

# ScholarWorks@UMass Amherst

## OPEN ACCESS DISSERTATIONS

### **Title**

Learning the Structure of Bayesian Networks with Constraint Satisfaction

### **Author**

**Andrew Scott Fast**, *University of Massachusetts Amherst* Follow

### **Date of Award**

2-2010

### **Document Type**

Open Access Dissertation

### **Degree Name**

Doctor of Philosophy (PhD)

### **Degree Program**

Computer Science

### **First Advisor**

David Jensen

### **Second Advisor**

Oliver Brock

## Third Advisor

Victor Lesser

## Subject Categories

Computer Sciences

## Abstract

A Bayesian network is graphical representation of the probabilistic relationships among set of variables and can be used to encode expert knowledge about uncertain domains. The structure of this model represents the set of conditional independencies among the variables in the data. Bayesian networks are widely applicable, having been used to model domains ranging from monitoring patients in an emergency room to predicting the severity of hailstorms. In this thesis, I focus on the problem of learning the structure of Bayesian networks from data. Under certain assumptions, the learned structure of a Bayesian network can represent causal relationships in the data. Constraint-based algorithms for structure learning are designed to accurately identify the structure of the distribution underlying the data and, therefore, the causal relationships. These algorithms use a series of conditional hypothesis tests to learn independence constraints on the structure of the model. When sample size is limited, these hypothesis tests are prone to errors. I present a comprehensive empirical evaluation of constraint-based algorithms and show that existing constraint-based algorithms are prone to many false negative errors in the constraints due to run- ning hypothesis tests with low statistical power. Furthermore, this analysis shows that many statistical solutions fail to reduce the overall errors of constraint-based algorithms. I show that new algorithms inspired by constraint satisfaction are able to produce significant improvements in structural accuracy. These constraint satisfaction algo- rithms exploit the interaction among the constraints to reduce error. First, I introduce an algorithm based on constraint optimization that is sound in the sample limit, like existing algorithms, but is guaranteed to produce a DAG. This new algorithm learns models with structural accuracy equivalent or better to existing algorithms. Second, I introduce an algorithm based constraint relaxation. Constraint relaxation combines different statistical techniques to identify constraints that are likely to be incorrect, and remove those constraints from consideration. I show that an algorithm combining constraint relaxation with constraint optimization produces Bayesian networks with significantly better structural accuracy when compared to existing structure learning algorithms, demonstrating the effectiveness of constraint satisfaction approaches for learning accurate structure of Bayesian networks.

## Recommended Citation

Fast, Andrew Scott, "Learning the Structure of Bayesian Networks with Constraint Satisfaction" (2010). *Open Access Dissertations*. 182.  
[https://scholarworks.umass.edu/open\\_access\\_dissertations/182](https://scholarworks.umass.edu/open_access_dissertations/182)

[Download](#)

DOWNLOADS

Since April 07, 2010

Included in

[Computer Sciences Commons](#)

Share

COinS