# The Dynamics of Probabilistic Structural Relevance

*L.C. van der Gaag* & *J.-J.Ch. Meyer*
Utrecht University, Department of Computer Science
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands

**Abstract**

Probabilistic inference with a belief network in general is computationally expensive. Since the concept of *structural relevance* provides for identifying parts of a belief network that are irrelevant to a context of interest, it allows for alleviating to some extent the computational burden of inference: inference can be restricted to the network's relevant part. The structurally relevant part of a belief network, however, is not static. It may change dynamically as reasoning progresses. We address the dynamics of structural relevance and introduce the concept of an independence projection to capture these dynamics.

## 1 Introduction

Complex problem domains that are fraught with uncertainty are in the focus of attention of artificial-intelligence research and have been so for some time now. One of the most promising frameworks for dealing with uncertainty that have emerged from this research is the *framework of (Bayesian) belief networks* [Pearl, 1988]. This framework is firmly rooted in probability theory. It provides a powerful and intuitively appealing formalism for representing a probability distribution; informally speaking, a belief network consists of a qualitative part, encoding a domain's variables and the probabilistic independences among them in a directed graph, and a quantitative part, encoding probabilities over these variables. In addition, the framework offers a set of algorithms for probabilistic inference. The belief-network framework is becoming increasingly popular for building knowledge-based systems, and more and more real-life applications employing the framework are being realised.

As applications of the belief-network framework grow larger, the networks involved increase in size accordingly. For large belief networks, probabilistic inference shows a tendency to become rather time-consuming. Since probabilistic inference is known to be NP-hard [Cooper, 1990], this tendency may not be denied in general. In many real-life problem domains, however, reasoning with a belief network concentrates on only some variables of interest. In a medical diagnostic application, for example, the main objective is to establish values for the variables modeling possible disorders and reasoning will concentrate on these variables. In essence, every variable in a belief network may at some time during reasoning be relevant to the variables of interest. Yet, rarely will all of them be of direct relevance. Knowledge of relevance to the variables of interest can be exploited to identify part of a network to which probabilistic inference can be restricted, thereby providing for a reduction of the computational burden involved.

The concept of relevance can be captured in many different ways [Druzdzel & Suermondt, 1994]. One way of capturing relevance in belief networks is by *structural relevance*. Structural relevance builds on the probabilistic independences among the domain's variables that are reflected by a belief network: a variable is deemed structurally relevant to a variable of interest if and only if the topological properties of the network's digraph do not imply these variables being independent. At any time during reasoning with a belief network, the set of variables that are structurally relevant to the variables of interest is unique. However, as reasoning progresses and evidential information is processed, this set may change: the evidence may cause previously structurally relevant variables to become irrelevant, and vice versa. These changes in relevance are not arbitrary but are strictly defined by the belief network at hand. In this paper, we introduce the concept of an *independence projection* to investigate these changes. Building on this concept, we will show that it is possible not only to identify at any time during reasoning the set of variables that are relevant to the variables of interest, but also to identify sets of variables that will never become relevant (anymore) no matter which further evidential information may be processed.

The paper is organised as follows. In Section 2, we review the belief-network formalism. The concept of structural relevance is detailed in Section 3. In Section 4, we introduce the concept of an independence projection for investigating structural relevance; in Section 5, we distinguish between strong and weak independence projections to allow for predicting changes in relevance. In Section 6, we briefly address the computation of the various types of independence projection. The paper is rounded off with some conclusions and directions for further research in Section 7.

## 2 The Belief-Network Formalism

The belief-network formalism provides for a concise representation of a probability distribution on a set of statistical variables; such a representation is called a *(Bayesian) belief network*. A belief network comprises a *qualitative* part and a *quantitative* part. The qualitative part of a belief network is a graphical representation of the independences among the variables holding in the distribution at hand. It takes the form of an acyclic directed graph. In this digraph, each vertex represents a statistical variable that can take one of a finite set of values. The arcs of the digraph with each other model the independences among these variables. Informally speaking, we take an arc $X_i \rightarrow X_j$ in the digraph to represent a direct causal relationship between the linked variables $X_i$ and $X_j$; the direction of the arc designates $X_j$ as the effect of the cause $X_i$. Absence of an arc between two vertices means that the corresponding variables do not influence each other directly, and hence are (conditionally) independent. Associated with the qualitative part of a belief network is a set of probabilities from the distribution at hand, constituting the network's quantitative part. As in this paper we will be concerned with the qualitative part of a belief network only, we will not elaborate on its quantitative part; for further information, the reader is referred to [Pearl, 1988].

We consider once more the set of arcs of the digraph of a belief network. These arcs with each other model the independences holding in the probability distribution that is represented by the network. We define the probabilistic meaning assigned to acyclic digraphs more formally.

**Definition 2.1** *Let $G = (V, A)$ be an acyclic digraph. Let $s$ be a chain in $G$ and let $Y \subseteq V$. Then, $s$ is* blocked *by $Y$ in $G$, denoted $\langle s/Y \rangle_G$, if $s$ contains three consecutive variables $V_1$, $V_2$, $V_3$, for which one of the following conditions holds:*

1. *arcs $V_1 \leftarrow V_2$ and $V_2 \rightarrow V_3$ are on the chain $s$, and $V_2 \in Y$;*
2. *arcs $V_1 \rightarrow V_2$ and $V_2 \rightarrow V_3$ are on the chain $s$, and $V_2 \in Y$;*
3. *arcs $V_1 \rightarrow V_2$ and $V_2 \leftarrow V_3$ are on the chain $s$, and $\sigma^*(V_2) \cap Y = \varnothing$, where $\sigma^*(V_2)$ denotes the set of variables comprising $V_2$ as well as all its descendants.*

In defining the concept of a blocked chain, we have distinguished between three conditions. Figure 1 serves as a reference for these conditions; in the chains representing the conditions 1 and 2, variable $V_2$ is drawn with shading to indicate that it is included in the blocking set for the chain at hand.

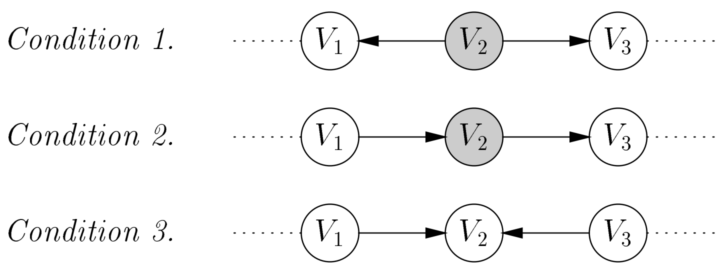


Figure 1: Chain Blocking.

The concept of blocking is defined for single chains only. Building on this concept we define the d-separation criterion to apply to sets of chains.

**Definition 2.2** *Let $G = (V, A)$ be an acyclic digraph. Let $X, Y, Z \subseteq V$. The set of variables $Y$ is said to* d-separate *the sets $X$ and $Z$ in $G$, denoted $\langle X \mid Y \mid Z \rangle_G^d$, if for every chain $s$ between any variable from $X$ and any variable from $Z$ we have that $\langle s/Y \rangle_G$.*

The following definition relates the d-separation criterion, which is graph-theoretic in nature, to the probabilistic concept of independence.

**Definition 2.3** *Let $G = (V, A)$ be an acyclic digraph. The* independence model *$M_G$ of $G$ is the set of statements $I(X, Y, Z)$ such that $I(X, Y, Z) \in M_G$ if and only if $\langle X \mid Y \mid Z \rangle_G^d$, for all mutually disjoint sets of variables $X, Y, Z \subseteq V$.*

A statement $I(X, Y, Z)$ of an independence model is coined an *independence statement*; the statement signifies that the sets of variables $X$ and $Z$ are *conditionally independent* given the set of variables $Y$. Note that independence statements apply to mutually disjoint sets of variables only; in the sequel, we will assume disjointness of sets of variables, unless explicitly stated otherwise.

