

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Research Objectives	11
1.3	Thesis Structure	12
1	Foundations	15
2	Digitalization of Companies	19
2.1	Assessing Digital Maturity	21
2.2	A Comprehensive Digitalization Check	24
2.3	Processes as a Digital Dimension	28
3	Business Process Management and Process Mining	31
3.1	The Horus Method	32
3.2	Imperative Process Modeling	33
3.3	Declarative Process Modeling	36
3.4	Multi-perspective Declarative Constraints	41
3.5	Process Mining	43
3.5.1	History and Foundations	45
3.5.2	Event Data	47
3.5.3	Process Discovery	54
3.5.4	Conformance Checking	58
3.5.5	Declarative Process Mining	59

3.5.6	Application in Practice	59
4	Data Mining Techniques for Pattern Discovery	67
4.1	Association Rule Mining	71
4.1.1	Apriori	72
4.1.2	FP-Growth	74
4.1.3	Association Rule Generation	78
4.2	Sequential Pattern Mining	79
4.2.1	Generalized Sequential Pattern Mining (GSP) . .	81
4.2.2	Sequential Rule Mining	87
II	Combining Data and Process Mining	91
5	Applying Pattern and Rule Mining on Process Event Data	95
5.1	General Setup of the Implementation	95
5.2	Association Rule Mining on Event Logs	97
5.2.1	Translating Association Rules to Declare Constraints	102
5.3	Sequential Pattern Mining on Event Logs	111
5.3.1	Translating Sequential Patterns to Sequential Rules	116
5.3.2	Translating Sequential Patterns or Rules to Declare Constraints	118
6	Evaluation	137
6.1	Evaluation Method	138
6.2	Application on Sample Event Data	138
6.2.1	Compensation Request Process	139
6.2.2	Production Process	151
6.2.3	Purchase Order Process of Production Firm . . .	162

6.3	Overall Evaluation Results Discussion	169
III	Related Research and Conclusions	173
7	Related Work	177
7.1	Prior Work on Declarative Process Discovery	177
7.2	Multi-perspective Declarative Process Modeling and Mining	182
7.3	Related Research for Declarative Process Modeling and Mining	188
8	Conclusions	191
8.1	Summary and Implications	191
8.2	Integration into the Horus Method	193
8.3	Outlook	196
	Bibliography	199
	List of Web Pages	215
	List of Abbreviations	219
A	Overview of the Declare Template Set	221

List of Figures

1.1	Discovery of Purchase Order Process in Disco [4]	7
2.1	Digital Maturity Model with Seven Dimensions [VEGD20]	22
2.2	Radar Chart for the Maturity Values in Table 2.1	27
3.1	Overview of the Horus Method (Source: adopted from [SVOK12])	33
3.2	Sample Process Model in BPMN Notation (Source: adopted from [3])	35
3.3	Payment Process in the Horus Business Modeler [10] (based on Figure 3.2)	35
3.4	Sample Declare Constraints (based on [ADCH ⁺])	38
3.5	Declare Representation of Payment Process (Figures 3.2 and 3.3)	40
3.6	Multi-perspective Declare Model of Payment Process (based on Figure 3.5)	44
3.7	Relationships of BPM, Data Science and Process Mining (based on [Wil16])	46
3.8	Import of Insurance Compensation Request Process Event Log into Disco [4]	50
3.9	AND-split in Petri Net Notation (Source: adopted from [Wil16])	55
3.10	Discovery of Compensation Request Process in Disco [4]	57

List of Figures

3.11	Discovery of Compensation Request Process in Disco [4] (100% Paths)	58
3.12	The L* Life-cycle Model for Process Mining Projects (Source: based on [vdAAdM ⁺ 12])	62
3.13	The PM ² Process Mining Project Methodology (Source: [VELLVDA15, 26])	63
3.14	The Process Mining Project Methodology (ProMiPM)	65
4.1	Three Cases of Insurance Compensation Request Process	68
4.2	The Apriori Algorithm for Transactions in Table 4.1 (Minimum Support 0.5)	74
4.3	The FP-Growth Algorithm for Transactions in Table 4.1 (T_1 to T_5)	76
4.4	The FP-Growth Algorithm for Transactions in Table 4.1 (T_6 to T_7)	76
4.5	Prefix paths of {Keyboard} and {Monitor, Keyboard} (based on Last FP-tree in Figure 4.4)	77
5.1	Process of Association Rule Mining on Event Logs in RapidMiner [6]	97
5.2	RESPONDED EXISTENCE Constraint for Two Activities	102
5.3	RESPONDED EXISTENCE Constraint for Three Activities (1 Premise, 2 Conclusions)	103
5.4	RESPONDED EXISTENCE Constraint for Three Activities (2 Premises, 1 Conclusion)	104
5.5	RESPONDED EXISTENCE Constraint for Four Activities (2 Premises and 2 Conclusions)	105
5.6	RESPONDED EXISTENCE Constraint for Four Activities (1 Premise and 3 Conclusions)	106
5.7	RESPONDED EXISTENCE Constraint for Four Activities (3 Premises and 1 Conclusion)	106

