Abstract

The use of acyclic, directed graphs (often called 'DAG's) to simultaneously represent causal hypotheses and to encode independence and conditional independence constraints associated with those hypotheses has proved fruitful in the construction of expert systems, in the development of efficient updating algorithms (Pearl, 1988, Lauritzen et al. 1988), and in inferring causal structure (Pearl and Verma, 1991; Cooper and Herskovits 1992; Spirtes, Glymour and Scheines, 1993).

In section 1 I will survey a number of extensions of the DAG framework based on directed graphs and chain graphs (Lauritzen and Wermuth 1989; Frydenberg 1990; Koster, Madigan and Perlman 1996). Those based on directed graphs include models based on directed cyclic and acyclic graphs, possibly including latent variables and/or selection bias (Pearl, 1988; Spirtes, Glymour and Scheines 1993; Spirtes 1995; Spirtes, Meek, and Richardson 1995; Richardson 1996a, 1996b; Koster 1996; Pearl and Dechter 1996; Cox and Wermuth, 1996).

In section 2 I state two properties, motivated by causal and spatial conditional independencies entailed by a graphical model might satisfy. I proceed to show that the sets of independencies entailed by (i) an undirected graph via separation, and (ii) a (cyclic or acyclic) directed graph (possibly with latent and/or selection variables) via d-separation, satisfy both properties. By contrast neither of these properties, in general, will hold in a chain graph under the Lauritzen-Wermuth-Frydenberg (LWF) interpretation. One property holds for chain graphs under the Andersson-Madigan-Perlman (AMP) interpretation, the other does not. The examination of these properties and others like them may provide insight into the current vigorous debate concerning the applicability of chain graphs under different global Markov properties.

1. Graphs and Probability Distributions

An undirected graph $U$ is an ordered pair $\langle V, U \rangle$, where $V$ is a set of vertices and $U$ is a set of undirected edges $X\rightarrow Y$ between vertices. Similarly, a directed graph $DG$ is an ordered pair $\langle V, D \rangle$ where $D$ is a set of directed edges $X\rightarrow Y$ between vertices. A directed cycle consists of a sequence of edges $X_1\rightarrow X_2\ldots\rightarrow X_n\rightarrow X_1$ (n≥2). If a directed graph $DG$ contains no directed cycles it is said to be acyclic; otherwise it is cyclic. An edge $X\rightarrow Y$ is said to be out of $X$ and into $Y$; $X$ and $Y$ are the endpoints of the edge. Note that if cycles are permitted there may be more than one edge between a given pair of vertices e.g. $X\leftrightarrow Y$.

---

1 I thank P. Spirtes, C. Glymour, D. Madigan, M. Perlman and J. Besag for helpful conversations. Research for this paper was supported by the Office of Naval Research through contract number N00014-93-1-0568.

2 Bold face (X) indicate sets; plain face (X) indicates individual elements; italics (U) indicates a graph or a path.
I will consider directed graphs (cyclic or acyclic) in which \( V \) is partitioned into three disjoint sets \( O \) (Observed), \( S \) (Selection) and \( L \) (Latent), written \( DG(O,S,L) \). The interpretation of this definition is that \( G \) represents a causal mechanism, the subset of the variables observed, \( S \) represents a set of variables which, due to the nature of the mechanism selecting the sample, are conditioned on in the subpopulation from which the sample is drawn, the variables not observed and for this reason are called 'latent'.

A mixed graph contains both directed and undirected edges. A partially directed cycle in a mixed graph \( G \) is a sequence of \( n \) distinct \( X \) such that \( (1 \leq i \leq n) \) either \( X_i \rightarrow X_{i+1} \) or \( X_i \leftarrow X_{i+1} \), and \( (1 \leq i \leq n) \) such that \( X_j \rightarrow X_{j+1} \).

A chain graph \( CG \) is a mixed graph in which there are no partially directed cycles. Koster (1996) considers a class of reciprocal graphs containing directed and undirected edges in which partially directed cycles are allowed. I do not consider such graphs separately here, though many comments which apply to LWF chain graphs apply also to reciprocal graphs since the former are a subclass of the latter.