## 3 Structural Relevance

A probability distribution may embed various independences among its variables. With a distribution Pr we associate an *independence model* including all of Pr's independence statements: the independence model $M_{\mathrm{Pr}}$ of Pr is the set of statements $I(X, Y, Z)$ such that $I(X, Y, Z) \in M_{\mathrm{Pr}}$ if and only if the sets of variables $X$ and $Z$ are conditionally independent given $Y$ in Pr. Upon representation of a probability distribution by a belief network, a digraph is constructed that captures this distribution's independence model. Unfortunately, there are probability distributions whose independence model cannot be represented faithfully by a directed graph, that is, there are probability distributions Pr for which there does not exist a digraph $G$ such that $M_G = M_{\mathrm{Pr}}$ [Pearl, 1988]. For representing a probability distribution Pr by a belief network, therefore, a digraph $G$ is created whose independence model $M_G$ is a (*maximal*) *subset* of the independence model $M_{\mathrm{Pr}}$ of Pr. Note that any independence statement that can be read from such a digraph $G$ by the d-separation criterion is guaranteed to actually hold in the distribution Pr; some independences holding in Pr, however, may have escaped representation in $G$. The concept of *structural relevance* now applies to the independences that can be read from the digraph of a belief network [Druzdzel & Suermondt, 1994].

**Definition 3.1** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. Let $X, Y, Z \subseteq V$. We say that the sets of variables $X$ and $Z$ are* structurally relevant *to each other given the set $Y$ if and only if $I(X, Y, Z) \notin M_G$.*

From the above observations, it will be evident that the concept of structural relevance does not coincide with the concept of dependence. Building on structural relevance for restricting probabilistic inference to the relevant part of a belief network, may therefore in general not preclude all irrelevant computation. Determining structural relevance, however, requires less computational effort than determining dependence: structural relevance can be determined qualitatively by inspecting a belief network's digraph, whereas determining dependence requires extensive manipulation of probabilities. We will return to this observation in Section 6. In the sequel, we build on the concept of structural relevance; for ease of exposition, however, we will use the phrases structural (ir)relevance and (in)dependence interchangeably.

## 4 Independence Projections

The independence model of the digraph of a belief network comprises *all* independence statements that can be read from the digraph by means of the d-separation criterion. It may therefore be looked upon as a *static* description of all independences among the variables discerned, conditional on all possible sets of variables. We observe that, at any time during reasoning with the network, there is a *unique* set of variables for which evidence has actually been processed; in the sequel, we will refer to this set of variables as the current *body of evidence*. To the current body of evidence only *some* of the independence statements of the digraph's model apply; these are the statements that are conditional on the body of evidence. The set of independence statements that apply to the current body of evidence is termed the *independence projection* of the model given the evidence.

**Definition 4.1** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. Let $Y \subseteq V$ be the current body of evidence. The* independence projection *of $M_G$ given $Y$, denoted $I_G(Y)$, is the set*

$$I_G(Y) = \{(X, Z) \mid I(X, Y, Z) \in M_G\}$$

*The* dependence projection *of $M_G$ given $Y$, denoted $D_G(Y)$, is the set*

$$D_G(Y) = \{(X, Z) \mid I(X, Y, Z) \notin M_G\}$$

Note that when the current body of evidence is empty, the independence projection of a digraph's model includes only unconditional independence statements. The following properties are easily verified for any body of evidence $Y$:

- $I_G(Y) \cap D_G(Y) = \varnothing$;
- $I_G(Y) \cup D_G(Y) = \mathcal{P}(V) \times \mathcal{P}(V)$;

where $\mathcal{P}(V)$ denotes the power set of the set of variables $V$; furthermore, we have that

- $\{I(X, Y, Z) \mid (X, Z) \in I_G(Y)\} \subseteq M_G$, for any $Y \subseteq V$;
- $\bigcup_{Y \subseteq V} \{I(X, Y, Z) \mid (X, Z) \in I_G(Y)\} = M_G$;

that is, the independence projection of a digraph's model given any body of evidence $Y$ indicates a subset of the entire set of independence statements that can be read from the digraph, and the independence projections for all $Y \subseteq V$ with each other once again span the entire model of the digraph.