5.8	RESPONDED EXISTENCE Constraint with Activation Condition	107
5.9	RESPONDED EXISTENCE Constraint with Activation Condition on Target Activity	107
5.10	RESPONDED EXISTENCE Constraint with Two Activation Conditions	108
5.11	RESPONDED EXISTENCE Constraint with Correlation Condition	109
5.12	Two EXISTENCE Constraints with Activation Conditions	111
5.13	Process of Sequential Pattern Mining on Event Logs in RapidMiner [6]	113
5.14	RESPONSE Constraint for Sequential Pattern with Two Activities	118
5.15	Chained RESPONSE Constraints for Sequential Pattern with Three Activities	118
5.16	Unchained RESPONSE Constraint for Sequential Pattern with Three Activities (1 Premise, 2 Conclusions)	119
5.17	Unchained RESPONSE Constraint for Sequential Pattern with Three Activities (2 Premises, 1 Conclusion)	120
5.18	Unchained RESPONSE Constraint for Sequential Pattern with Four Activities (2 Premises, 2 Conclusions)	121
5.19	EXISTENCE Constraint for <i>decide</i> Activity with Activation Condition	122
5.20	RESPONSE Constraint for Sequential Pattern with One Activity/Attribute Combination	123
5.21	RESPONSE Constraint for Sequential Pattern with Two Activity/Attribute Combinations	123
5.22	Chained RESPONSE Constraints with Multi-perspective Part	124

5.23	Unchained RESPONSE Constraints for Three Activities (Multi-perspective)	125
5.24	Generalized Unchained RESPONSE Constraint (Four Activities)	126
5.25	RESPONSE Constraint for Sequential Pattern with Stand-alone Resource Attribute	127
5.26	EXISTENCE Constraint for Sequential Pattern with Case Attribute	129
5.27	EXISTENCE/RESPONSE Constraints for Sequential Pattern with Case Attribute and Two Activities	130
5.28	RESPONSE Constraints for Sequential Pattern with Case Attribute/Activity Combination	131
5.29	PRECEDENCE Constraint for Sequential Pattern with Two Activities	133
5.30	Chained RESPONSE Constraints with Costs Attributes	134
6.1	EXISTENCE and RESPONDED EXISTENCE Constraints for Compensation Request Process	142
6.2	CO-EXISTENCE Representation of Association Rules for Compensation Request Process	143
6.3	CO-EXISTENCE Representation of Association Rules for Compensation Request Process (w/ Resource Attributes)	145
6.4	Declarative Representation of Sequential Patterns for Compensation Request Process	147
6.5	Declarative Representation of Sequential Patterns for Compensation Request Process with Resource Information (based on Table 6.6)	149
6.6	Discovery of Production Process in Disco [4]	154
6.7	Discovery of Production Process in Disco [4] with Maximum Number of Activities and 67% of Paths	155

6.8	Declare Representation of Sequential Patterns in Table 6.8	158
6.9	Declare Representation of Sequential Patterns in Table 6.8 Enriched with Results in Table 6.10	160
6.10	Discovery of Purchase Order Process in Disco [4]	164
6.11	Declare Representation of Sequential Patterns in Table 6.12	166
6.12	Declare Representation of Sequential Patterns in Table 6.13	168
8.1	Procedure Model of the Horus Method Including Process Mining	194

List of Tables

2.1	Sample Calculation of Total Digital Maturity	26
3.1	In-table Versioning of Personal Data (Source: adopted from [dMRVDA19])	49
3.2	Sample Event Log Footprint Matrix (Source: adopted from [Wil16])	55
4.1	Customer Transactions in a Computer Store (Source: based on [13])	69
4.2	Computer Store Transaction Data (Source: based on [AS95])	80
4.3	Computer Store Transactions per Customer (Source: based on [AS95])	80
4.4	Computer Store Transactions with Time Differences (Source: based on [SA96])	82
4.5	Joined Candidate Sequences of Length Two (Transactions of Table 4.4)	85
4.6	Joined Candidate Sequences of Length Two After Pruning (1/2)	87
4.7	Joined Candidate Sequences of Length Two After Pruning (2/2)	88
4.8	Joined Candidate Sequences of Length Three After Pruning	88
5.1	Compensation Request Log with Activities Aggregated by Case ID	98

List of Tables

5.2	Frequent Itemsets of Compensation Request Process Event Log Activities	99
5.3	Association Rules for Compensation Request Process Event Log	100
5.4	Association Rules for Compensation Request Process Event Log with Resource Information	101
5.5	Event Log Excerpt in Binominal Format (Activity and Resource Attributes)	114
5.6	GSP Operator Parameters in RapidMiner [6]	115
5.7	Excerpt of the GSP Output for the Compensation Request Process Dataset	116
6.1	Key Facts of Compensation Request Process	139
6.2	Frequent Itemsets for Compensation Request Process (No Minimum)	140
6.3	Association Rules for Compensation Request Process	141
6.4	Association Rules for Compensation Request Process (w/ Resource Attributes)	144
6.5	Sequential Patterns for Compensation Request Process	146
6.6	Sequential Patterns for Compensation Request Process Including Resource and Costs Information	150
6.7	Key Facts of Production Process [20]	152
6.8	Sequential Patterns for Production Process (No Optional Attributes)	156
6.9	Sequential Rules and Confidence Values (based on Table 6.8)	157
6.10	Sequential Patterns for Production Process (Four Optional Attributes)	159
6.11	Key Facts of Purchase Order Process	162

6.12	Sequential Patterns for Purchase Order Process (No Optional Attributes)	165
6.13	Sequential Patterns for Purchase Order Process (Three Optional Attributes)	167
A.1	Unary Relations in the Declare Template Set (based on [DCM13, KMDCDF16, FG19])	221
A.2	Binary Relations in the Declare Template Set (based on [DCM13, KMDCDF16, FG19])	222
A.3	Binary Relations in the Declare Template Set (based on [DCM13, KMDCDF16, FG19]), continued	223