To make clear which kind of graph is being referred to I will fuse undirected graph \( UG \) for directed graphs, \( AG \) for acyclic directed graphs, \( CG \) for chain graphs, and \( G \) to denote a graph which may be any one of these.

A path between \( Xin \) and \( Graph \) of whatever type) consists of a sequence \( E_{1}, \ldots, E_{n} \) such that there exists a sequence of distinct vertices \( X_{1}, \ldots, X_{n+1}, Y \) where \( E_{i} \) has endpoints \( i \) and \( i+1 \) (\( 1 \leq i \leq n \)), i.e. \( E_{i} \) is \( X_{i} \rightarrow X_{i+1} \) or \( X_{i} \leftarrow X_{i+1} \). A directed path from \( X \) to \( Y \) is a path of the form \( X \rightarrow \ldots \rightarrow Y \).

1.2 Global Markov Properties Associated with Graphs

A Global Markov Property associates a set of conditional independence relations with a graph \( G \). In an undirected graph \( UG \), for disjoint sets of vertex \( X, Y \) and \( Z \), \( (Z \) may be empty), if there is no path from a variable \( X \in X \), to a variable \( Y \in Y \), that does not include some variable in \( Z \), then \( X \) and \( Y \) are said to be separated by \( Z \).

Undirected Global Markov Property

\[
\text{\textit{Undirected Global Markov Property}} \quad \downarrow_{s} \quad Y | Z \quad \text{if} \quad X \text{and} \ Y \text{are separated by} \ Z \text{in} \ UG. \]

In a graph \( G \), \( X \) is parent of \( Y \), (and \( Y \) is child of \( X \)) if there is an edge \( X \rightarrow Y \) in \( G \). \( X \) is an ancestor of \( Y \) (and \( Y \) is descendant of \( X \)) if \( X=Y \), or there is a directed path \( X \rightarrow \ldots \rightarrow Y \) from \( X \).

---

3 Note that we use the terms 'variable' and 'vertex' interchangeably.
4 Path is defined here as a sequence of edges, rather than vertices; in a cyclic graph a sequence of vertices does not in general define a unique path, since there may be more than one edge between a given pair of vertices.
5 Often global Markov conditions are introduced as a means for deriving the consequences of a set of local Markov conditions. Here I merely define the Global property in terms of the relevant graphical criterion.
6 \( X \downarrow_{s} Y | Z \) means that \( X \) is independent of \( Y \) given \( Z \); if \( Z=\emptyset \), the abbreviation \( X \downarrow Y \) is used. \( X \downarrow Y \) and/or \( Z \) are singleton sets \( \{V\} \), then brackets are omitted e.g. \( \forall \downarrow Y | Z \), instead of \( \{\forall\} Y | Z \).
to \( Y \). A pair of consecutive edges on a path in \( G \) are said to collide at vertex \( A \), if both edges are into \( A \), (i.e. \( \rightarrow A \rightarrow \)), in this case \( A \) is called a non-collider on \( P \), otherwise \( A \) is a non-collider on \( P \).

For distinct vertices \( X \) and \( Y \), and \( Z \in V \setminus \{X,Y\} \), a path \( P \) between \( X \) and \( Y \) gives to every collider in the path (if any) are such that \( Y \) is between \( X \) and \( W \) and \( Y \) is neither an ancestor of a non-collider on \( P \) is in \( Z \). Disjoint sets \( X \) and \( Y \) are said to be \( d \)-connected given \( Z \) if there is an \( X \in X \), and \( Z \in Y \), and \( Y \), where there is such a path which \( d \)-connects \( X \) and \( Y \) given \( Z \) (see Pearl, 1988).

**Global Markov for Property Graphs; \( \models_{DS} d \)-separation**

\[
\models_{DS} \nmid X | Y \mid Z \quad \text{if and only if} \quad X \nmid Y | Z \quad \text{in} \quad DG
\]

For \( DG(O,S,L) \), and disjoint subseb \( S \subset Y \subset Z \subset O \) we define:

\[
DG(O,S,L)|\models_{DS} \nmid Y \mid Z \quad \text{if and only if} \quad DG \models_{DS} \nmid Y \mid Z \cup S
\]

Since, under the interpretation \( DG(O,S,L) \), the only observed variables \( O \) are not observe conditional independence relations involving variables \( I \). Similarly, since samples are drawn from a subpopulation in which all variables are conditioned upon in every conditional independence relation we observe to hold in the original graph, this definition gives the set of conditional independencies in the observed distribution \( P(\cdot|O,S) \). (See Spirtes and Richardson, this volume; Spirtes, Meek and Richardson, 1996; Cox and Wermuth, 1996.)