The d-separation criterion provides for reading from a belief network's digraph any independence projection, as is stated more formally in the following lemma.

**Lemma 4.2** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. Let $Y \subseteq V$ and let $I_G(Y)$ be the independence projection of $M_G$ given $Y$. Then, for all sets of variables $X, Z \subseteq V$, we have that $(X, Z) \in I_G(Y)$ if and only if $\langle X \mid Y \mid Z \rangle^d_G$.*

When reasoning with a belief network, the body of evidence changes: it *grows* monotonically as a result of processing evidential information. Also, the set of independence statements given the evidence may change: variables that are independent given the current body of evidence may become dependent upon observing an additional piece of evidence, and vice versa. Note that these changes may be *non-monotonic* in nature.

We consider the digraph $G$ of a belief network and its associated independence model $M_G$. Let $Y$ be the current body of evidence and let $I_G(Y)$ be the independence projection of $M_G$ given $Y$. Now suppose that new evidential information becomes available for some of the variables in the network and that the new body of evidence equals the set of variables $Y'$ with $Y \subset Y'$. We compare the new independence projection $I_G(Y')$ given this set $Y'$ with the independence projection $I_G(Y)$. For two sets of variables $X$ and $Z$, it is possible that both $(X, Z) \in I_G(Y)$ and $(X, Z) \in I_G(Y')$.

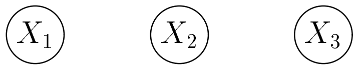


Figure 2: $(\{X_1\},\{X_3\}) \in I_G(\varnothing)$ and $(\{X_1\},\{X_3\}) \in I_G(\{X_2\})$.

**Example 4.3** Consider the digraph $G$ shown in Figure 2. Exploiting the d-separation criterion, we find that $\{X_1\}$ and $\{X_3\}$ are d-separated by $\varnothing$ as well as by $\{X_2\}$, that is, $\langle\{X_1\} \mid \varnothing \mid \{X_3\}\rangle_G^d$ and $\langle\{X_1\} \mid \{X_2\} \mid \{X_3\}\rangle_G^d$. From Lemma 4.2, it follows that $(\{X_1\},\{X_3\}) \in I_G(\varnothing)$ and $(\{X_1\},\{X_3\}) \in I_G(\{X_2\})$. So, the variables $X_1$ and $X_3$ are independent, and remain to be so when evidence for $X_2$ is processed. □

For two sets of variables $X$ and $Z$, it is also possible that $(X,Z) \notin I_G(Y)$ and $(X,Z) \notin I_G(Y')$.

**Example 4.4** Consider the digraph $G$ shown in Figure 3. From this digraph, it is easily seen that the chain $X_1 \rightarrow X_3$ between the variables $X_1$ and $X_3$ cannot be blocked, and hence that the sets of variables $\{X_1\}$ and $\{X_3\}$ can never be d-separated — not by $\varnothing$, and not by $\{X_2\}$. From Lemma 4.2, it follows that $(\{X_1\},\{X_3\}) \notin I_G(\varnothing)$ and $(\{X_1\},\{X_3\}) \notin I_G(\{X_2\})$. So, the variables $X_1$ and $X_3$ are dependent and remain to be so when evidence for the variable $X_2$ is processed. □

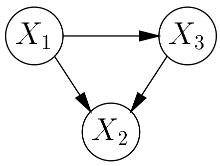


Figure 3: $(\{X_1\},\{X_3\}) \notin I_G(\varnothing)$ and $(\{X_1\},\{X_3\}) \notin I_G(\{X_2\})$.

