

Contents

1 Introduction	1
1.1 Business Process Modeling	1
1.2 Process Mining	3
1.3 Book Outline	4
1.4 Website	7

Part I: State of the Art: BPM, Data Mining and Process Mining

2 Introduction to Business Processes, BPM, and BPM Systems	11
2.1 Introduction to Business Processes	11
2.1.1 Petri Nets	14
2.1.2 BPMN	15
2.1.3 YAWL	17
2.1.4 Declare	18
2.1.5 Other Formalisms	19
2.2 Business Process Management Systems	19
3 Data Generated by Information Systems (and How to Get It)	23
3.1 Information Extraction from Unstructured Sources	23
3.2 Evaluation with the F_1 Measure	24
4 Data Mining for Information System Data	27
4.1 Classification with Nearest Neighbor	27
4.2 Neural Networks Applied to Estimation	28
4.3 Association Rules Extraction	29
4.4 Clustering	29
4.4.1 Clustering with Self-organizing Maps	29
4.4.2 Clustering with Hierarchical Clustering	30
4.5 Profiling Using Decision Trees	31
5 Process Mining	33
5.1 Process Mining as Control-Flow Discovery	35
5.2 Other Perspectives of Process Mining	45

5.2.1	Organizational Perspective	45
5.2.2	Conformance Checking	46
5.2.3	Data Perspective	46
5.3	Performance Evaluation of Process Mining Algorithm	46
6	Quality Criteria in Process Mining	49
6.1	Model-to-Log Metrics	51
6.2	Model-to-Model Metrics	52
7	Event Streams	53
7.1	Data Streams	53
7.1.1	Data-Based Mining	54
7.1.2	Task-Based Mining	54
7.2	Common Stream Mining Approaches	54
7.3	Stream Mining and Process Mining	54
 Part II: Obstacles to Process Mining in Practice		
8	Obstacles to Applying Process Mining in Practice	59
8.1	Typical Deploy Scenarios	59
8.2	Problems with Data Preparation	60
8.3	Problems During the Mining Phase	62
8.4	Problems with the Interpretation of the Mining Results and Extension of Processes	62
8.5	Incremental and Online Process Mining	63
9	Long-term View Scenario	65
9.1	A Target Scenario	65
9.2	Discussion	66
 Part III: Process Mining as an Emerging Technology		
10	Data Preparation	71
10.1	Process Mining in New Scenarios	72
10.2	Working Framework for Event Logs	73
10.3	Identification of Process Instances	75
10.3.1	Exploiting A-priori Knowledge	75
10.3.2	Selection of the Identifier	76
10.3.3	Results Organization and Filtering	78
10.3.4	Deriving a Log to Mine	79
10.4	Experimental Results	80
10.5	Similar Problems and Solutions	81
10.6	Summary	83
11	Heuristics Miner for Time Interval	85
11.1	Heuristics Miner	86
11.2	Activities as Time Interval	88

11.3	Experimental Results	91
11.4	Summary	94
12	Automatic Configuration of Mining Algorithm	97
12.1	The Problem of Selecting the Right Parameters	98
12.2	Parameters of the Heuristics Miner++ Algorithm	99
12.3	Facing the Parameters Setting Problem	101
12.4	Discretization of the Parameters Values	102
12.5	Exploration of the Hypothesis Space	103
12.6	Improved Exploration of the Hypothesis Space	105
12.6.1	Factorization of the Search Space	105
12.6.2	Searching for the Best Hypothesis	106
12.7	Experimental Results	108
12.7.1	Experimental Setup	108
12.7.2	Results	109
12.8	Summary	112
13	User-Guided Discovery of Process Models	113
13.1	Clustering for Process Mining	114
13.2	Results on Clustering for Process Mining	114
13.3	Implementation	115
13.4	Summary	118
14	Extensions of Business Processes with Organizational Roles	119
14.1	Working Framework	120
14.2	Rules for Handover of Roles	123
14.2.1	Rule for Strong No Handover	124
14.2.2	Rule for No Handover	124
14.2.3	Degree of No Handover of Roles	124
14.2.4	Merging Roles	125
14.3	Algorithm Description	126
14.3.1	Step 1: Handover of Roles Identification	126
14.3.2	Step 2: Roles Aggregation	127
14.3.3	Generation of Candidate Solutions	128
14.3.4	Partition Evaluation	130
14.4	Experiments	130
14.5	Other Approaches Dealing with Organizational Perspective	134
14.6	Summary	135
15	Results Interpretation and Evaluation	137
15.1	Comparing Processes	138
15.1.1	Problem Statement and the General Approach	139
15.1.2	Process Representation	140
15.1.3	A Metric for Processes Comparison	144

15.2	A-Posteriori Analysis of Declarative Processes	147
15.2.1	Declare	147
15.2.2	An Approach for A-Posteriori Analysis	148
15.2.3	An Algorithm to Discriminate Fulfillments from Violations.	151
15.2.4	Healthiness Measures.	154
15.2.5	Experiments	156
15.3	Implementations	161
15.4	Summary	161
16	Hands-On: Obtaining Test Data.	163
16.1	A Process and Logs Generator.	164
16.1.1	The Processes Generation Phase	165
16.1.2	Execution of a Process Model.	169
16.2	Implementation	172
16.3	Summary	173
 Part IV: A New Challenge in Process Mining		
17	Process Mining for Stream Data Sources	177
17.1	Basic Concepts	178
17.2	Heuristics Miners for Streams	181
17.2.1	Baseline Algorithm for Stream Mining.	181
17.2.2	Stream-Specific Approaches	183
17.2.3	Stream Process Mining with Lossy Counting (Evolving Stream).	187
17.3	Error Bounds on Online Heuristics Miner	189
17.4	Results	191
17.4.1	Models Description	192
17.4.2	Algorithms Evaluation	192
17.5	Implementation	201
17.6	Summary	203
 Part V: Conclusions and Future Work		
18	Conclusions and Future Work	207
18.1	Wrap-Up	207
18.2	Future Work	210
References.		211

Process Mining Techniques in Business Environments
Theoretical Aspects, Algorithms, Techniques and Open
Challenges in Process Mining

Burattin, A.

2015, XII, 220 p. 101 illus., Softcover

ISBN: 978-3-319-17481-5