FUNCTIONAL CONNECTIVITY MODELLING IN FMRI BASED ON CAUSAL NETWORKS

F.F. Deleus, P.A. De Mazière, & M.M. Van Hulle

Laboratorium voor Neuro- en Psychofysiologie
Katholieke Universiteit Leuven
Campus Gasthuisberg
Herestraat 49
B-3000 Leuven, BELGIUM
Tel.: +32 16 34 59 61 Fax.: +32 16 34 59 60
Email: {filip, patrick, marc}@neuro.kuleuven.ac.be

Abstract.

We apply the principle of causal networks to develop a new tool for connectivity analysis in functional Magnetic Resonance Imaging (fMRI). The connections between active brain regions are modelled as causal relationships in a causal network. The causal networks are based on the notion of *d-separation* **in a graph-theoretic context or, equivalently, on the notion of** *conditional independence* **in a statistical context. Since relationships between brain regions are believed to be non-linear in nature [1], we express the conditional dependencies between the brain regions' activities in terms of conditional mutual information. The density estimates needed for computing the conditional mutual information are obtained with topographic maps, trained with the kernel-based Maximum Entropy Rule (kMER).**

INTRODUCTION

Network modelling techniques have been used in a variety of application domains. In functional neuroimaging, network modelling has been used for modelling connections between active brain regions. Functional network modelling, as it is called, has first been applied to Positron Emission Tomography (PET) by McIntosh and co-workers [2, 3], and later also to functional Magnetic Resonance Imaging (fMRI)[1] [4]. The technique used in both cases is

---

[1]FMRI is a non-invasive technique for measuring metabolic-related brain activity given a certain task or sequence of tasks.

Structural Equation Modelling (SEM). In SEM, a linear system of equations is used to describe the interactions between the nodes of a network model, and the free parameters are optimised in such a manner that the covariance pattern of the recorded signals is sufficiently closely modelled [5]. Alternatives to SEM have been introduced in fMRI by Büchel and Friston: in one case, the relationships are obtained with a regression technique (Kalman filtering and a fixed smoothing process) [6], and in another case, with a higher-order convolution technique (using Volterra kernels) [1] (for an overview, see [7]). Here, we will introduce a still different and new approach to functional network modelling in neuroimaging. The technique relies on the principle of causal networks which are instantiations of a Directed Acyclic Graph (DAG) [8]. The causal network is then adopted as our functional network model. The advantage of this technique is that it allows for non-linear connections in the functional network model.

The DAGs were originally conceived for discrete variables, and have led to the well-known Bayesian networks [9]. The main principle behind them is a test for *conditional independence* between the network nodes' activities. The correspondence between graphical causal models (here, DAGs) and linear causal models, for continuous variables, (*e.g.*, Structural Equations Models, SEM [5]) is investigated in [10]. However, by applying a non-linear conditional independence test, we can go beyond the linearity restriction of SEM, and discover causal *non-linear* relationships.

The article is organised as follows. First, we introduce DAGs and review the conditions required for a causal interpretation of a DAG. We then describe the basic algorithm for obtaining the causal DAG, called SGS-algorithm, after their inventors Spirtes, Glymour and Scheines. In the next section, we briefly describe the conditional dependencies between the brain regions' activities in terms of conditional mutual information. The density estimates needed for computing the latter are determined with kernel-based topographic maps, trained with the kernel-based Maximum Entropy Rule (kMER) [11, 12]. Next, we apply our network modelling to a human fMRI data set [13] obtained in a block design heading study, and discuss the causal non-linear relationships between the obtained brain regions. Finally, we summarise our results and indicate some open ends which require further research.

CAUSAL NETWORKS

DAG. We will consider a functional network model of fMRI activations as an instantiation of a DAG [8]. We first need a number of definitions. A graph consists of nodes joined by edges. A DAG is then defined as a set of nodes[2] **X** and a set of directed edges **E**, in such a way that the DAG is loopfree. To

---

[2]We denote sets of variables by boldface capital letters $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ and individual variables by capital letters such as X, Y, Z.

each node of a DAG corresponds a variable, *i.c.*, the recorded fMRI signal. When a *directed edge* exists from nodes X to Y, then X is a *parent* of Y and Y is a child of X. A *descendant* of node X is any node Y for which a *directed path* exists from X to Y.
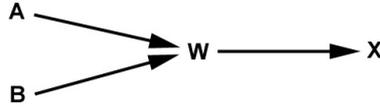
Figure 1: Causal Markov Condition for a simple network. See text for explication.

