On Causal Discovery and Inference from Observational Data

by

Fujin Zhu

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

Centre for Artificial Intelligence (CAI), School of Computer Science
Faculty of Engineering and Information Technology (FEIT)
University of Technology Sydney

August, 2019

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by the Australian Government Research Training Program

Signature of Candidate

To my parents, brother, and wife.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisors Prof. Guangquan Zhang and Prof. Jie Lu for their continuous encouragement and guidance in my PhD career. I really appreciate them for offering me the opportunity to study and do my research in the Centre for Artificial Intelligence (CAI, former QCIS), Faculty of Engineering and Information Technology (FEIT), University of Technology Sydney (UTS).

As my principle supervisor, Prof. Zhang is more like a father to me. I still remember the meetings we had together. He always gives me sufficient freedom to learn new knowledge and explore unknowns. When my research got stuck and I started to vacillate and loss faith in myself, he believed in me and encouraged me to calm down, take a break and then carry on. He teaches me to look faraway and to become an independent researcher. He gives me so much patience even though I learned far too slow. I am always grateful to him.

I am very lucky to have Prof. Lu as my co-supervisor. Prof. Lu is an excellent and respectable researcher. She loves her students and influences us with her continual passion for almost everything. Prof. Lu shares with us her experience in study, research and life. She teaches me how to write professional papers, discusses and listens to my problems in research, and provides comments and suggestions which are really constructive and make this thesis possible. I learn a lot from her.

I am also indebted to the rest of my committee – Dr. Haiyan Lu and Dr. Ling Chen – who have kindly contributed their time to make me a better researcher. I would like to thank all the other faculty, students, and staff at CAI and FEIT who have helped and influenced me. They are Camila Cremonese, Lily Qian, Janet Stack, Yi Zhang, Hongshu Chen, Junyu Xuan, Shirui Pan, Jia Wu, Mingmig Gong, Weiwei Liu, Anjin Liu, Fan Dong, Qian Zhang, Shan Xue, Wei Wang, Peng Hao, Guanjin Wang, Chenlian Hu, Yiliao Song, Feng Liu, Adi Lin, Dian Ouyang, Daokun Zhang, Yuangang Pan, Pingbo Pan, Ruiqi Hu, Di Wu, Zhibin Li and Xiaofeng Xu.

I would also like to thank my friends, Daniel Kim and Kai Austin Zhang, for their companion and the joyful days we spent together in a foreign country. The scholarship for my study and funding for my academic travels have come from a variety of sources. They are the China Scholarship Council, UTS, CAI, the School of Computer Science, and the FEIT.

Finally and above all, I would like to thank my family for their support all the way. They are my parents, brother and wife. During this long journey, they always believed and supported me unconditionally. I dedicate this thesis to them. I especially thank my wife, Mang Chen, who has been taking care of my everyday life and shared all my pain, sorrow and joy during my study and research. Her optimism and love makes all the dark days hopeful. No words could express my gratitude and love to her.

Fujin Zhu

Sydney, Australia, 2019

ABSTRACT

Causality is a fundamental component in all fields of science. In contrast to associational dependencies that are widely used in existing predictive machine learning and data-mining methods, causality implies the mechanism of how variables take their values and how the change of causes would lead to the change in the outcome. In the era of big data, for scientific discovery and rational decision-making, we fundamentally need methods for learning causal relationships between variables and estimating causal effects from observational data.

In this thesis, we aim to develop new models and algorithms for learning causal relationships and estimating causal effects using observational data. In particular, for the purpose of modelling and learning causal relationships from observational data, we study dynamic causal systems with feedbacks. To overcome the weakness of existing models that are unable to model both instantaneous and cross-temporal causal relations simultaneously, we propose a First-order Causal Process (FoCP) model and a causal structure learning algorithm to learn the causal graph of FoCPs from time series. For the purpose of estimating treatment effects, we investigate a range of existing methods for causal effect estimation, and propose three new methods using advanced machine learning techniques. First, to relieve the high-variance issue of the classic Inverse Propensity Weighting (IPW) estimator and thus to get more stable treatment effect estimates, we reframe it to the importance sampling framework and propose a novel Pareto-smoothing method using the generalized Pareto distribution from the extreme value statistics. Second, for causal inference with unobserved confounders, we take advantage of proxy variables and use deep latent variable models to model the underlying data-generating process. Building on recent advances in Bayesian inference and deep generative models, we propose a Causal Effect Implicit Generative Model (CEIGM). Finally, with an observation that most of existing methods for causal inference are essentially indirect in that they estimate the target treatment effects by first estimating other auxiliary quantities, we propose the idea of direct treatment effect estimation. Based on this idea, we further propose two deep neural networks for direct treatment effect estimation.

We evaluate all the methods proposed in this thesis using simulated, semi-simulated or real-world data. Experiment results show that they perform generally better than their competitors. Given the key importance of learning causality and causal inference in both theory and real-world applications, we argue that our proposed models and algorithms are of both theoretical and practical significance.

Dissertation directed by

Prof. Guangquan Zhang, Dist. Prof. Jie Lu, and Prof. Donghua Zhu

Center for Artificial Intelligence, School of Computer Science, FEIT

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES