

Contents

1	Formal Concept Analysis and Extensions for Complex Data Analytics	1
	Léonard Kwuida and Rokia Missaoui	
1.1	Introduction	1
1.2	Background	2
1.2.1	Formal Concepts and Line Diagrams	2
1.2.2	Non Binary Data	4
1.2.3	Implication Computation	6
1.3	Extensions to FCA	7
1.3.1	Logical FCA	7
1.3.2	Fuzzy FCA	7
1.3.3	Relational Concept Analysis	7
1.3.4	Triadic Concept Analysis	8
1.3.5	Approximation	9
1.4	Complex Data Analytics	10
1.5	Contributions	11
	References	13
2	Conceptual Navigation in Large Knowledge Graphs	17
	Sébastien Ferré	
2.1	Introduction	17
2.2	Graph-FCA: Extending FCA to Knowledge Graphs	19
2.2.1	Graph Context	19
2.2.2	Graph Patterns	21
2.2.3	Graph Concepts	21
2.2.4	Graph Concept Lattice	23
2.3	Conceptual Navigation in Graph-FCA Lattices	23
2.3.1	Abstract Conceptual Navigation (ACN)	25
2.3.2	Graph-ACN: Instantiating ACN to Knowledge Graphs	27
2.4	Scaling to Large RDF Graphs with SPARQL Endpoints	30
2.4.1	From Graph-FCA to RDF and SPARQL	31

2.4.2	Computing the Result, Index, and Links	32
2.4.3	Living with Partial Results	35
2.5	Rising in Expressivity	35
2.5.1	An Algebraic Form of Queries	36
2.5.2	Extensions of the Query Algebra	37
2.6	The Sparklis Tool and Application Cases	38
2.6.1	Sparklis	38
2.6.2	Application Cases	40
2.7	Conclusion and Perspectives	42
	References	43
3	FCA2VEC: Embedding Techniques for Formal Concept Analysis . . .	47
	Dominik Dürschnabel, Tom Hanika, and Maximilian Stubbemann	
3.1	Introduction	47
3.2	Related Work	48
3.3	Foundations	49
3.3.1	Formal Concept Analysis	49
3.3.2	Word2Vec	50
3.4	Modeling	52
3.4.1	Retrieving FCA Features Through Closure2Vec	52
3.4.2	Object2Vec and Attribute2Vec	57
3.5	Experiments	60
3.5.1	Object2Vec and Attribute2Vec	61
3.5.2	FCA Features Through Closure2Vec	66
3.6	Conclusion	70
	References	71
4	Analysis of Complex and Heterogeneous Data Using FCA and Monadic Predicates	75
	Karell Bertet, Christophe Demko, Salah Boukhetta, Jérémy Richard, and Cyril Faucher	
4.1	Introduction	75
4.2	The NEXTPRIORITYCONCEPT Algorithm	76
4.2.1	Formal Concept Analysis	76
4.2.2	NEXTPRIORITYCONCEPT	77
4.3	Use Cases	88
4.3.1	Binary and Categorical Characteristics with the Lenses Dataset	88
4.3.2	Numerical Characteristics with the Iris Dataset	91
4.3.3	Sequential Characteristics with the Daily-actions Dataset	96
4.3.4	Sequential Characteristics with the Wine City Dataset	98
4.4	Conclusion	101
	References	102

5	Dealing with Large Volumes of Complex Relational Data Using RCA	105
	Agnès Braud, Xavier Dolques, Alain Gutierrez, Marianne Huchard, Priscilla Keip, Florence Le Ber, Pierre Martin, Cristina Nica, and Pierre Silvie	
5.1	Introduction	105
5.2	Background	107
5.3	Related Work	111
5.4	RCA for Environmental Data	114
5.4.1	Two Complex Datasets from the Environmental Domain	114
5.4.2	Experimenting RCA Algorithms	115
5.4.3	Discussion	121
5.5	Analysing Sequences from Water Quality Monitoring Using RCA	122
5.5.1	RCA-Seq	123
5.5.2	Experiments	126
5.5.3	Navigating the Resulting Hierarchy of Graphs	127
5.6	Conclusion	129
	References	130
6	Computing Dependencies Using FCA	135
	Jaume Baixeries, Victor Codocedo, Mehdi Kaytoue, and Amedeo Napoli	
6.1	Introduction	135
6.2	Notation	137
6.2.1	Equivalence Relation	137
6.2.2	Tolerance Relations	138
6.3	FCA and Database Dependencies	140
6.3.1	Functional Dependencies	140
6.3.2	Similarity Dependencies	141
6.3.3	Formal Concept Analysis	141
6.3.4	Functional Dependencies as Implications	142
6.3.5	Pattern Structures	143
6.4	Results	144
6.4.1	Characterization of Functional Dependencies with Pattern Structures	144
6.4.2	Similarity Dependencies	146
6.5	Discussion	146
6.6	Conclusions	148
	References	148
7	Leveraging Closed Patterns and Formal Concept Analysis for Enhanced Microblogs Retrieval	151
	Meryem Bendella and Mohamed Quafafou	
7.1	Introduction	151
7.2	Related Work	152