Causal Markov Condition. The DAG G represents the conditional independence properties of the probability distribution of its nodes $P(\mathbf{X})$. For random vectors X, Y and Z, X and Y are conditionally independent given Z, denoted as $X \perp Y \mid Z$, if X and Y are independent after removing the effects caused by Z. The *Causal Markov Condition* (CMC) lists the *independence* relationships that are specified by the graph. For example, consider Fig. 1. The CMC states that variable X is independent of its grandparents A and B, given its parent W: we can say that A and X are blocked by W and B and X are blocked by W. Indeed, once one knows the value of W, one can predict the value of X using W. In the literature on causality, it is said that $\{A, B\}$ and X are *d-separated* by W. The notion of d-separation is a graph theoretic criterion for detecting conditional independencies. More generally, the CMC states that each variable is independent of its grandparents, given its parent variables (for more information, see [9]). Formally, for a given graph G and node X, let *Parents* (X) be the set of parents of X, and *Descendants* (X) the set of descendants of X, then the DAG G over a set $\mathbf{X}$ and a probability distribution $P(\mathbf{X})$ satisfies the CMC if and only if for every W in $\mathbf{X}$:

$$X \perp \mathbf{X} \backslash (\mathbf{Descendants}(X) \cup \mathbf{Parents}(X)) \mid \mathbf{Parents}(X)$$

Causal Faithfulness Condition. Conclusions about the structure of the DAG of an observed sample (*i.c.*, an fMRI data set) can only be reached if a second assumption is adopted in the analysis: in general, a probability distribution P on a graph G satisfying the CMC may include other independence relations besides those entailed by the CMC. One such assumption is the *Causal Faithfulness Condition* (CFC) specifying the *dependence* relationships in the graph, stating that variables are independent only when their independence is implied by the CMC. Both CMC and CFC connect probability with causality and allow for a causal interpretation of the network.

SGS-algorithm. The SGS-algorithm is one of the basic algorithms for obtaining a causal DAG from a data set. The algorithm, as well as its correctness proof with respect to the CMC and CFC, are described in [14, 15]. The algorithm assumes causal sufficiency, which means that there are measurements available for each node (*i.e.*, no hidden variables). Since, in the field of functional neuroimaging, one is interested in a connectivity analysis given the recorded activations, this algorithm suffices for our purposes. The original SGS-algorithm is formulated using the notion of *d-separation* [14]. However, as was done in [15], since in the discovery phase of a DAG, *d-separation* of nodes is determined by conditional independence, we will write "conditional independence" instead of "d-separation between X and Y".

Algorithm

A) Compose the complete undirected graph G for the set of variables $\mathbf{X}$.

B) For each pair of nodes X and Y, if there exists a subset $\mathbf{S}$ of $\mathbf{X}\backslash\{X,Y\}$ such that $X \perp Y \mid \mathbf{S}$, remove the edge between X and Y from G.

C) Let K be the undirected graph resulting from step B). For each triple of nodes X, Y, and Z such that X and Y are linked and also Y and Z (written as $X - Y - Z$), but the pair X and Z is not, add directions to the edges of $X - Y - Z$ as follows: $X \rightarrow Y \leftarrow Z$, if and only if there is no subset $\mathbf{S}$ of $\mathbf{X}\backslash\{X,Z\}$, but which contains Y, such that $X \perp Z \mid \mathbf{S}$.

D) In the partially directed graph resulting from step C), add as many directions to the remaining undirected edges as possible subject to two conditions: (i) the added direction should not create a new *v-structure*, *i.e.*, two converging arrows whose tails are not connected by an arrow, *e.g.*, $X \rightarrow Y \leftarrow Z$, and (ii) the added direction should not create a directed cycle.

Step D) of the algorithm can be implemented in several ways. We refer to [16] where four simple rules are described to implement this step and which were also used in our simulations.

