratory rate ($s_7 = 0.75$) together with severe fever ($s_5 = 1$) may be more prone to indicate some non-pneumonia disease ($np = 1$), pneumonia cannot be ruled out based on this information alone ($p = 0.75$).

The attentive reader may have noticed a disturbing information on the diagram which may seem a defect of the theory. According to C_{20}, even healthy subjects do present some degree of tachycardia. Fortunately for the theory, and unfortunately for some of the children from whom data has been collected for Table 5 to be composed, Pereira informs us that

> [...] analysis of the patients' records showed that, actually, six children in this group had mild tachycardia (three at a level of 0.25 and three at a level of 0.50), with the weight of four of them being below the 25th percentile. Thus, these children probably were not perfectly healthy but were malnourished and perhaps had the not rarely accompanying condition of anemia, which could account for the tachycardia[12].

4.2.4 Final considerations concerning applications

Why did we choose to interpret the information in Fig. 15, rather than Fig. 12 or one of the diagrams in Fig. 14? The choice of 5-valued Łukasiewicz implication of $3, 4$-valued implications is for a simple reason. By allowing Łukasiewicz implication to work with more possible truth-values we get more detailed information. However, we do not allow ourselves to increase the number of truth-values too much as this would give us excessive information, which would then be useless.

Now, comparing Gödel's implication and Łukasiewicz implication, both have their own advantages and disadvantages. First of all, the lattice constructed from either

[11] See (PEREIRA et al., 2004, p. 708)
[12] See (PEREIRA et al., 2004, p. 705) .

lattice allows us to read back the information of the corresponding fuzzy context[13]. As an example, concept C_5 in Fig. 15 is

$$C_5 = \left\langle \frac{0.50}{s_1} + \frac{1.00}{s_2} + \frac{0.75}{s_3} + \frac{1.00}{s_4} + \frac{0.75}{s_5} + \frac{0.75}{s_6} + \frac{1.00}{s_7} \right. ,$$
$$\left. \frac{1.00}{p} + \frac{0.25}{np} + \frac{0.00}{h} \right\rangle .$$

Because it is the greatest concept (with the greatest fuzzy extent) which has p as a full member of its intent (if we move up from C_5, the membership of p decreases), we conclude that the first column of Table 6 is the sequence

$$\langle 0.50, 1.00, 0.75, 1.00, 0.75, 0.75, 1.00 \rangle ,$$

corresponding to the extent of C_5.

Now, because when working with Gödel's implication (Fig. 12) we work with the original information (Table 5), Gödel's implication has a clear advantage over 5-valued Łukasiewicz implication as application of the later implies lost of (parts of) the original information. Only rough approximations remain. An attempt to remedy this problem would incur in the aforementioned problem of excessive information.

On the other hand, Łukasiewicz implication breaks down the information at hand, giving us an understanding of truly fuzzy information, such as the degrees of clinical signs associated with some non-pneumonia disease.

Thus, a good practice would be the following: if everything you need is to consider attributes (or objects) with membership degree 1 (e. g., if we are considering a fuzzy context of relations between models of cars and ages of consumers, typically it is the full model that shall be considered), it may be more interesting to construct the lattice by using Gödel's implication as the resulting lattice will preserve the original information — if there is excess of information, the fuzzy context may be adapted.

If on the other hand both objects and attributes are fuzzy, the approach that uses Łukasiewicz implication may prove itself more useful.

[13] That is, from Fig. 12 we can read Table 5, and from Fig. 15 we can read Table 6.

Conclusion

Formal Concept Analysis is an elegant mathematical theory closely related to the theory of complete lattices. As shown in this work the theory can be generalized, allowing us to work with formal contexts in which objects and attributes are only partially related to each other and, in fact, the “degree" of such a relationship may be drawn from arbitrary complete residuated lattices.

In order to guide the reader in a process of understanding how such a generalization can be achieved, we have attempted to build the main ideas up from little assumptions on the reader’s mathematical knowledge (although the ability to follow mathematical reasoning is required).

Thus, the reader is introduced to the ideas of orders, (complete) lattices and closure operators on such lattices. That — together with notions of naïve set theory — is what is necessary for an understanding of classical FCA.

