Structure Learning of Causal Bayesian Networks: A Survey

Ashique Mahmood
Department of Computing Science
University of Alberta, Edmonton, Canada
Email: ashique@cs.ualberta.ca

Abstract

Causality is a fundamental concept in reasoning. The effectiveness of many reasoning tasks depends on the understanding of the underlying cause-effect relationships. Therefore, the notion of causality has been explored in a wide range of disciplines. Causal discovery, however, was not modeled as a machine learning task until recently. Many learning approaches have recently been developed and applied to capture causation. The most frequently used approach among them is learning causal Bayesian networks (CBNs). A powerful calculus, capable of causal reasoning, has been formalized through CBNs. In this paper, we reviewed the fundamentals of learning causal structures using CBNs. We distinguished between observation and intervention, a crucial concept for learning CBNs. We reviewed some methods for learning from observational and interventional data. We have noted that, as a growing field of research, learning CBN structure is being investigated with increasingly difficult problems and possibilities are arising for incorporating it to other learning problems, such as active learning.

1 Introduction

Causal reasoning has become a central point of attention in many practical areas. For example, in medical diagnosis, researchers are concerned in discovering the conditions, events or genes that are the causes of a certain disease. Causal reasoning is also essential in resolving problems in finance such as prediction of market prices and risk management [16].

In recent decades, artificial intelligence (AI) research has a growing interest in causal reasoning [11]. Discovery of the underlying model of a world often requires knowledge about cause-effects. Hence, automated causal reasoning has recently developed as an active field of machine learning.

Many different methods have been developed to capture causal relationships [7],[12]. The most frequently used approach is using Bayesian networks (BNs), stemming from the works of Pearl [10]. In fact, BNs as well as many other directed acyclic graphical models in statistical and AI applications, were originally intended as a formalism for causal reasoning [9]. BNs are frequently used for capturing the probabilistic model.

BNs are inadequate for capturing causal relationships. It is because probabilistic calculus itself is incapable of modeling causality. Pearl illustrated it through Simpson's paradox, a phenomenon well known to statistical community [9]. It occurs when conditioning on an event, a probability measure increases but in each of its subpopulation, the probability measure decreases. Therefore, a conclusion at subpopulation level contradicts with a conclusion at total-population level. Pearl showed that, an attempt to model causality with probabilities can fall into this paradox.

Pearl formalized causal calculus, which can avoid the Simpson's paradox. He introduced causal Bayesian networks (CBNs) that are more effective than general BNs for causal discovery. A central concept of CBNs is intervention: perturbation of a variable from an external agent. When an external agent forces a system variable to take a specific value, its previous causal relationships are cut off. CBNs can answer intervention queries, such as, whether a patient gets well if she is given a specific medicine. It is different than the case where a patient takes the medicine on her own. CBNs are able to capture causal structures from observational as well as interventional data.

In this paper, we survey different methods for learning CBN structures. In section 2, we develop the conceptual framework necessary for the understanding of the subsequent discussions. In section 3, we describe some methods used for learning CBNs from observational data. In section 4, we describe some methods for learning CBNs from interventional data. In section 5, we conclude with a summary of the trends of such works.

2 Conceptual Framework

To understand structure learning in CBNs, we first need to define Bayesian networks and described their properties. Then, we define intervention and causal Bayesian networks.

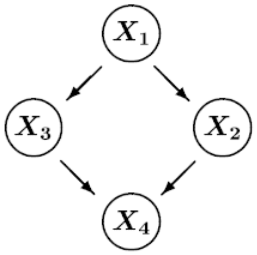

Figure 1. Example of a Bayesian network

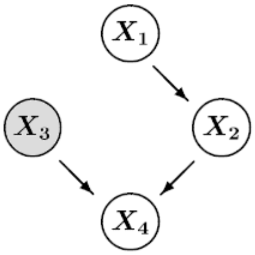

Figure 2. Example of a mutilated network

2.1 Bayesian Networks

A Bayesian network (BN) is a directed acyclic graph (DAG) G where nodes represent random variables and arcs represent probabilistic dependencies among them. An example of such a BN with four variables is depicted in Figure 1. A BN encodes the joint probability P over a set of variables $V = \{X_1, X_2, ..., X_n\}$ and decomposes it into a product of the conditional probability distributions over each variable given its parents in the graph. That is,

$$P(X_1, X_2, ..., X_n) = \prod_{i=1}^{n} P(X_i \mid Pa(X_i)) \qquad (2.1)$$

$Pa(X_i)$ is the set of parents of X_i. Given that the distribution factorizes according to the G as above, the *local Markov condition* holds, that is,

$$P(X_i \mid Pa(X_i), Z) = P(X_i \mid Pa(X_i))$$

where, Z is any set of non-descendant nodes of X_i except the nodes in $Pa(X_i)$. P is said to be a *Markov relative* to G.

