



Universiteit  
Leiden  
The Netherlands

## Unraveling temporal processes using probabilistic graphical models

de Paula Bueno, M.L.

### Citation

De Paula Bueno, M. L. (2020, February 11). *Unraveling temporal processes using probabilistic graphical models*. SIKS Dissertation Series. Retrieved from <https://hdl.handle.net/1887/85168>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/85168>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/85168> holds various files of this Leiden University dissertation.

**Author:** De Paula Bueno, M.L.

**Title:** Unraveling temporal processes using probabilistic graphical models

**Issue Date:** 2020-02-11

# Unraveling Temporal Processes using Probabilistic Graphical Models

MARCOS LUIZ DE PAULA BUENO

Cover design: Matheus de Paula Bueno  
Cover background image: Davide Guglielmo  
Printed by Gildeprint  
ISBN: 9789464020519



SIKS Dissertation Series No. 2020-02.

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

This thesis was supported by the Netherlands Organization for Scientific Research (NWO) as part of the “Careful” project (62001863), and by project “NORTE-01-0145-FEDER-000016” (NanoSTIMA) financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF).

# **Unraveling Temporal Processes using Probabilistic Graphical Models**

Proefschrift

ter verkrijging van  
de graad van Doctor aan de Universiteit Leiden,  
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,  
volgens besluit van het College voor Promoties  
te verdedigen op dinsdag 11 februari 2020  
klokke 15.00 uur

door

**MARCOS LUIZ DE PAULA BUENO**

geboren te Catalão, Brazilië  
in 1984

Promotor: Prof. dr. P.J.F. Lucas  
Copromotor: Dr. A.J. Hommersom (Open Universiteit)  
Promotiecommissie:  
Prof. dr. T.H.W. Bäck (secretaris)  
Prof. dr. A. Plaat (voorzitter)  
Prof. dr. M. Druzdzel (University of Pittsburgh, USA)  
Dr. S. Renooij (Universiteit Utrecht)

---

## CONTENTS

---

<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	The relevance of temporal information . . . . .	1
1.2	Probabilistic graphical models . . . . .	2
1.3	Modeling sequential behaviors . . . . .	3
1.4	Adding more expressive power . . . . .	5
1.4.1	Time-dependent representation . . . . .	5
1.4.2	Factor-dependent representation . . . . .	5
1.4.3	Subprocess representation . . . . .	6
1.5	Thesis outline . . . . .	6
<b>2</b>	<b>PRELIMINARIES</b>	<b>9</b>
2.1	Notation . . . . .	9
2.2	Bayesian networks . . . . .	10
2.2.1	Origin . . . . .	10
2.2.2	Representation . . . . .	10
2.3	Learning Bayesian networks . . . . .	13
2.3.1	Parameter learning . . . . .	13
2.3.2	Structure learning . . . . .	14
2.3.3	Decomposable scores . . . . .	16
2.4	Dynamic Bayesian networks . . . . .	17
2.4.1	Representation . . . . .	17
2.4.2	Learning . . . . .	18
2.5	Hidden Markov models . . . . .	19
2.5.1	Model architectures . . . . .	20
2.5.2	Families of HMMs . . . . .	21
2.5.3	Learning . . . . .	23
2.6	Learning with latent variables . . . . .	23
2.6.1	The expectation-maximization algorithm . . . . .	23
2.6.2	The Baum-Welch algorithm . . . . .	25
2.6.3	Number of latent states . . . . .	26
2.6.4	Structure learning with missing data . . . . .	27
<b>3</b>	<b>ASYMMETRIC HIDDEN MARKOV MODELS</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	Basic notions . . . . .	30
3.3	Asymmetric hidden Markov models . . . . .	32
3.3.1	Model specification . . . . .	32
3.3.2	Parameterization . . . . .	33
3.3.3	Representation aspects . . . . .	35

3.4	Learning . . . . .	37
3.4.1	Learning setting . . . . .	37
3.4.2	Expectation step . . . . .	37
3.4.3	Maximization step . . . . .	38
3.5	Assessment via simulations . . . . .	41
3.5.1	Model selection . . . . .	41
3.5.2	Datasets . . . . .	42
3.5.3	Results for symmetric models . . . . .	42
3.5.4	Results for asymmetric models . . . . .	44
3.6	Experiments with real-world datasets . . . . .	48
3.6.1	Datasets . . . . .	48
3.6.2	Results . . . . .	50
3.6.3	Problem insight . . . . .	53
3.7	Related work . . . . .	55
3.8	Conclusions . . . . .	57
3.A	Proofs . . . . .	58
4	PREDICTING DISEASE DYNAMICS: A CASE STUDY OF PSYCHOTIC DEPRESSION	61
4.1	Introduction . . . . .	61
4.2	Related work . . . . .	62
4.3	A probabilistic framework for capturing disease dynamics . . . . .	63
4.3.1	Latent variable modeling . . . . .	63
4.3.2	State trajectories . . . . .	64
4.3.3	Exploring medical outcomes . . . . .	65
4.3.4	Selecting states . . . . .	65
4.4	Data . . . . .	67
4.4.1	Patients . . . . .	67
4.4.2	Baseline and follow-up variables . . . . .	67
4.4.3	Depression assessment . . . . .	68
4.5	A model for psychotic depression . . . . .	68
4.5.1	General and intervention-specific model . . . . .	68
4.5.2	Model parameters and structure . . . . .	69
4.6	Results . . . . .	70
4.6.1	Model dimension . . . . .	70
4.6.2	Identified states . . . . .	71
4.6.3	Dynamics . . . . .	72
4.6.4	Comparing interventions . . . . .	72
4.6.5	Reachability trend per treatment . . . . .	73
4.6.6	Reachability trend per starting state . . . . .	73
4.7	Validation . . . . .	73
4.7.1	Model validation . . . . .	73
4.7.2	Outcome validation . . . . .	76
4.8	Conclusions . . . . .	76
4.A	Model selection scores . . . . .	77