Two different Global Markov properties have been proposed for Chain Graphs. In both definitions a conditional independence relation is entailed if sets \( X \) and \( Y \) are separated by \( Z \) in an undirected graph whose edges are a superset of those in the original chain graph.

A vertex \( V \) in a chain graph is said to be an \( \nmid \) if there is a path from \( V \) to some \( W \in \mathcal{W} \) in which all directed edges \( (X \rightarrow Y) \) on the path (if any) are such that \( Y \) is between \( X \) and \( W \) on \( P \), \( \text{Ant}(W) = \{V \mid V \text{ is anterior to } W \} \). Let \( CG(W) \) denote the induced subgraph \( CG \) obtained by removing all vertices \( \mathcal{W} \) and edges with an endpoint in \( \mathcal{W} \) and replacing all directed edges with undirected edges. \( CG \) is an induced subgraph with the following form: \( V_1 \leftarrow \ldots \leftarrow V_n \rightarrow Y \) (\( n \geq 1 \)). A complex is a moralized by adding the undirected edge \( X \rightleftharpoons Y \). Moral (\( CG \)) is the undirected graph formed by moralizing all complexes in \( CG \), and then replacing all directed edges with undirected edges.

**Lauritzen-Wermuth-Frydenberg Global (\( \models_{LWF} \) Markov Property**

\[
\models_{LWF} \nmid X | Y \mid Z \quad \text{if} \quad X \nmid Y | Z \quad \text{in} \quad CG(\text{Ant}(X \cup Y \cup Z))
\]

In a chain graph vertices \( V \) and \( W \) are said to be in a path containing only undirected edges between \( V \) and \( W \), \( \text{Con}(W) = \{V \mid V \text{ is connected to some } W \in \mathcal{W} \} \). The extended subgraph, \( \text{Ext}(CG(W)) \), has vertex set \( \text{Con}(W) \) and contains all directed edges in \( CG(W) \), and all undirected edges in \( CG \) (\( \text{Con}(W) \)). A triple of vertices \( \langle X,Y,Z \rangle \) is said to form a triple in \( CG \) if the induced subgraph \( CG(X,Y,Z) \) is either \( X \leftarrow Y \leftarrow Z \), or \( X \rightarrow Y \leftarrow Z \). A triple is augmented by adding the \( X \rightarrow Z \) edge. \( \text{Aug}(CG) \) is the undirected graph formed by augmenting all triples in \( CG \) and replacing all directed edges with undirected edges.

409
Andersson-Madigan-Perlman Global Markov Property

\[ \text{AMP} \models Y \perp Z \mid X \text{ if } X \text{ is separated from } Y \text{ by } Z \text{ in } \text{Aug(Ext(} CG(\text{Anc}(X \cup Y \cup Z)))\text{))} \]

where \( \text{Anc}(W) = \{ V \mid V \text{ is an ancestor of some } W \in W \} \).

Both LWF and AMP properties coincide with separation (d-separation) for the special case of a chain graph which is an undirected (acyclic, directed) graph. In this sense the chain graphs with either property are a generalization of both acyclic, directed graphs and undirected graphs.

Examples:

\[
\begin{align*}
\text{CG}_1 & \quad A \rightarrow C \\
& \quad B \rightarrow D
\end{align*}
\]

\[
\begin{align*}
\text{CG}_2 & \quad A \rightarrow D \\
& \quad B \rightarrow C
\end{align*}
\]

\[
\begin{align*}
\text{CG}_1 & \models_{\text{AMP}} \perp B; A \perp D \mid \{B,C\}; B \perp C \mid \{A,D\} \\
\text{CG}_2 & \models_{\text{AMP}} B; B \perp C; B \perp D \mid \{A,C\}
\end{align*}
\]