It asserts that, given its parents, the probability of each variable is independent of its non-descendants in the graph, concisely described as $I(X_i \perp Z \mid Pa(X_i))$. Note that, whenever X is independent of Y given Z, it can be written as $I(X \perp Y \mid Z)$.

All such independence assertions derived from G are denoted as $I(G)$. There can be several Gs that represent the same set of independencies. Given two different graphs $G_1 = \langle V, E_1 \rangle$ and $G_2 = \langle V, E_2 \rangle$ with set of arcs, $E_1 \neq E_2$, if $I(G_1) = I(G_2)$, then G_1 and G_2 are called I*-equivalent*. All such graphs form an I*-equivalence class*.

Note that, according to *local Markov condition*, it is sufficient to know the conditional probability distribution (CPD) of X_i in the form of $P(X_i \mid Pa(X_i))$. The graph G and CPDs associated with the nodes form a full description of a BN.

2.2 Causal Bayesian Networks

A causal Bayesian network (CBN) is a BN where each arc is interpreted as a cause-effect relationship from the parent to the child. A CBN satisfies the *causal Markov condition*: given the direct causes, the phenomenon associated with a node is independent of its non-effects. This assumption permits the joint distribution of the variables in a CBN to be factored as in Equation 2.1.

A powerful feature of a CBN is that it can answer intervention queries as well as probabilistic queries. An *intervention* over a random variable X_i to x_i, denoted as $do(X_i = x_i)$, is a perturbation from an external agent, forcing the value of X_i to be fixed at x_i. An intervention query takes a form of $P(Y \mid do(X = x))$. In the following, we describe observational data and interventional data.
Observational data is passively observed data. Values taken by each variable are determined by the causal interactions between them within the system.
Interventional data is a record of phenomena perturbed by external agents. Some variables in the system are given fixed values by an external force, most often by an experimenter.

Intervention essentially separates X_i from its direct causes. An intervention can be described by a mutilated network derived from the original network of the joint distribution, by removing all the incoming arcs to X_i and setting the CPD of X_i as $P(X_i \mid Pa(X_i)) = I(X_i = x_i)$. The network in Figure 2 is a mutilated network resulting from the network in Figure 1 by intervening on variable X_3.

Pearl defines CBNs more formally as follows:

Definition (*causal Bayesian network*). A DAG G is said to be a *causal Bayesian network* compatible with every possible interventional distribution $P(v \mid do(X = x))$, where v is an assignment of values to V and $X \subset V$, allowing $X = \phi$, if and only if:

- Every $P(v|do(X = x))$ is a *Markov relative* to G;
- $P(X_i \mid do(X = x)) = I(X_i = x_i)$ with $X_i \subset X$;
- $P(X_i \mid Pa(X_i), do(X = x)) = P(X_i \mid Pa(X_i))$ with $X_i \not\subset X$.

Therefore, if a CBN can capture the true causal relationships between the variables, then any interventional query on such variables can be answered by deriving the mutilated network from the CBN and applying a truncated factorization,

$$P(v \mid do(X = x)) = \prod_{i | V_i \notin X} P(v_i \mid pa_i)$$

Note that an I-equivalence of a CBN can represent the same probabilistic independencies but not the same causal relationships.

3 Learning CBN from Observational Data

Here we will consider the task of learning CBNs from completely observational data. Structure learning of CBNs in observational data is essentially the same as structure learning of BNs.

There are two classes of methods for structure learning of BNs: constraint-based methods and score-based methods. Here we discuss some seminal approaches of these methods.

3.1 Constraint-based Approaches

These approaches are based on the qualitative properties of the probability distribution. In these approaches, the set of independencies are derived from the empirical distribution and used as constraints in the structure learning.

