Experimental Learning of Causal Models with Latent Variables

Sam Maes Stijn Meganck Philippe Leray

October 12, 2006

1 Introduction

This article discusses graphical models that can handle latent variables without explicitly modeling them quantitatively. There exist several paradigms for such problem domains. Two of them are *semi-Markovian causal models* and *maximal ancestral graphs*. Applying these techniques to a problem domain consists of several steps, typically: structure learning from observational and experimental data, parameter learning, probabilistic inference, and, quantitative causal inference.

A problem is that research in each of the existing approaches only focuses on one or a few of all the steps involved in the process of modeling a problem including latent variables. In other work we have investigated the integral process from observational and experimental data unto different types of efficient inference. The goal of this article is to focus on learning the structure of causal models in the presence of latent variables from a combination of observational and experimental data.

Semi-Markovian causal models (SMCMs) are an approach developed by Pearl and Tian [3, 6]. They are specifically suited for performing quantitative causal inference in the presence of latent variables. However, at this time no efficient parametrisation of such models is provided and there are no techniques for performing efficient probabilistic inference. Furthermore there are no techniques to learn these models from data issued from observations, experiments or both.

Maximal ancestral graphs (MAGs) are an approach developed by Richardson et al. [4]. They are specifically suited for structure learning in the presence of latent variables from observational data. However, the techniques only learn up to Markov equivalence and provide no clues on which additional experiments to perform in order to obtain the fully oriented causal graph. See [1, 2] for that type of results for Bayesian networks without latent variables. Furthermore, as of yet no parametrisation for discrete variables is provided for MAGs and no techniques for probabilistic inference have been developed. There is some work on algorithms for causal inference, but it is restricted to causal inference quantities that are the same for an entire Markov equivalence class of MAGs [5, 7].

We have chosen to use SMCMs as a final representation in our work, because at this time they are the only formalism that allows to perform causal inference while fully taking into account the influence of latent variables. However, we will combine existing techniques to learn MAGs with newly developed methods to provide an integral approach that uses both observational data and experiments in order to learn fully oriented semi-Markovian causal models.

We will start the paper by introducing SMCMs and MAGs and by discussing their semantical and other differences. Then we will discuss the problems that have to be solved to tranform a complete partially ancestral graph (CPAG), this is a representative of the Markov equivalence class for MAGs, into a SMCM. We will see that in general experiments are necessary to be able to perform this transformation. Therefore, we will discuss what we understand by an experiment and more specifically how their results can be interpreted. After that we will introduce a method that transforms a CPAG learned from observational data into a SMCM with the help of experiments. The result of such a transformation is the single structure that represents all the immediate causal relations and the latent confounders between the observable variables and that is crucial knowledge in order to perform causal inference.

Finally we will discuss our results and give some indications for possible future research.

References

[1] F. Eberhardt, C. Glymour, and R. Scheines. On the number of experiments sufficient and in the worst case necessary to identify all causal relations among n variables. In *Proc. of the 21st Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 178–183, 2005.

[2] S. Meganck, P. Leray, and B. Manderick. Learning causal bayesian networks from observations and experiments: A decision theoretic approach. In *Modeling Decisions in Artificial Intelligence, LNCS*, pages 58–69, 2006.

[3] J. Pearl. *Causality: Models, Reasoning and Inference.* MIT Press, 2000.

[4] T. Richardson and P. Spirtes. Ancestral graph markov models. Technical Report 375, Dept. of Statistics, University of Washington, 2002.

[5] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction and Search.* MIT Press, 2000.

[6] J. Tian and J. Pearl. On the identification of causal effects. Technical Report (R-290-L), UCLA C.S. Lab, March 2002.

[7] Jiji Zhang. *Causal Inference and Reasoning in Causally Insufficient Systems.* PhD thesis, Carnegie Mellon University, July 2006.