7.3	FCA-Based Query Expansion	154
7.3.1	Patterns Discovery	154
7.4	Patterns and Word Embeddings Based Query Expansion	155
7.4.1	Word Embeddings: Word2Vec Model	156
7.4.2	Expansion Terms Selection	157
7.5	Experiments	158
7.5.1	Dataset Description	158
7.5.2	Retrieval Model	159
7.5.3	Experimental Protocol	159
7.5.4	Experimental Results	160
7.6	Conclusion	163
	References	164
8	Scalable Visual Analytics in FCA	167
	Tim Pattison, Manuel Enciso, Ángel Mora, Pablo Cordero, Derek Weber, and Michael Broughton	
8.1	Introduction	167
8.1.1	Scalable Visual Analytics in FCA	168
8.1.2	Organisation	169
8.2	Graph-Theoretic Introduction to FCA	169
8.2.1	Formal Context	170
8.2.2	Formal Concepts	170
8.2.3	Concept Lattice Digraph	171
8.2.4	Line Diagram	172
8.2.5	Simplifying Implications	172
8.2.6	Visualising Implications	173
8.2.7	Coordinating Views of Implications and Concepts	174
8.3	Introduction to Visual Analytics	175
8.3.1	Algorithmic Analysis	176
8.3.2	Graph Drawing	177
8.3.3	Information Visualisation	177
8.3.4	Multiple Coordinated Views	178
8.3.5	Tight Coupling	178
8.4	Layout, Visualisation and Interaction	179
8.4.1	Reducing Digraph Size	179
8.4.2	Layout of Line Diagram	179
8.4.3	Interactive Visualisation	180
8.4.4	Discovering or Imposing Tree Structure	181
8.4.5	Demand for Enhanced Tool Support	181
8.4.6	Implications	182
8.5	Three FCA Prototypes	182
8.5.1	Hierarchical Parallel Decomposition	182
8.5.2	User-Guided FCA	183
8.5.3	Structural Navigation	186

8.6	Discovering Insightful Implications	188
8.6.1	Visualisation of Implications	188
8.6.2	Our Data Visualisation Approach	192
8.7	Conclusions and Future Work	196
	References	196
9	Formal Methods in FCA and Big Data	201
	Domingo López-Rodríguez, Emilio Muñoz-Velasco, and Manuel Ojeda-Aciego	
9.1	Introduction	201
9.2	Context and Concept Lattice Reduction Methods	204
9.3	Improved Management of Implications	209
9.4	Minimal Generators to Represent Knowledge	214
9.5	Probably Approximately Correct Implication Bases	216
9.6	Summary and Possible Future Trends	219
	References	221
10	Towards Distributivity in FCA for Phylogenetic Data	225
	Alain Gély, Miguel Couceiro, and Amedeo Napoli	
10.1	Motivation	225
10.2	Models: Lattices, Semilattices, Median Algebras and Median Graphs	227
10.2.1	Lattices and FCA	227
10.2.2	Distributive Lattices	230
10.2.3	Median Graphs	232
10.3	Algorithm to Produce a Distributive \vee -Semilattice	233
10.4	A Counter-Example for the Existence of a Minimum Distributive \vee -Semilattice	235
10.5	Discussion and Perspectives	236
	References	236
11	Triclustering in Big Data Setting	239
	Dmitry Egurnov, Dmitry I. Ignatov, and Dmitry Tochilkin	
11.1	Introduction	239
11.2	Prime Object-Attribute-Condition Triclustering	241
11.3	Triclustering Extensions	244
11.3.1	Multimodal Clustering	244
11.3.2	Many-Valued Triclustering	245
11.4	Implementations	245
11.4.1	Map-Reduce-Based Multimodal Clustering	245
11.4.2	Implementation Aspects and Used Technologies	248
11.4.3	Parallel Many-Valued Triclustering	249
11.5	Experiments	249
11.5.1	Datasets	250
11.5.2	Results	251