PC Algorithm

Spirtes et al. devised PC algorithm for recovering DAG structures [13]. The algorithm takes a set of independencies derived from independence test over the variables in domain V (found in the empirical distribution) and outputs a partially directed graph.

The algorithm starts with a complete undirected graph among the variables in V. The first major step of the method is to delete an edge between x and y, if there exists a $Z \subset V \setminus \{x, y\}$ such that $I(x \perp y \mid Z)$. The basis of this step is that, if there exists an edge between x and y, they cannot be conditionally independent for any Z. This step builds the skeleton of the graph.

The second major step is, if there is a triplet $x - y - z$ with no $x - z$ and there exists no $Z \subset V \setminus \{x, z\}$ such that $I(x \perp z \mid Z \cup y)$, then it is replaced with $x \to y \leftarrow z$. The basis is that, if the orientation is not so, $I(x \perp z \mid y)$ will be immediately true. So, if there is no such conditional independence between x and z and the undirected orientation is $x - y - z$, then it must be $x \to y \leftarrow z$.

The final step ensures that the arc-directions preserve the independencies. For example, direction is imposed if that is necessary to avoid a directed cycle.

This algorithm has worst-case complexity bounded by $O(|V|^q)$ with high probability, where q is the maximum number of adjacent nodes for any node in the graph. The algorithm usually results into a partially directed acyclic graph (PDAG). The PDAG represents an I-equivalence class. These approaches are highly sensitive to erroneous independence test results, more likely when data sets are small.

3.2 Score-Based Approaches

Score-Based approaches quantitatively distinguish between BN structures. There are two separate tasks in such methods. First, a measure is defined to score BNs: structures that better represent the empirical distribution should be scored higher. The second task is to devise a search algorithm (e.g., local, global, heuristic) that uses the score metric.

K2 Algorithm

Cooper et al. devised a score metric and an algorithm, namely K2, for learning DAG structures [1]. It requires some assumptions, such as, discrete value assumption, independent data assumption, complete data assumption and parameter independence assumption. The score metric helps to find the most probable structure G given the data D. It essentially maximizes $P(G \mid D)$. However, as $\frac{P(G_1|D)}{P(G_2|D)} = \frac{P(G_1,D)}{P(G_2,D)}$, K2 algorithm looks for a network structure G that maximizes $P(G, D)$.

They represented the probability in terms of local parent-child subgraphs. Their measure is:

$$P(G, D)$$

$$= P(G) \int P(D \mid G, \theta_G) P(\theta_G \mid G) d\theta_G \tag{3.1}$$

$$= P(G) \int \prod_{h=1}^{m} P(C_h \mid G, \theta_G) P(\theta_G \mid G) d\theta_G \tag{3.2}$$

$$\begin{aligned} = P(G) \int \prod_{h=1}^{m} \prod_{i=1}^{n} & P(X_i = d_{ih} \mid Pa(X_i) = d_{Pa(X_i)h}, \theta_G) \\ & \times P(\theta_G \mid G) d\theta_G \end{aligned} \tag{3.3}$$

$$\begin{aligned} = P(G) \int \prod_{i=1}^{n} \prod_{j=1}^{q_i} \prod_{k=1}^{r_i} & P(X_i = v_{ik} \mid Pa(X_i) = w_{ij}, \theta_G)^{N_{ijk}} \\ & \times P(\theta_G \mid G) d\theta_G \end{aligned} \tag{3.4}$$

$$= P(G) \prod_{i=1}^{n} g(X_i, Pa(X_i)) \tag{3.5}$$

$$g(X_i, Pa(X_i)) = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod N_{ijk}$$

where, θ_G is CPD parameters associated to the nodes of G, C_h is the hth data instance in D. d_{ih} is the value X_i takes in hth data instance. v_{ik} is the kth value of X_i and w_{ij} is the jth assignment of $Pa(X_i)$. N_{ijk} is the number of data instances in which $X_i = v_{ik}$ and $Pa(X_i) = w_{ij}$. m is the total number of instances, n is the total number of variables, r_i is the total number of values X_i can take and q_i is the total number of assignments $Pa(X_i)$ can have. $g(X_i, Pa(X_i))$ is a measure for the local subgraphs consisting child and parents.