1.3 Completeness

For a given global Markov property \( R \), and graph \( G \) with vertex set \( V \), a distribution \( P \) is said to be \( G \)-Markovian \( \models_{R} X \perp Y \mid Z \) if \( X \perp Y \mid Z \) in \( P \). A given global Markov property is said to be weakly complete if for all disjoint sets \( X, Y \) and \( Z \), such that \( G \not=_{R} X \perp Y \mid Z \) there is a \( G \)-Markovian \( \models_{R} X \perp Y \mid Z \) distribution \( P \) in which \( X \perp Y \mid Z \) in \( P \). The property \( R \) is said to be strongly complete if there is a \( G \)-Markovian \( \models_{R} X \perp Y \mid Z \) distribution \( P \) in which \( G =_{R} X \perp Y \mid Z \) if and only if \( X \perp Y \mid Z \) in \( P \).

Except for the AMP property, all of the global Markov properties here are known to be weakly complete (Geiger, 1990; Frydenberg, 1990). For general directed graphs, d-separation, and for chain graphs, the LWF Markov property, have been shown to be strongly complete. (Spirtes 1995; Meek 1995; Spirtes et al. 1993; Studeny and Bouckaert, 1996.)

2 Inseparability and Related Markov Properties

In this section I will introduce two properties, motivated by spatial and causal intuitions.

Distinct vertices \( X \) and \( Y \) are inseparable \( \models_{R} X \perp Y \mid W \) in \( G \) under Markov Property \( R \) if there is no set \( W \) such that \( G =_{R} X \perp Y \mid W \). If \( X \) and \( Y \) are not inseparable \( \models_{R} X \sim Y \mid W \) in \( G \) under \( R \) then \( X \sim Y \mid W \). Let \( G^\text{in}_R \) be the undirected graph in which there is an edge \( X \rightarrow Y \) if and only if \( X \perp Y \mid \text{Anc}(X) \cup \text{Anc}(Y) \) under \( R \). Note that in accord with \( \models_{R} \) the definition of
For an undirected graph model $G_{\text{ins}}^\text{uns}$ is just the undirected graph. For an acyclic, directed graph (without latent or selection variables) under d-separation, or a chain graph under either LWF or $A^{\text{dir}}_{R_{\text{ins}}}$ is simply the undirected graph formed by replacing undirected edges. In any graphical model, if there is an edge (directed or undirected) between a pair of variables then those variables are inseparable. For undirected graphs, acyclic directed graphs, and chain graphs, inseparability is both a necessary and a sufficient condition for the existence of an edge between a pair of variables. However, in a directed graph with cycles, or in a (cyclic or acyclic) directed graph with latent and/or selection variables (recall that in $DG(O,S,L)$, we restrict ourselves to the observed conditional independencies), inseparability is not a sufficient condition for there to be an edge between a pair of variables. An inducing path between $X$ and $Y$ is a path $P$ between $X$ and $Y$ on which (i) every vertex in $O \cup S$ is a collider on $P$, and (ii) every collider is an ancestor of $X, Y, or S$.

In a directed graph $DG(O,S,L)$, variables $X,Y \in O$ are inseparable if and only if there is an inducing path between $X$ and $Y$ in $DG(O,S,L)$.

$\equiv X_0, X_1, \ldots \notin B, \ldots X_n \in Y >$ such that each consecutive pair of vertices $X_i, X_{i+1}$ in the sequence are inseparable in $G$ under $R$. Clearly $B$ will be between $X$ and $Y$ if and only if $B$ lies on a path between $X$ and $Y$ in $G$. The set of vertices between $X$ and $Y$ under property $R$ is denoted $\text{Between}_R(X,Y)$.

**Between Separated** A model $G$ is between separated if for all pairs of vertices $X, Y$ and sets $W(X,Y \in W)$: $G|_{R} \perp \! \! \! \! \perp Y | W \Rightarrow G|_{R} \perp \! \! \! \! \perp Y | W \cap \text{Between}_R(X,Y)$

It follows that if $G$ is between separated, then in order to make some (separable) pair of vertices $X$ and $Y$ conditionally independent, it is always sufficient to condition (empty) of the vertices that lie on paths between $X$ and $Y$.