CONDITIONAL INDEPENDENCE

Mutual Information. As a measure of statistical independence between two variables X and Y, we use mutual information $\mathrm{MI}(X,Y)$. Indeed, MI can be considered as a generalised correlation coefficient, generalised with respect to the nature of the relationship between the variables, thus not only linear relationships, as in the usual correlation coefficient, r. Mutual information can be defined in terms of differential entropy. Let $\mathbf{X}$ be a random vector with density $p(\mathbf{X})$, the differential entropy is defined as:

$$\mathrm{H}(\mathbf{X}) = -\int_X p(\mathbf{X})log_2(p(\mathbf{X}))d\mathbf{X}.$$

The mutual information between two variables X and Y of $\mathbf{X}$ is then:

$$\mathrm{MI}(X,Y) = \mathrm{H}(X) + \mathrm{H}(Y) - \mathrm{H}(X,Y). \tag{1}$$

Topographic maps and density estimation. To estimate the marginal and joint densities of Eq. (1), we use the kernel-based Maximum Entropy Rule (kMER) which develops a topographic map for which each neuron i possesses a radially-symmetric Gaussian kernel in the data space, centered at w_i and with radius σ_i. A neuron is only activated when it falls in the hypersphere with center w_i and radius σ_i. The learning rule updates the weights and the radii so that a topology-preserving map is developed in which all neurons have the same activation probabilities. These kernels can then be used for density estimation [12]:

$$p_{\rho_s}(\mathbf{v}) = \frac{1}{N} \sum_{i=1}^{N} \frac{\exp\left(-\frac{\|\mathbf{v}-w_i\|^2}{2(\rho_s \sigma_i)^2}\right)}{(\sqrt{2\pi}\rho_s \sigma_i)^d}, \tag{2}$$

in which N denotes the number of kernels, and ρ_s a common factor which controls the smoothness of the resulting density estimate. We obtain an optimal value for the parameter ρ_s by maximum likelihood estimation, *i.e.*, by adjusting ρ_s so that the data has maximum probability to have been drawn from the true input density. This is obtained by using the following formula with M denoting the number of data points.

$$\rho_s = \arg\min_{\rho_s} \left(-\sum_{j=1}^{M} \log\left(p_{\rho_s}(v_j)\right)\right).$$

Conditional density estimation. We need to estimate the $d_{\mathbf{X}}$-dimensional joint conditional density of a set of $d_{\mathbf{X}}$ variables $\mathbf{X} = \{X_1, X_2, ..., X_{d_{\mathbf{X}}}\}$, given another set of $d_{\mathbf{C}}$-dimensional variables $\mathbf{C}$ which are fixed at a particular vector of values γ. We first put a $d_{\mathbf{C}}$-dimensional kernel at position γ and with width $\sigma_{\mathbf{C}}$, and evaluate the N density kernels located at positions $\mathbf{w}_i$, $\forall i$, by only taking into account the components of $\mathbf{w}_i$ which are directed along the $d_{\mathbf{C}}$ dimensions defined by $\mathbf{C}$. This is denoted by $\mathbf{w}_i^{\mathbf{C}}$. In this way, we obtain a weighted contribution for each density kernel. We then introduce these weights into the density estimate as described in Eq. (2). This leads to the following definition of conditional density estimation:

$$p_{\rho_s}(\mathbf{v}^{\mathbf{X}} \mid \mathbf{C} = \gamma) = \sum_{i=1}^{N} \left(\frac{\exp\left(-\frac{\|\gamma - \mathbf{w}_i^{\mathbf{C}}\|^2}{2(\sigma_{\mathbf{C}})^2}\right)}{\sum_{j=1}^{N} \exp\left(-\frac{\|\gamma - \mathbf{w}_j^{\mathbf{C}}\|^2}{2(\sigma_{\mathbf{C}})^2}\right)} \cdot \frac{\exp\left(-\frac{\|\mathbf{v}^{\mathbf{X}} - \mathbf{w}_i^{\mathbf{X}}\|^2}{2(\rho_s \sigma_i^X)^2}\right)}{(\sqrt{2\pi}\rho_s \sigma_i^X)^{d_{\mathbf{X}}}}\right)$$