New independences may arise as further evidential information is processed, that is, for two sets of variables $X$ and $Z$ it is possible that $(X,Z) \notin I_G(Y)$ and $(X,Z) \in I_G(Y')$.

**Example 4.5** Consider the digraph $G$ shown in Figure 4. From this digraph, we read that $\{X_1\}$ and $\{X_3\}$ are not d-separated by the empty set. From Lemma 4.2, it follows that $(\{X_1\},\{X_3\}) \notin I_G(\varnothing)$. However, $\{X_1\}$ and $\{X_3\}$ are d-separated by the set $\{X_2\}$, and therefore $(\{X_1\},\{X_3\}) \in I_G(\{X_2\})$. So, the variables $X_1$ and $X_3$ initially are dependent but become independent upon processing evidence for the variable $X_2$. □

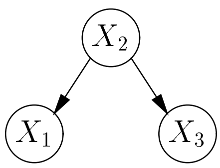


Figure 4: $(\{X_1\},\{X_3\}) \notin I_G(\varnothing)$ and $(\{X_1\},\{X_3\}) \in I_G(\{X_2\})$.

To conclude, we observe that independences may also disappear as reasoning progresses, that is, for two sets of variables $X$ and $Z$ it is possible that $(X,Z) \in I_G(Y)$ and $(X,Z) \notin I_G(Y')$.

**Example 4.6** Consider the digraph $G$ shown in Figure 5. From this digraph, we read that $\{X_1\}$ and $\{X_3\}$ are d-separated by the empty set; we conclude that $(\{X_1\}, \{X_3\}) \in I_G(\varnothing)$. However, $\{X_1\}$ and $\{X_3\}$ are not d-separated by the set $\{X_2\}$, and therefore $(\{X_1\}, \{X_3\}) \notin I_G(\{X_2\})$. So, the variables $X_1$ and $X_3$ initially are independent, yet become dependent upon processing evidence for the variable $X_2$. $\square$

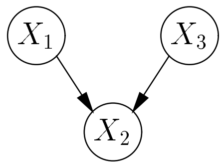


Figure 5: $(\{X_1\}, \{X_3\}) \in I_G(\varnothing)$ and $(\{X_1\}, \{X_3\}) \notin I_G(\{X_2\})$.

The following lemma summarises the above observations.

**Lemma 4.7** *Let $V$ be a set of (at least three) variables. For any digraph $G = (V, A)$ and any set of variables $Y \subseteq V$, let $I_G(Y)$ be the independence projection of the independence model $M_G$ of $G$ given $Y$. Then,*

- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in I_G(Y)$ and $(X, Z) \in I_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \notin I_G(Y)$ and $(X, Z) \notin I_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \notin I_G(Y)$ and $(X, Z) \in I_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in I_G(Y)$ and $(X, Z) \notin I_G(Y')$.*

# 5 Strong and Weak Independence Projections

In the previous section, we have shown that an independence projection may change as the body of evidence grows during reasoning with a belief network. In addition, we have argued that the changes incurred may be non-monotonic in nature. Now, for exploiting the concept of structural relevance for restricting probabilistic inference to part of a network, it is worthwhile to distinguish between independences that hold given the current body of evidence and remain to hold no matter which information may further be processed, and independences that may be invalidated as further evidence becomes available. For this purpose, we define the concepts of *strong* and *weak* independence projection.

**Definition 5.1** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. Let $Y \subseteq V$ be the current body of evidence. The* strong independence projection *of $M_G$ given $Y$, denoted $S_G(Y)$, is the set*

$$S_G(Y) = \{(X, Z) \mid I(X, Y', Z) \in M_G \textit{ for all sets of variables } Y' \textit{ with } Y \subseteq Y' \subseteq V\}$$

*The* weak independence projection *of $M_G$ given $Y$, denoted $W_G(Y)$, is the set*

$$W_G(Y) = \{(X, Z) \mid I(X, Y, Z) \in M_G \text{ and } I(X, Y', Z) \notin M_G \text{ for some set of variables } Y' \text{ with } Y \subseteq Y' \subseteq V\}$$

The following properties are easily verified for any body of evidence $Y$:

- $S_G(Y) \cap W_G(Y) = \varnothing$;
- $S_G(Y) \cup W_G(Y) = I_G(Y)$.