$\text{Between}_R(X,Y) = \{A,B,C,E\}$ CoCon $R(X,Y) = \{A,B,C,D,E,F\}$

$P,Q,R,S,T$ are vertices not in CoCon $R(X,Y)$

---

7The notion of an inducing path was first introduced, for acyclic directed graphs with latent variables in Verma and Pearl (1990), it was subsequently extended in Spirtes, Meek and Richardson (1995).

8Inseparability is a necessary and sufficient condition for there to be an edge between a pair of variables in a Partial Ancestral Graph (PAG), (Richardson 1996a), which represents structural features common to a given Markov equivalence class of directed graphs, possibly with latent and/or selection variables.
The intuition that only vertices on paths between X and Y are independent is related to the idea, fundamental to much of graphical modelling, that if vertices are *dependent* then they should be *connected* graphically. This is a natural correspondence, present in the spatial intuition that only contiguous regions interact directly, and also in causal principles which state that if two quantities are dependent then they are connected. 

**Theorem 1**

(i) All undirected graphs \( H \) are between separated.
(ii) All directed graphs \( DG(O,S,L) \) are between \( DS \) separated.

**Proof:** We give here the proof for undirected graph models. It is easy to see that the proof carries over directly to directed graphs without selection or latent variables (i.e. \( V=O, S=L=\emptyset \)) replacing 'separated' by 'd-separated', and 'connected' by 'd-connected'. The proof for directed graphs with latent and/or selection variables is in the appendix.

Suppose, for a contradiction, \( UG \models_S \perp Y \mid W \), but \( U \not\models_S \perp Y \mid W \) \( \text{Between}_R(X,Y) \). Then there is a path in \( UG \) connecting X and Y given \( W \) \( \text{Between}_R(X,Y) \). Since this path does not connect given \( W \), it follows that there is some vertex \( V \) on \( P \), and \( V \in W \) \( \text{Between}_S(X,Y) \). But if \( V \) is on \( P \), then \( P \) constitutes a sequence of \( X_0, X_1, \ldots X_m, Y \) such that consecutive pairs of vertices are inseparable (because there is an edge between each pair of variables). Hence \( V \in \text{Between}_S(X,Y) \), which is a contradiction. \( \therefore \)

In general chain graphs are not between \( LWF \) separated or between \( AMP \) separated. This is shown by \( CG_1 \) and \( CG_2 \) in figure 1:

\( CG_1 \) is \( LWF \) separated but \( CG_1 \) is not \( AMP \) separated. For the \( AMP \) property note \( G_2 \) separated but \( G_2 \) is not \( AMP \) separated.

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]

\[ CG_2 = \amp{B \mid A, C} \]
Clearly, $\text{Between}(X, Y) \subseteq \text{CoCon}(X, Y)$, so being co-connected is a weaker requirement than being between $X$ and $Y$. Both $\text{Between}(X, Y)$ and $\text{CoCon}(X, Y)$ are sets of vertices which are topologically "in between" $X$ and $Y$ in $G^\text{Ins}_R$.

A model $G$ will be said to be co-connection $\text{R}$ determined if for all pairs of vertices $X$, $Y$ and sets $W (X, Y \not\in W)$: $G \models \bot \; Y \mid W \iff G \models \bot \; Y \mid W \cap \text{CoCon}_R(X, Y)$

This principle states that the inclusion or exclusion of vertices that are not in $\text{CoCon}_R(X, Y)$ from some set $W$ is irrelevant to whether $X$ and $Y$ are entailed to be independent given $W$.

**Theorem 2**

(i) Undirected graph models are co-connection $\text{S}$ determined.

(ii) Directed graph models possibly with latent and/or selection variables are co-connection $\text{D}$ determined.

(iii) Chain graphs are co-connection $\text{AMP}$ determined.

**Proof:** We present here the proof for undirected graphs. The proof for directed graph models is given in the appendix; For reasons of space the proof for AMP chain graphs is not included though it is quite similar to the proof for (i) and (ii).

Since $\text{Between}(X, Y) \subseteq \text{CoCon}(X, Y)$, an argument similar to that used in the proof of Theorem 1 (replacing $\text{Between}(X, Y)$ with $\text{CoCon}(X, Y)$) suffices to show that if $G \models \bot \; Y \mid W$ then $G \models \bot \; Y \mid W \cap \text{CoCon}(X, Y)$.