Equation 3.2 is due to the independent data assumption. Equation 3.3 is due to the complete data assumption that allows the factorization. Equation 3.4 is due to discrete variable and parameter independence assumption.

The algorithm assumes that there exists an ordering of the variables such that if X_j precedes X_i in the ordering then $X_i \to X_j$ is not allowed. Then the algorithm iterates over the following steps to find the parents of each node X_i.

First, it selects a $X_j \in \{X_1, X_2, , X_{i-1}\} - Pa(X_i)$ such that $g(X_i, Pa(X_i) \cup X_j)$ is maximum. Then X_j is included as X_i's parent if $\Delta = g(X_i, Pa(X_i) \cup X_j) - g(X_i, Pa(X_i))$ is greater than zero. The loop iterates until Δ falls below a threshold or $Pa(X_i)$ has become maximal. The time complexity of K2 algorithm is $O(Nu^2n^2r)$, where N is the number of data instances, u is the maximum number of parents allowed, n is the number of variables and r is the maximum number of possible values a variable can take.

BDe Metric

Heckerman et al. devised a Bayesian score-based algorithm that can use a BN for prior knowledge [6]. Such guidance through prior knowledge is a great assistance for structure learning of CBNs when some causal links are already known. They proposed the BDe metric (Bayesian metric with Dirichlet priors and equivalence) that requires additional assumptions in addition to the assumptions in the previous method, such as, parameter modularity assumption and likelihood equivalence assumption. The measure is:

$$P(G, D) = P(G) \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + N_{ijk})}{\Gamma(\alpha_{ijk})}$$

The derivation follows from Equation 3.4 after incorporating the additional assumptions. It gives the same score to all I-equivalent BNs. Their algorithm outputs an I-equivalence class of networks.

Although, among the BNs in I-equivalence class, any BN is as useful capturing the probabilistic model as the other, but they yield different cause-effect relationships, giving rise to different answers to causal queries. Many non-equivalent networks can have reasonably high scores. Therefore, a better approach is to use Bayesian model averaging that quantify the probability mass of structures. Other approaches include, combining both the constraint-based and score-based approaches.

However, it is not always possible to identify the complete CBN structure from observational data alone. To distinguish between structures further, we need interventional data.

4 Learning CBNs from Interventional Data

Learning CBN structure from interventional data is significant in many ways. First, CBNs are formulated to consider data obtained from interventions. Such data is available and arises in many scientific experiments. It is, hence, natural to incorporate interventional data for learning CBN structures. Second, learning from observational data can only identify models up to I-equivalence, where interventional data can help discover causal relationships further.

Both constraint-based and score-based approaches can be applied for learning CBNs from interventional data. However, enough data instances of any given type of intervention for a reliable independence test is generally not available. On the other hand, the score-based approach is flexible for dealing with a combination of observational and interventional data, having interventions at different nodes.

There are two major axes along which types of intervention can vary. One is whether an intervention is perfect or imperfect. A perfect intervention occurs when an intervention sets a variable to a fixed value. In imperfect intervention, the variable does not take a fixed value, but takes a different distribution of values than the original one, upon intervention.

Another axis asks whether an intervention is deterministic or nondeterministic. A deterministic intervention is an assured intervention on a variable by experimenter. In a nondeterministic intervention, the experimenter is uncertain whether the execution of the intervention will be successful or not. A classical example, as mentioned by Cooper et al., is when a group of patients volunteer to participate in a study and given a medicine to take but a patient can decide not to take it [2].

Not all the cases are interesting, practical or thoroughly explored. Here we discuss some of the works on these different cases.

4.1 Learning from Data with Perfect Intervention

One of the first works on learning CBN structures from mixed data is due to Heckerman [5]. He proposed a score-based method for mixed data where interventional data are perfect and deterministic. Together with the assumptions necessary for the BDe metric, derived by the same author, he proposed some additional assumptions: mechanism independence and component independence. These assumptions are similar to the parameter independence of BDe parameter but applicable in an interventional setting upon the intervened nodes.

The most important step is, for each intervened variable X_i in a data instance, every incoming arc to X_i is removed and $p(X_i \mid Pa(X_i))$ is changed to $I(X_i = x_i^*)$, where x_i^* is the value at which X_i is intervened. Essentially, this changes the conditional probability distribution (CPD) to be 1 when $X_i = x_i^*$ and 0 elsewhere.