Afterwards, we introduced the theory of fuzzy sets which, together with fuzzy connectives (particularly t-norms and their residua), provided us with the tools necessary to extend the classical theory. It is our belief that the examples provided throughout the text have been illuminating as a guide for intuition.

Although we have developed the results in Section 3.1 independently, we are aware that they fall within the scope of a more general approach based on complete residuated lattices, which was developed (independently) by (POLLANDT, 1997) and (BĚLOHLÁVEK, 1999). In fact, we have shown that the unit interval together with a lower semicontinuous t-norm and its residuum constitute a complete residuated lattice. Nonetheless, this more general approach tends to remain on purely theoretical grounds. We have presented one example based on a non-linear (complete) residuated lattice merely for illustrative purposes.

We are aware that potential connections between the theory here presented and other fields such as Bayesian networks and ontologies have not even been drafted in this text. We notice nonetheless that POELMANS et al. have previously performed a survey on fuzzy formal concept analysis (2014) in which they have presented references for connections with ontology engineering.

To the best of our knowledge, there are no efficient algorithms for computing the L-fuzzy concept lattice when the residuated lattice L is not linearly ordered. It is our intention to explore more on non-linearly ordered lattices L in the future and, perhaps, to develop an algorithm for (directly and efficiently) solving the problem of computing the concept lattice in such cases.

Bibliography

ANGIONI, L. *Introdução à teoria da predicação em Aristóteles*. Campinas: Editora da UNICAMP, 2006. Citado na página 65.

BARROS, L. C. de; BASSANEZI, R. C.; LODWICK, W. A. *A first course in fuzzy logic, fuzzy dynamical systems, and biomathematics: theory and applications*. Springer Berlin Heidelberg, 2017. (Studies in Fuzziness and Soft Computing). ISBN 9783662533246. Disponível em: <https://books.google.com.br/books?id=tu4QDQAAQBAJ>. Citado 2 vezes nas páginas 50 and 56.

BIRKHOFF, G. *Lattice theory*. New York: American Mathematical Society, 1948. Citado 2 vezes nas páginas 21 and 34.

BĚLOHLÁVEK, R. Lattices generated by binary fuzzy relations. *Tatra Mountains Mathematical Publications*, v. 16, 01 1999. Citado 2 vezes nas páginas 86 and 103.

BOURBAKI, N. *Elements of mathematics: general topology*. Dordrecht: Addison-Wesley, 1966. v. 1. Citado 2 vezes nas páginas 67 and 70.

BRITO, A.; BARROS, L.; LAUREANO, E.; BERTATO, F.; CONIGLIO, M. Fuzzy formal concept analysis. In: BARRETO, G. A.; COELHO, R. (Ed.). *Fuzzy Information Processing*. Cham: Springer International Publishing, 2018. p. 192–205. ISBN 978-3-319-95312-0. Citado 3 vezes nas páginas 6, 73, and 79.

BĚLOHLÁVEK, R. Fuzzy galois connections. *Mathematical Logic Quarterly*, Wiley-Blackwell, v. 45, n. 4, p. 497–504, 1999. Citado 2 vezes nas páginas 86 and 87.

BĚLOHLÁVEK, R.; VYCHODIL, V. What is a fuzzy concept lattice? *Proc. CLA 2005, 3rd Int. Conference on Concept Lattices and Their Applications*, v. 162, p. 34–45, 01 2005. Citado na página 85.

BĚLOHLÁVEK, R. Concept lattices and order in fuzzy logic. *Annals of Pure and Applied Logic*, v. 128, n. 1, p. 277 – 298, 2004. ISSN 0168-0072. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168007204000168>. Citado na página 87.

BĚLOHLÁVEK, R.; DE BAETS, B.; OUTRATA, J.; VYCHODIL, V. Computing the lattice of all fixpoints of a fuzzy closure operator. *IEEE Transactions on Fuzzy Systems*, v. 18, n. 3, p. 546–557, June 2010. ISSN 1063-6706. Citado na página 89.

FALMAGNE, J. C.; DOIGNON, J. P. Galois connections. *Learning Spaces: Interdisciplinary Applied Mathematics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 133–150. Citado na página 36.