11.6 Experiments with Parallelisation 253

11.7 Conclusion 254

References 256

Index 259

1.1 Introduction 1

1.2 The Problem 2

1.3 The Solution 3

1.4 The Architecture 4

1.5 The Implementation 5

1.6 The Results 6

1.7 The Conclusion 7

2.1 The Problem 8

2.2 The Solution 9

2.3 The Architecture 10

2.4 The Implementation 11

2.5 The Results 12

2.6 The Conclusion 13

3.1 The Problem 14

3.2 The Solution 15

3.3 The Architecture 16

3.4 The Implementation 17

3.5 The Results 18

3.6 The Conclusion 19

4.1 The Problem 20

4.2 The Solution 21

4.3 The Architecture 22

4.4 The Implementation 23

4.5 The Results 24

4.6 The Conclusion 25

5.1 The Problem 26

5.2 The Solution 27

5.3 The Architecture 28

5.4 The Implementation 29

5.5 The Results 30

5.6 The Conclusion 31

6.1 The Problem 32

6.2 The Solution 33

6.3 The Architecture 34

6.4 The Implementation 35

6.5 The Results 36

6.6 The Conclusion 37

7.1 The Problem 38

7.2 The Solution 39

7.3 The Architecture 40

7.4 The Implementation 41

7.5 The Results 42

7.6 The Conclusion 43

8.1 The Problem 44

8.2 The Solution 45

8.3 The Architecture 46

8.4 The Implementation 47

8.5 The Results 48

8.6 The Conclusion 49

9.1 The Problem 50

9.2 The Solution 51

9.3 The Architecture 52

9.4 The Implementation 53

9.5 The Results 54

9.6 The Conclusion 55

10.1 The Problem 56

10.2 The Solution 57

10.3 The Architecture 58

10.4 The Implementation 59

10.5 The Results 60

10.6 The Conclusion 61

11.1 The Problem 62

11.2 The Solution 63

11.3 The Architecture 64

11.4 The Implementation 65

11.5 The Results 66

11.6 The Conclusion 67

12.1 The Problem 68

12.2 The Solution 69

12.3 The Architecture 70

12.4 The Implementation 71

12.5 The Results 72

12.6 The Conclusion 73

13.1 The Problem 74

13.2 The Solution 75

13.3 The Architecture 76

13.4 The Implementation 77

13.5 The Results 78

13.6 The Conclusion 79

14.1 The Problem 80

14.2 The Solution 81

14.3 The Architecture 82

14.4 The Implementation 83

14.5 The Results 84

14.6 The Conclusion 85

15.1 The Problem 86

15.2 The Solution 87

15.3 The Architecture 88

15.4 The Implementation 89

15.5 The Results 90

15.6 The Conclusion 91

16.1 The Problem 92

16.2 The Solution 93

16.3 The Architecture 94

16.4 The Implementation 95

16.5 The Results 96

16.6 The Conclusion 97

17.1 The Problem 98

17.2 The Solution 99

17.3 The Architecture 100

17.4 The Implementation 101

17.5 The Results 102

17.6 The Conclusion 103

18.1 The Problem 104

18.2 The Solution 105

18.3 The Architecture 106

18.4 The Implementation 107

18.5 The Results 108

18.6 The Conclusion 109

19.1 The Problem 110

19.2 The Solution 111

19.3 The Architecture 112

19.4 The Implementation 113

19.5 The Results 114

19.6 The Conclusion 115

20.1 The Problem 116

20.2 The Solution 117

20.3 The Architecture 118

20.4 The Implementation 119

20.5 The Results 120

20.6 The Conclusion 121

21.1 The Problem 122

21.2 The Solution 123

21.3 The Architecture 124

21.4 The Implementation 125

21.5 The Results 126

21.6 The Conclusion 127

22.1 The Problem 128

22.2 The Solution 129

22.3 The Architecture 130

22.4 The Implementation 131

22.5 The Results 132

22.6 The Conclusion 133

23.1 The Problem 134

23.2 The Solution 135

23.3 The Architecture 136

23.4 The Implementation 137

23.5 The Results 138

23.6 The Conclusion 139

24.1 The Problem 140

24.2 The Solution 141

24.3 The Architecture 142

24.4 The Implementation 143

24.5 The Results 144

24.6 The Conclusion 145

25.1 The Problem 146

25.2 The Solution 147

25.3 The Architecture 148

25.4 The Implementation 149

25.5 The Results 150

25.6 The Conclusion 151