Conversely, $G \models \bot \; Y \mid W \cap \text{CoCon}(X, Y)$ implies that $X$ and $Y$ are separated by $W \cap \text{CoCon}(X, Y)$ in $UG$. Since $W \cap \text{CoCon}(X, Y)$, it follows that $X$ and $Y$ are separated by $W$ in $UG$.

In fact, for undirected graphs $G \models \bot \; Y \mid W \iff G \models \bot \; Y \mid W \cap \text{Between}(X, Y)$, i.e. undirected graphs could be said to be between $\text{S}$ determined.

Chain graphs are not co-connection $\text{LWF}$ determined. In $CG_1$ B and C are separable $\text{LWF}$, since $CG_1 \models \bot \; C \mid \{ A, D \}$, but $CG_1 \models \bot \; C \mid \{ D \}$. In contrast, chain graphs are co-connection $\text{AMP}$ determined.

**2.3 Discussion**

The two Markov properties presented here are based on the intuition that only vertices which, in some sense, come "between" $X$ and $Y$ should be relevant to whether or not two vertices in a graph are entailed to be independent. Both of these properties are satisfied by undirected graphs, and by all forms of directed graph model. Since neither of these properties are satisfied by chain graphs under the LWF interpretation these properties capture a qualitative difference between undirected and directed graphs, and LWF chain graphs. In this respect, at least, AMP chain graphs are dissimilar to directed and undirected graphs since chain graphs are co-connection $\text{AMP}$ determined.
Since the pioneering work of Sewall Wright (1921), models based on directed graphs have been used to model causal relations, and data generating processes. Strotz and Wold (1958), Spirtes et al. (1993) and Pearl (1995) develop a theory of intervention for directed graph models which makes it possible to calculate the effect of intervening in a system in certain ways. Directed graphs with cycles, have been used for over 50 years in econometrics, possible of representing certain kinds of feedback, or two-way interaction. Besag (1974) gives several spatial-temporal data generating processes whose limiting spatial distributions satisfy the Markov property with respect to a naturally associated undirected graph.

In contrast Cox (1993) states that chain graphs under the LWF Markov property do "not satisfy the requirement of specifying a direct mode of data generation." This statement is given additional support by the failure of LWF chain graphs to satisfy either of the properties given above. AMP chain graphs seem more compatible with a data generating co-connection AMP determined (See also Andersson et al. 1996).

Models which are co-connection determined have a very different character from those which are not; in a co-connection determined model, the inclusion or exclusion of co-connected to X and Y from some irrelevant to the question of whether X and Y are entailed to be independent given W. Given that a large class of well-understood models which can be interpreted directly as data generating processes possess this property, it would seem that a researcher would have to have a quite particular justification for using an LWF chain graph to model a given system.

3. Appendix - Proofs

In \( DG(O,S,L) \) suppose that \( U \) is a path that d-connects X and Y given \( Z \cup S \). C is a collider on \( U \), and C is not an ancestor of S. Let \( \text{length}(C,Z) = 0 \) if C is a member of \( Z \); otherwise it is the length of a shortest directed path from C to a member of \( Z \). Let \( T(U) = \{ C \mid C \) is a collider on \( U \), and C is not an ancestor of \( S \} \). Then let

\[
\text{size}(U) = |T(U)| + \sum_{C \in T(U)} \text{length}(C,Z)
\]

If \( U \) d-connects X and Y and there is no other path that d-connects X and Y given \( Z \) such that \( \text{size}(U) < \text{size}(U) \). If there is a path that d-connects X and Y given \( Z \) then there is at least one minimal d-connecting path between X and Y. In the following proofs \( U(A,B) \) denotes the subpath of \( U \) between vertices A and B.
Lemma 1: If $U$ is a minimal d-connecting path between $X$ and $Y$ given $S$ in $DG(O,S,L)$ then for each collider $\mathcal{C}$ on $U$ that is not an ancestor of $S$ there is a directed path from $\mathcal{C}$ to some vertex in $Z$, such that $P$ intersects $U$ only at $\mathcal{C}$, $D_i$ and $D_j$ do not intersect and no vertex on any path $D_i$ is in $S$.