FUENTES-GONZÁLEZ, R.; BURUSCO, A. The study of the *L*-fuzzy concept lattice. *Mathware & soft computing*, v. 1, n. 3, p. 209–218, 01 1994. ISSN 1134-5632. Citado 2 vezes nas páginas 85 and 86.

GANTER, B.; WILLE, R. *Formal concept analysis: mathematical foundations*. Berlin: Springer, 1999. Citado 6 vezes nas páginas 19, 30, 37, 41, 44, and 45.

GOGUEN, J. A. *L*-fuzzy sets. *Journal of Mathematical Analysis and Applications*, v. 18, p. 145–174, 1967. Citado na página 75.

GRÄTZER, G. *Lattice theory: first concepts and distributive lattices*. San Francisco: W. H. Freeman and Company, 1971. Citado na página 21.

HÁJEK, P. *Metamathematics of fuzzy logic*. Dordrecht: Kluwer, 1998. Citado na página 72.

HRBÁČEK, K.; JECH, T. *Introduction to set theory*, Third edition, revised and expanded. Taylor & Francis, 1999. (Chapman & Hall/CRC Pure and Applied Mathematics). ISBN 9780824779153. Disponível em: <https://books.google.com.br/books?id=Er1r0n7VoSEC>. Citado 6 vezes nas páginas 21, 26, 29, 32, 49, and 50.

JOSEPH, M. *The Trivium: The liberal arts of logic, grammar and rhetoric*. Philadelphia: Paul Dry Books, 2002. Citado 3 vezes nas páginas 19, 37, and 39.

KLEMENT, E. P.; MESIAR, R.; PAP, E. *Triangular norms*. Dordrecht: Kluwer, 2000. Citado 3 vezes nas páginas 56, 57, and 61.

LINDIG, C. Fast concept analysis. In: *Working with Conceptual Structures - Contributions to ICCS 2000*. [S.l.]: Shaker Verlag, 2000. p. 152–161. Citado 3 vezes nas páginas 47, 48, and 89.

LÓPEZ, P. B.; FUENTEZ-GONZÁLES, R. Perturbations on l-fuzzy sets. In: ______. Santander: Universidad de Cantabria, 1991. p. 43–54. Citado na página 86.

MARGARIS, A. *First order mathematical logic*. Dover, 1990. (Dover books on advanced mathematics). ISBN 9780486662695. Disponível em: <https://books.google.com.br/books?id=q3mObeaklrIC>. Citado 4 vezes nas páginas 35, 56, 57, and 65.

MENDELSON, E. *Introduction to mathematical logic*. CRC Press, 2009. (Discrete Mathematics and Its Applications). ISBN 9781584888772. Disponível em: <https://books.google.com.br/books?id=NJ7SBQAAQBAJ>. Citado 4 vezes nas páginas 23, 35, 56, and 65.

PEREIRA, J.; TONELLI, P.; BARROS, L.; ORTEGA, N. Clinical signs of pneumonia in children: association with and prediction of diagnosis by fuzzy sets theory. *Brazilian Journal of Medical and Biological Research*, scielo, v. 37, p. 701–709, 05 2004. ISSN 0100-879X. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-879X2004000500012&nrm=iso>. Citado 2 vezes nas páginas 94 and 101.

POELMANS, J.; IGNATOV, D. I.; KUZNETSOV, S. O.; DEDENE, G. Fuzzy and rough formal concept analysis: a survey. *International Journal of General Systems*, v. 43, n. 2, p. 105–134, 2014. Citado na página 103.

POLLANDT, S. *Fuzzy Begriffe*: Formale begriffsanalyse unscharfer daten. [S.l.]: Springer-Verlag Berlin Heidelberg, 1997. ISBN 978-3-642-60460-7. Citado 2 vezes nas páginas 86 and 103.

ŠEŠELJA, B.; TEPAVČEVIĆ, A. Fuzzy ordering relation and fuzzy poset. In: GHOSH, A.; DE, R. K.; PAL, S. K. (Ed.). *Pattern Recognition and Machine Intelligence*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 209–216 Citado na página 55.

ZADEH, L. Fuzzy sets. *Information and Control*, v. 8, n. 3, p. 338 – 353, 1965. ISSN 0019-9958. Disponível em: <http://www.sciencedirect.com/science/article/pii/S001999586590241X>. Citado 2 vezes nas páginas 50 and 75.