As for independence projections in general, the d-separation criterion provides for reading from a network's digraph any strong or weak independence projection. Before stating this property more formally, we distinguish between two different ways of blocking a chain.

**Definition 5.2** *Let $G = (V, A)$ be an acyclic digraph. Let $s$ be a chain in $G$ and let $Y \subseteq V$. Then, $s$ is blocked by $Y$ in $G$* by presence of information, *denoted $\langle s/Y, \in\rangle_G$, if $s$ contains three consecutive variables $V_1, V_2, V_3$, for which one of the following conditions holds:*

- *arcs $V_1 \leftarrow V_2$ and $V_2 \rightarrow V_3$ are on the chain $s$, and $V_2 \in Y$;*
- *arcs $V_1 \rightarrow V_2$ and $V_2 \rightarrow V_3$ are on the chain $s$, and $V_2 \in Y$.*

*The chain $s$ is blocked by $Y$ in $G$* by absence of information, *denoted $\langle s/Y, \notin\rangle_G$, if $s$ is blocked by $Y$ in $G$ and $s$ is not blocked by $Y$ in $G$ by presence of information.*

Building on the two different ways of blocking a chain, we distinguish between strong and weak d-separation.

**Definition 5.3** *Let $G = (V, A)$ be an acyclic digraph. Let $X, Y, Z \subseteq V$. The set of variables $Y$ is said to* strongly *d-separate the sets of variables $X$ and $Z$ in $G$, denoted $\langle X \mid Y \mid Z\rangle_G^{S_d}$, if for every chain $s$ between any variable from $X$ and any variable from $Z$ we have that $\langle s/Y, \in\rangle_G$. The set $Y$ is said to* weakly *d-separate the sets $X$ and $Z$, denoted $\langle X \mid Y \mid Z\rangle_G^{W_d}$, if $\langle X \mid Y \mid Z\rangle_G^d$ and $Y$ does not strongly d-separate $X$ and $Z$.*

Note that in a digraph $G$ two sets of variables $X$ and $Z$ are weakly d-separated by a set $Y$ if they are d-separated by $Y$ and there exists at least one chain in $G$ between a variable from $X$ and a variable from $Z$ that is blocked by $Y$ by absence of information.

The following lemma now states that any strong independence projection can be read from a belief network's digraph by exploiting the strong d-separation criterion, that is, by inspecting the chains in the digraph that are blocked by presence of information. The basic idea underlying the lemma is that, once a chain is blocked by presence of information, it can never become unblocked upon processing further evidence.

**Lemma 5.4** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. Let $Y \subseteq V$ and let $S_G(Y)$ be the strong independence projection of $M_G$ given $Y$. Then, for all sets of variables $X, Z \subseteq V$ we have that $(X, Z) \in S_G(Y)$ if and only if $\langle X \mid Y \mid Z\rangle_G^{S_d}$.*

Note that the property stated in the lemma allows for verifying strong independence without having to check the independence projections for all bodies of evidence larger than the current one.

The following lemma states that weak independence projections can be read from a belief network's digraph by exploiting the weak d-separation criterion. The basic idea underlying the lemma is that only chains that are blocked by absence of information can become unblocked upon processing further evidence.