4.B	Dynamics of intervention-specific models . . . . .	78
4.C	Confidence intervals of reachability trend differences . . . . .	78
5	<b>UNDERSTANDING MULTIMORBIDITY THROUGH CLUSTERS OF HIDDEN STATES</b>	81
5.1	Introduction . . . . .	81
5.2	Health-care event data . . . . .	83
5.2.1	Representation . . . . .	83
5.2.2	Modeling . . . . .	84
5.3	Identifying transition patterns . . . . .	84
5.3.1	Clusters of states . . . . .	84
5.3.2	Transition patterns . . . . .	84
5.4	Case study . . . . .	86
5.4.1	Variables and observations . . . . .	86
5.4.2	Sample . . . . .	87
5.4.3	Number of hidden states . . . . .	87
5.4.4	Clinical interpretation of clusters . . . . .	88
5.5	Experimental results . . . . .	89
5.5.1	Model dimension . . . . .	89
5.5.2	Clusters . . . . .	89
5.5.3	Transition patterns . . . . .	90
5.5.4	Clinical interpretation of clusters . . . . .	90
5.5.5	Are the clusters needed? A comparison to Markov chains	93
5.6	Related work . . . . .	94
5.7	Conclusions . . . . .	94
6	<b>PARTITIONED DYNAMIC BAYESIAN NETWORKS</b>	97
6.1	Introduction . . . . .	97
6.2	Related work . . . . .	99
6.3	Partitioned dynamic Bayesian networks . . . . .	100
6.3.1	Model specification . . . . .	100
6.3.2	A heuristic search procedure . . . . .	102
6.4	Empirical evaluation via simulations . . . . .	104
6.4.1	Simulation parameters . . . . .	104
6.4.2	Learning and evaluating PDBNs . . . . .	105
6.4.3	Results and discussion . . . . .	106
6.4.4	Small datasets . . . . .	108
6.5	Learning temporal models of psychotic depression . . . . .	110
6.5.1	Bayesian networks in psychiatry . . . . .	110
6.5.2	Problem description and data . . . . .	111
6.5.3	Heuristic learning . . . . .	112
6.5.4	Transition structures . . . . .	114
6.6	Model assessment from a clinical perspective . . . . .	114
6.6.1	Marginals of symptoms over time . . . . .	114
6.6.2	Predictive symptoms over time . . . . .	115

6.7	Conclusions . . . . .	120
7	<b>EXCEPTIONAL MODEL MINING USING DYNAMIC BAYESIAN NETWORKS</b>	123
7.1	Introduction . . . . .	123
7.1.1	Motivating example . . . . .	124
7.2	Related work . . . . .	125
7.3	Temporal exceptional model mining . . . . .	125
7.3.1	Temporal targets . . . . .	125
7.3.2	Subgroups . . . . .	126
7.3.3	Comparing subgroups . . . . .	128
7.3.4	Exceptional subgroups . . . . .	128
7.3.5	Problem statement . . . . .	129
7.4	Exceptional dynamic Bayesian networks . . . . .	129
7.4.1	Dynamic Bayesian networks . . . . .	129
7.4.2	Distance function . . . . .	130
7.4.3	Scoring function . . . . .	131
7.4.4	Exceptional subgroups . . . . .	131
7.5	Identifying exceptional subgroups . . . . .	132
7.5.1	Distribution of false discoveries . . . . .	132
7.5.2	Subgroup search . . . . .	132
7.5.3	Exceptionality test . . . . .	133
7.5.4	Search optimization . . . . .	134
7.6	Experiments with simulated data . . . . .	135
7.6.1	Data . . . . .	135
7.6.2	Evaluation . . . . .	135
7.6.3	Results . . . . .	136
7.6.4	Similar ground truth models . . . . .	137
7.6.5	Discussion . . . . .	138
7.7	Data of funding applications . . . . .	139
7.7.1	Data . . . . .	139
7.7.2	Discovered subgroups . . . . .	140
7.7.3	Validation . . . . .	140
7.8	Conclusions . . . . .	140
8	<b>DISCUSSION</b>	143
8.1	Contributions . . . . .	143
8.1.1	Asymmetry in models . . . . .	143
8.1.2	Generation of hypotheses on processes . . . . .	144
8.1.3	Capturing hidden (non-observed) aspects of processes . . . . .	144
8.1.4	Taking into account the size of datasets . . . . .	144
8.1.5	Temporal subgroups . . . . .	145
8.2	Future work . . . . .	145
8.2.1	Asymmetry in models . . . . .	145
8.2.2	Generation of hypotheses on processes . . . . .	145
8.2.3	Capturing hidden (non-observed) aspects of processes . . . . .	146

8.2.4	Taking into account the size of datasets . . . . .	146
8.2.5	Temporal subgroups . . . . .	147
<b>BIBLIOGRAPHY</b>		<b>149</b>
<b>SUMMARY</b>		<b>163</b>
<b>SAMENVATTING</b>		<b>165</b>
<b>ACKNOWLEDGMENTS</b>		<b>167</b>
<b>CURRICULUM VITAE</b>		<b>169</b>
<b>SIKS DISSERTATIONS</b>		<b>171</b>