The conditional entropy of $\mathbf{X}$ given $\mathbf{C}$, *i.e.*, H($\mathbf{X} \mid \mathbf{C}$), is defined as the expected value of H($\mathbf{X} \mid \mathbf{C} = \gamma$) over all different $\gamma' s$. This expected value can

be approximated as a sample mean:

$$H(\mathbf{X} \mid \mathbf{C}) \approx \frac{1}{M} \sum_{j=1}^{M} H(\mathbf{X} \mid \mathbf{C} = \gamma_j)$$

$$H(\mathbf{X} \mid \mathbf{C} = \gamma_j) = -\int_{\mathbf{X}} p_{\rho_s}(\mathbf{v}^{\mathbf{X}} \mid \mathbf{C} = \gamma_j) \log_2 \left(p_{\rho_s}(\mathbf{v}^{\mathbf{X}} \mid \mathbf{C} = \gamma_j) \right) d\mathbf{v}^{\mathbf{X}}$$

Analogous to our definition of conditional entropy as an average over $\mathbf{C}$, we also define the conditional mutual information between two variables X and Y, given a set of variables $\mathbf{C}$, *i.e.*, $\mathrm{MI}(X, Y \mid \mathbf{C})$, as an average over $\mathbf{C}$:

$$\mathrm{MI}(X, Y \mid \mathbf{C}) \approx \frac{1}{M} \sum_{j=1}^{M} \mathrm{MI}(X, Y \mid \mathbf{C} = \gamma_j)$$

$$\mathrm{MI}(X, Y \mid \mathbf{C} = \gamma_j) = H(X \mid \mathbf{C} = \gamma_j) + H(Y \mid \mathbf{C} = \gamma_j) - H(X, Y \mid \mathbf{C} = \gamma_j)$$

As stated above, we use conditional mutual information to decide about the presence of conditional independence relations, and hence, the presence of a link in the functional network (thus, purely the topology). However, as opposed to the usual correlation coefficient, r, mutual information does not take values between -1 and 1. We propose a normalisation of the MI result between 0 and 1 by observing that the differential entropy $H(X)$ is an upper bound for $\mathrm{MI}(X, Y)$, and 0 the lower bound:

$$\mathrm{MI}_{norm}(X, Y) = 2 \frac{\mathrm{MI}(X, Y)}{H(X) + H(Y)}$$

If X and Y are independent, then $\mathrm{MI}(X, Y) = 0$ and also $\mathrm{MI}_{norm}(X, Y) = 0$. If X and Y are maximally dependent, *i.e.*, when $X=Y$, then $\mathrm{MI}(X, Y)$ can be written as $\mathrm{MI}(X, X)$ which is equal to $H(X)$ and, hence, $\mathrm{MI}_{norm}(X, Y)$ becomes equal to 1. Evidently, this normalisation can also be used for evaluating the conditional mutual information:

$$\mathrm{MI}_{norm}(X, Y \mid \mathbf{C}) = 2 \frac{\mathrm{MI}(X, Y \mid \mathbf{C})}{H(X \mid \mathbf{C}) + H(Y \mid \mathbf{C})}$$

EXPERIMENT

In order to demonstrate our new network modelling approach, we consider the human fMRI data set described in [13], namely, a block design heading study. Human subjects had to judge the heading of a ground plane optic flow pattern. From the observed fMRI activations, and the anatomical evidence, Peuskens and co-workers hypothesise that the network specifically involved in heading consists of two motion sensitive areas: human MT/V5+, including an

inferior satellite, and dorsal intraparietal sulcus area (DIPSM/L), predominantly in the right hemisphere, plus a dorsal premotor region bilaterally. However, they did not apply a network modelling technique.

We will now use the experimental design referenced in [13] as fMRI3 for network modelling. In this fMRI experiment, five tasks were considered: one fixation task, which is used as a reference, two heading tasks, and two dimming tasks, which did not contain any heading information (control tasks). The difference in brain activations between the heading and the control tasks should only be due to the heading component in the stimulus.