Proof: Let $D_i$ be a shortest acyclic directed path from a collider $C_i$ on $U$ to a member of $Z$, where $C_i$ is not an ancestor of $S$. We will prove that $D_i$ does not intersect $U$ except at $C_i$ by showing that if such a point of intersection existed $U$ would not be minimal, contrary to our assumption. See the figure 4 below:

![Figure 4](image)

Next, we will show that, if minimal, $D_i$ and $D_j$ do not intersect. Suppose this is false. See figure 5 below:

![Figure 5](image)

and $size(U) < size(U)$ because $C_i$ and $C_j$ are not colliders on $U$, the only collider on $U$ that may not be on $U$ is $R$, and the length of a shortest path from $R$ to a member of $Z$ is less than the length of a shortest path from $C_i$ to a member of $Z$. Hence $U$ is not minimal, contrary to the assumption. Since each $C_i$ is not an ancestor of $S$, it follows directly that no vertex on any path $D_i$ is in $S$. 

\[ \bigcup S, \text{ and } size(U) < size(U) \text{ because } D_i \text{ and } D_j \text{ do not intersect. Suppose this is false. See figure 5 below:} \]

\[ \bigcup S \]
Lemma 2: If $U$ is a minimal d-connecting path between $X$ and $Y$ in $O$, $S$, and $L$ given $R \cup S$, $B$ is a vertex in $O$, and $B \subseteq R \cup S$. Then there is $\exists x_0, \ldots, x_n \subseteq R \cup S$ such that $x_0$ and $x_{n+1}$ are inseparable in $D(G(O,S,L))$.

Proof: Since $U$ is a d-connecting path given $S \cup R$ every collider on $U$ that is not an ancestor of $S$ is an ancestor of a vertex $Z$. Denote the colliders of $U$, that are not ancestors of $C_i, \ldots, C_k$. Let $D_j$ be a shortest directed path from $C_j$ to some vertex $R_j \subseteq R$. It follows by the previous Lemma that $D_j$ and $U$ intersect only at $i$, and that $D_j$ and $D_j'(j \neq j')$ do not intersect. We now construct a sequence of vertices $X_i$ in $O$, s.t. each $X_i$ is either on $U$ or is on a directed path $D_j$ from $C_j$ to $R_j$.

Base Step: $X_0 \subseteq X_i$.

Inductive Step: If $X_i$ is on some path $D_j$ then let $W$ be the next vertex after $W$, such that $W$. If there is no vertex between $W$ and $V$ on $U$, then let $X_{i+1} \subseteq V$. Otherwise let $C_{j*}$ be the first collider on $U$ that is not an ancestor of $S$, and let $X_{i+1}$ be the first vertex in $O$ on the directed path $D_{j*}$ (such a vertex is guaranteed to exist since $R_{j*}$, the endpoint of $D_{j*}$ is in $O$).

It follows from the construction that if $B$ is on $U$, and $B \subseteq O$, then for some $i$, $X_i \subseteq B$.

Claim: $X_i$ and $X_{i+1}$ are inseparable in $D(G(O,S,L))$ under d-separation.

If $X_i$ and $X_{i+1}$ are both on $U$, then $U(X_i, X_{i+1})$ is a path on which no vertex, except the endpoints, is in $O$, and every collider is an ancestor of $U \in \{X_i, X_{i+1}\}$ d-connects $X_i$ and $X_{i+1}$ given $R \cup S$ for any $R \subseteq O \setminus \{X_i, X_{i+1}\}$. So $X_i$ and $X_{i+1}$ are inseparable.

If $X_i$ lies on some path $D_j$, but $X_{i+1}$ is on $U$, then the path $P$ formed by concatenating the directed path $X_i \subseteq \cdots \subseteq C_j$ and $U(C_j, X_{i+1})$ again is such that, excepting the endpoints, no vertex of $P$ is in $O$, and every collider on $P$ is an ancestor of $S$, hence again $X_i$ and $X_{i+1}$ are inseparable in $D(G(O,S,L))$. The cases in which either $X_{i+1}$ alone, or both $X_i$ and $X_{i+1}$ are not on $U$, can be handled similarly.