Let's see how it affects the score metric. We are primarily interested in finding the network structure G that maximizes $P(G \mid D)$. It is essentially the same as maximizing $P(G, D)$ for the same data. Now, $P(G, D)$ follows the derivation as Equation 3.1-3.4, except that in Equation 3.4 for this case, N_{ijk} means the number of data instances in which $X_i = v_{ik}$ and $Pa(X_i) = w_{ij}$ where X_i *is not intervened*. It differs from the N_{ijk} in K2 algorithm as there was no interventional data in that setting. The difference is due to the fact that whenever X_i in data instance h is intervened, $P(X_i = d_{ih} \mid Pa(X_i) = d_{Pa(X_i)h}, \theta_G)$ is set to 1.

Now, consider a network over two variables X and Y. To find which one is cause and which one is effect, if any, we look for $P(X \to Y \mid D)$ and $P(Y \to X \mid D)$. Given only observational data, both networks $X \to Y$ and $X \leftarrow Y$ would be I-equivalent and hence indistinguishable. This is because, the count N_{ijk} is symmetrical for a pair of variables in observational data. However, in the interventional data, N_{ijk} will render an asymmetrical count for this pair of variables, given that only one of them is intervened in a data instance. Therefore, the metric will score these two networks differently. Note that, intervening on both variables here is of no use.

Perfect Nondeterministic Intervention

Cooper et al. proposed a method for learning from data with perfect nondeterministic intervention [2]. They introduced an extra variable, M_i, for the role of the experimenter in intervention on variable X_i. $M_i = 0$ when X_i is passively observed. $M_i = k$ (from 1 to r_i), when the experimenter wishes to intervene X_i at value k. It helps to see nondeterministic intervention as a general case of deterministic intervention to get the difference between them. In case of a deterministic intervention under this description (with added M_i), $P(X_i \mid Pa(X_i), M_i = k, \theta_G) = I(X_i = k)$, while for nondeterministic intervention, it is not so. The authors mentioned that, adding M_is to the set of variables and carrying out the same analysis as in the deterministic intervention would derive the CBN. The addition of M_i will require information about whether experimenter wished to intervene or not, for each data instance.

Korb et al. generalized this model in terms of effectiveness [8]. They formulated the uncertainty of the success of intervention through a latent indicator R_i, where $R_i = 1$ when the intervention is successful and $R_i = 0$ when it is not. Therefore, $P(X_i \mid Pa(X_i), M_i = k, \theta_G)$ becomes a mixture model. Under perfect nondeterministic intervention, it is described by $P(R_i = 0)P(X_i \mid Pa(X_i), \theta_G) + P(R_i = 1)I(X_i = k)$.

4.2 Learning from Data with Imperfect Intervention

Tian et al. termed an imperfect intervention as a *mechanism change* [14].

Definition (*Mechanism change*). A Mechanism change is a transformation of causal model $M =< G, \theta_G >$at a variable X_i to a new model $M_{X_i} =< G, \theta'_G >$, where $\theta'_G = \Psi'_i \cup (\theta_G \setminus \Psi_i)$ and Ψ'_i is a set of parameters having different values than in Ψ_i.

Hence, we set,

$$p(X_i \mid Pa(X_i), M_i = 0, \theta_G) = p(X_i \mid Pa(X_i), \Psi_i)$$
$$\text{and} \quad p(X_i \mid Pa(X_i), M_i \neq 0, \theta_G) = p(X_i \mid Pa(X_i), \Psi'_i)$$

With these assumptions, the Equation 3.3 can be partitioned into cases where X_i is passively observed and cases where X_i is intervened [4].

$P(G, D)$

$$
\begin{aligned}
=&P(G) \int \prod_{h:M_{ih}=0} \prod_{i=1}^{n} P(X_i = d_{ih} \mid Pa(X_i) = d_{Pa(X_i)h}, \Psi_i) \\
&\times P(\Psi_i \mid G) d\Psi_i \\
&\times \int \prod_{h:M_{ih}\neq 0} \prod_{i=1}^{n} P(X_i = d_{ih} \mid Pa(X_i) = d_{Pa(X_i)h}, \Psi'_i) \\
&\times P(\Psi'_i \mid G) d\Psi'_i
\end{aligned}
$$

Note that, the formulation of non-deterministic intervention by Korb et al. is also applicable on imperfect nondeterministic interventions.