For the sake of simplicity, we use data from a single subject only. The raw fMRI time series are corrected for head movement and are realigned to the Talairach-Tournoux space, using SPM99 (Wellcome Department of Cognitive Neurology, London), a well-known package for fMRI analyses. Furthermore, we restrict our network modelling to the fMRI time series, the coordinates of which are listed in Table 3 in [13] and that are part of the right brain hemisphere. The original fMRI time series are linearly detrended to remove any drift caused by the fMRI scanner, and uniformly scaled to fit the $[0,1]$ interval, followed by a zero-mean centering. Higher-order detrending is not performed, since it could hamper the detection of paradigm-related patterns. The fMRI time series are shown in Fig. 2.
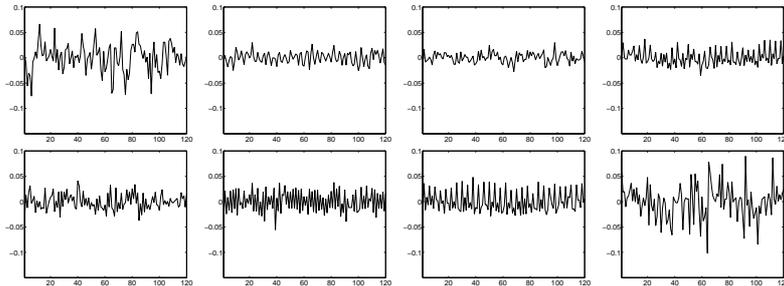
Figure 2: fMRI signals that are actively involved in the heading task. The eight fMRI signals shown, are the ones mentioned in Table 3 in Peuskens *et al.*, and pre-processed as discussed in the text. The horizontal axis represents time, while the vertical one represents the signal's amplitude.

We now apply our causal network modelling tool. The density estimation is performed with a 4×4 topographic map. When determining the conditional independence relations, we take $MI_{norm}(X, Y \mid \mathbf{C}) = 0.05$. The resulting network is shown in Fig. 3. When considering a 3×3 topographic map, only a few links are detected and a relatively large number of nodes are not connected to any other node. This is due to the fact that the map is too small to capture the probability distribution function accurately and, hence, it cannot be used for reliable conditional independence tests. The results for 4×4 and 5×5 topographic maps are similar, indicating that the density estimate for the 4×4 case is accurate enough for the current data set.
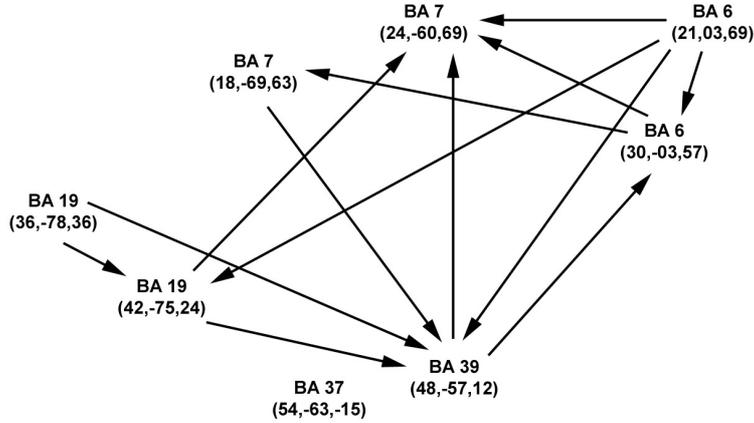
Figure 3: Network as found by applying the SGS algorithm to the fMRI signals.

DISCUSSION

The nodes in Fig. 3 represents the brain areas that are assumed to be responsible for the judgement of heading. This result corresponds rather well to what is known about this part of the cortex, as well as the hypothesis formulated by Peuskens and co-workers [13].

First, we briefly discuss the name and function of the brain areas indicated in Fig. 3. Two nodes labelled with BA 7 represent the dorsal intraparietal sulcus, medial and lateral (DIPSM/L), which is considered to be important in the integration of sensory inputs with the decision-making apparatus, including the focusing of attention and the initiation of purposeful body movements [17]. The node labelled with BA 39 represents an area that is slightly anterior (*i.e.*, towards the forehead) to the middle temporal area (hMT/V5+). Areas BA 37 and BA 19, and to a lesser extent also BA 39, all belong to hMT/V5+, which is part of the cortical visual system. Finally, BA 6 is the premotor area which is responsible for more complex body movements [17].

We now discuss the plausibility of the network connections shown in Fig. 3. The network indicates that, prior to a body movement (BA 6), input is received from the visual system (BAs 39 and 19) and from BA 7, which integrates sensory and decision-making information. The functionality of the latter also explains why arrows, *e.g.*, originating from BA 6, are pointing to it, since, when one is actually performing a body movement, feedback will have to be received by BA 7. Note that BA 37 seems to have no interaction at all with the other brain areas, and that BA 6 even projects directly to BA 19, however, we note that there is not much known about this connection (Peuskens & Orban, personal communications).