This completes the proof. $	herefore$

Corollary If $1B$ lies on a minimal d-connecting path between $X$ and $Y$ in $D(G(O,S,L))$, $C$ is a collider on $U$ that is an ancestor of $Z$, but not $S$, $D$ is a shortest directed path from $C$ to some $Z \subseteq Z$, then $\exists C \subseteq C \subseteq S \subseteq (X,Y)$.

Proof: This follows directly from Lemma 2.

Corollary 2 If $U$ is a minimal d-connecting path between $X$ and $Y$ in $D(G(O,S,L))$, $C$ is a collider on $U$ that is an ancestor of $Z$, but not $S$, $D$ is a shortest directed path from $C$ to some $Z \subseteq Z$, then $\exists C \subseteq C \subseteq S \subseteq (X,Y)$. Hence there is a sequence of vertices in $O$ such that consecutive pairs of vertices are inseparable in $D(G(O,S,L))$. Since, by hypothesis $C$ is not an ancestor of $S$, it follows that no vertex on $D$ is in $S$.
Hence $D(V_i, V_{i+1})$ is a directed path from $V_i$ to $V_{i+1}$ on which, with the exception of the endpoints, every vertex is in $\text{Lin}$. And $D_{out}$ is a non-collider $D_{out}$ follows that all $X_i$ are inseparable in $D = (\mathcal{G}, S, L)$. Thus $\forall X_0, \ldots X_n \neq Y \in V, X \Perp_{Y} \iff X_{n+1} \in V$. Hence $<Y = X_n, \ldots X_0 \neq C \in V, V_{i+1} \neq Z>$ establish that $\not\Perp_{DS} \text{CoCon}(X, Y)$ in $D(G, O, S, L)$.

**Theorem 1:** A directed graph $D(G, O, S, L)$ is between $DS$ separated under d-separation.

**Proof:** Suppose, for a contradiction, $\not\Perp_{DS} Y \mid DS \Perp_{DS} O, S, L$, but $O, S, L \not\Perp_{DS} Y$. In this case there is some minimal-path d-separating $X$ and $Y$ given $S \cup (W \cap \text{Between}DS(X, Y))$, but this path is not d-constant $S \cup W$. Clearly it is not possible for a collider on $P$ to have a descendant in $S \cup (W \cap \text{Between}DS(X, Y))$, but not $S \cup W$. Hence there is a path $P$, is $\subseteq S \cup W$, but not in $O \cup S \cup (W \cap \text{Between}DS(X, Y))$. Clearly this is also the $W$-Best with respect to $Y$, and $\not\Perp_{DS} Y \mid W$. But in this case by Corollary 1, $B \in \text{Between}DS(X, Y)$, which is a contradiction.

**Theorem 2** A directed graph $DG(O, S, L)$ is co-connection determined.

Since Between $X \in CoConDS(X, Y)$, the proof of Theorem 1 (replacing with between $\not\Perp_{DS} O, S, L \mid DS \Perp_{DS} Y \mid DS \Perp_{DS} X \mid DS \Perp_{DS} O, S, L$), but this path is not d-constant $S \cup W$. Let $P$ be a minimal d-connecting path between $X$ and $Y$ in $DG(O, S, L)$ given $W \cup S$. Clearly it is not possible for there to be a non-collider $D$ which is in $S \cup (W \cap \text{CoCon}DS(X, Y))$, but not in $S \cup W$. Hence it follows that there is some collider $D$ which has a descendant in $S \cup W$, but not in $S \cup (W \cap \text{CoCon}DS(X, Y))$. Hence $\not\Perp_{DS} W \mid \text{CoCon}DS(X, Y)$. Consider a shortest directed path $D$ from $C$ to some vertex $W$ in $W$. It follows from Lemma 1, and the minimality of $D$ that $D$ does not intersect except at $C$. It now follows by Corollary 2, that $W \in \text{CoCon}DS(X, Y)$.

Therefore if $C$ is an ancestor of $W$, then $W$ has a common ancestor of a vertex $C \in \text{CoCon}DS(X, Y)$, which is contradiction.

I do not include here the proof that chain graphs are co-connection determined. The property follows from the fact that when the extended subgraph is augmented, the only edges that are added are between vertices that are both d-separated from some common third vertex. This is an important difference between augmentation, used in the AMP Markov property, and moralization, used in the LWF Markov property.