5 Conclusion

This paper serves as a short summary of the recently flourished research area of learning CBNs. We have reviewed the conceptual background, developed a suitable taxonomy and summarized some learning methods along the divisions.

We have noted that, initial efforts only used observational data for causal discovery. However, with observational data alone, these approaches can only identify the structure up to I-equivalence. In the absence of interventional data, structure learning of CBNs is just the same as structure learning of BNs.

The use of interventional data has enabled further disambiguation of structures. Since CBNs were formalized, many works have been done on learning causal discovery from different types of interventional data. The difference in interventional data arises from different intervention setup in real world problems. The types of interventions have become subsequently challenging, from perfect to imperfect, from deterministic to non-deterministic. The trend of learning CBNs under increasingly difficult interventional setup continues, for example, Eaton et al. has recently worked on uncertain interventions where the effects of the intervention are unknown [4].

Learning CBNs has recently been incorporated with active learning, where the learning algorithm actively seeks appropriate data points, in this case interventional data instances, in order to optimize performance [15]. A great deal of research is to be done to extend the idea of CBNs and integrate it to other learning problems where causality is concerned.

References

[1] Gregory F. Cooper and Edward Herskovits. A bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 9:309–347, 1992.

[2] Gregory F. Cooper and Changwon Yoo. Causal discovery from a mixture of experimental and observational data. In Kathryn B. Laskey and Henri Prade, editors, *UAI*, pages 116–125. Morgan Kaufmann, 1999.

[3] Ramon López de Mántaras and David Poole, editors. *UAI '94: Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence, July 29-31, 1994, Seattle, Washington, USA*. Morgan Kaufmann, 1994.

[4] Daniel Eaton and Kevin Murphy. Exact bayesian structure learning from uncertain interventions. In *AI & Statistics*, 2007.

[5] David Heckerman. A bayesian approach to learning causal networks. In Philippe Besnard and Steve Hanks, editors, *UAI*, pages 285–295. Morgan Kaufmann, 1995.

[6] David Heckerman, Dan Geiger, and David Maxwell Chickering. Learning bayesian networks: The combination of knowledge and statistical data. In de Mántaras and Poole [3], pages 293–301.

[7] David Heckerman and Ross D. Shachter. A decision-based view of causality. In de Mántaras and Poole [3], pages 302–310.

[8] Kevin B. Korb, Lucas R. Hope, Ann E. Nicholson, and Karl Axnick. Varieties of causal intervention. In Chengqi Zhang, Hans W. Guesgen, and Wai-Kiang Yeap, editors, *PRICAI*, volume 3157 of *Lecture Notes in Computer Science*, pages 322–331. Springer, 2004.

[9] J. Pearl. *Causality*. Causality, by Judea Pearl, pp.˜400.˜ISBN 0521773628.˜Cambridge, UK: Cambridge University Press, March 2000., March 2000.

[10] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference*. Morgan Kaufmann, September 1988.

[11] Ramon Sangüesa and Ulises Cortés. Learning causal networks from data: A survey and a new algorithm for recovering possibilistic causal networks. *AI Commun.*, 10(1):31–61, 1997.

[12] Peter Spirtes. Detecting causal relations in the presence of unmeasured variables. In Bruce D'Ambrosio and Philippe Smets, editors, *UAI*, pages 392–397. Morgan Kaufmann, 1991.

[13] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search (Lecture Notes in Statistics)*. Springer-Verlag, 1993.

[14] Jin Tian and Judea Pearl. Causal discovery from changes. In Jack S. Breese and Daphne Koller, editors, *UAI*, pages 512–521. Morgan Kaufmann, 2001.

[15] Simon Tong and Daphne Koller. Active learning for structure in bayesian networks. In Bernhard Nebel, editor, *IJCAI*, pages 863–869. Morgan Kaufmann, 2001.

[16] L.k. West and W.F. Agbola. Causality links between asset prices and cash rate in australia. *International Journal of Applied Econometrics and Quantitative Studies*, 2(3):69–86, 2005.