CONCLUSION AND FUTURE RESEARCH

As demonstrated, our network modelling approach offers a quantitative way for studying cortical networks. However, to validate our approach, we should consider more cases for which the obtained networks can be compared with the known anatomy. Finally, concerning the tool itself, although we have only considered directed *acyclic* graphs, the assumptions for causal discovery can also be investigated with directed *cyclic* graphs (DCG) [18, 19]. We intend to incorporate conditional mutual information tests in DCGs, and apply DCGs to fMRI data in order to model recurrent connections between brain regions.

REFERENCES

[1] Friston, K.J., & Büchel, C. (2000). Attentional modulation of effective connectivity from V2 to V5/MT in humans. *PNAS* , vol. 97, **13**, 7591-7596.

[2] Friston, K.J., Frith, C.D., and Frackowiak, R.S.J. (1993). Time-Dependent Changes in Effective Connectivity Measured with PET. *Human Brain Mapping*, **1**, 69-79.

[3] McIntosh, A.R., & Gonzalez-Lima, F. (1994). Structural Equation Modeling and Its Application to Network Analysis in Functional Brain Imaging. *Human Brain Mapping*, **2**, 2-22.

[4] Büchel, C. & Friston, K.J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex*, **7**, 768-778.

[5] Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, **5**, 161-215.

[6] Büchel, C., & Friston, K.J. (1998). Dynamic Changes in Effective Connectivity Characterized by Variable Parameter Regression and Kalman Filtering. *Human Brain Mapping*, **6**, 403-408.

[7] Petersson, K.M, Nichols, T.E., Poline, J.-B., and Holmes, A.P. (1999a). Statistical limitations in functional neuroimaging I. Non-inferential methods and statistical models. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **354** (1387), 1239-1260.

[8] Jordan, M.I., and Sejnowski, T.J.(Eds.) (2001). *Graphical Models: Foundations of Neural Computation*. Cambridge MA: MIT Press.

[9] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Francisco, CA: Morgan Kaufmann.

[10] Pearl, J. (1997). Graphs, Causality and Structural Equation Models. Technical Report R-253, Cognitive Science Laboratory, UCLA.

[11] Van Hulle, M.M. (1998). Kernel-based equiprobabilistic topographic map formation. *Neural Computat.*, **10**(7), 1847-1871.

[12] Van Hulle, M.M. (2000). *Faithful Representations and topographic maps: From distortion- to information-based self-organization*. New York: Wiley.

[13] Peuskens, H., Sunaert, S., Dupont, P, Van Hecke, P, and Orban, G.A. (2001). Human brain regions involved in heading estimation. *J.Neuroscience*, **21**(7), 2451-2461.

[14] Spirtes, P., Glymour, C., and Scheines, R. (1990). Causality from Probability. *Evolving knowledge in Natural Science and Artificial Intelligence* (J.Tiles *et al.*, Eds.), pp. 181-199. London: Pitman.

[15] Spirtes, P., Glymour, C., and Scheines, R. (1993). *Causation, Prediction, and Search*. New York: Springer-Verlag.

[16] Verma, T., and Pearl, J. (1992). An algorithm for deciding if a set of observed independencies has a causal explanation. In *Proc. 8th Conf. Uncertainty in Artificial Intelligence* (Dubois, Wellman, D'Ambrosio, and Smets, Eds.), pp. 323-330. Stanford: Morgan Kaufmann.

[17] Greenstein, B., and Greenstein, A. (1999). *Color atlas of neuroscience*. Stuttgart - New York: Thieme.

[18] Spirtes, P. (1995). Directed Cyclic Graphical Representation of Feedback Models. In *Proc. 11th Conf. Uncertainty in Artificial Intelligence* (Besnard and Hanks, Eds.). San Mateo: Morgan Kaufmann Publ., Inc.

[19] Richardson, T. and Spirtes, P. (1999). Automated discovery of linear feedback models. In *Computation, Causation, and Discovery* (Glymour, and Cooper, Eds.), pp. 253-304. Cambridge, MA: MIT Press.