**Lemma 5.5** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. Let $Y \subseteq V$ and let $W_G(Y)$ be the weak independence projection of $M_G$ given $Y$. Then, for all sets of variables $X, Z \subseteq V$ we have that $(X, Z) \in W_G(Y)$ if and only if $\langle X \mid Y \mid Z\rangle_G^{W_d}$.*

We consider once more the digraph $G$ of a belief network and its associated independence model $M_G$. Let $Y$ be the current body of evidence and let $S_G(Y)$ and $W_G(Y)$ be the strong and weak independence projections of $M_G$ given $Y$, respectively. Now suppose that new evidential information is processed for some of the variables in the network and that the new body of evidence equals the set $Y'$ with $Y \subset Y'$. We compare the new independence projections $S_G(Y')$ and $W_G(Y')$ given this set $Y'$ with the independence projections $S_G(Y)$ and $W_G(Y)$, respectively.

For two sets of variables $X$ and $Z$, if $X$ and $Z$ are strongly independent given $Y$, then they are also strongly independent given $Y'$. This property follows directly from Definition 5.1; for completeness of presentation, the property is stated in the following lemma.

**Lemma 5.6** *Let $G = (V, A)$ be an acyclic digraph and let $M_G$ be its independence model. For any set of variables $Y \subseteq V$, let $S_G(Y)$ be the strong independence projection of $M_G$ given $Y$. For all sets of variables $X, Z \subseteq V$ and for all sets of variables $Y, Y' \subseteq V$ with $Y \subseteq Y'$, if $(X, Z) \in S_G(Y)$ then $(X, Z) \in S_G(Y')$.*

We now turn to the dynamics of weak independence projections. For two sets of variables $X$ and $Z$, it is possible that $(X, Z) \in W_G(Y)$ and $(X, Z) \in D_G(Y')$.

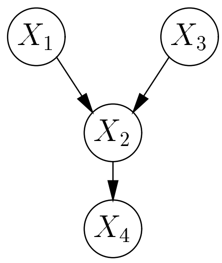


Figure 6: $(\{X_1\}, \{X_3\}) \in W_G(\varnothing)$ and $(\{X_1\}, \{X_3\}) \in D_G(\{X_4\})$.

**Example 5.7** Consider the digraph $G$ shown in Figure 6. Exploiting the d-separation criterion, we find that $\{X_1\}$ and $\{X_3\}$ are d-separated by $\varnothing$. Since for the chain $s$ between $X_1$ and $X_3$ we have that $\langle s/\varnothing, \notin\rangle_G$, we find that $\{X_1\}$ and $\{X_3\}$ are weakly d-separated by $\varnothing$. From Lemma 5.5 we conclude that $(\{X_1\}, \{X_3\}) \in W_G(\varnothing)$. From the digraph we further read that $\{X_1\}$ and $\{X_3\}$ are not d-separated by $\{X_4\}$. We conclude that $(\{X_1\}, \{X_3\}) \in D_G(\{X_4\})$. □

For two sets of variables $X$ and $Z$, it is also possible that $(X, Z) \in W_G(Y)$ and $(X, Z) \in W_G(Y')$.

**Example 5.8** Consider the digraph $G$ shown in Figure 7. Exploiting the d-separation criterion, we find that $\{X_1\}$ and $\{X_3\}$ are d-separated by $\varnothing$. Since for the chain $s$ between $X_1$ and $X_3$ we have that $\langle s/\varnothing, \notin\rangle_G$, we find that $(\{X_1\}, \{X_3\}) \in W_G(\varnothing)$. From the digraph we further read that $\{X_1\}$ and $\{X_3\}$ are also weakly d-separated by $\{X_4\}$. We conclude that $(\{X_1\}, \{X_3\}) \in W_G(\{X_4\})$. □

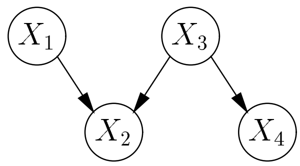


Figure 7: $(\{X_1\}, \{X_3\}) \in W_G(\varnothing)$ and $(\{X_1\}, \{X_3\}) \in W_G(\{X_4\})$.

To conclude, for two sets of variables $X$ and $Z$, it is possible that $(X, Z) \in W_G(Y)$ and $(X, Z) \in S_G(Y')$.

**Example 5.9** Consider the digraph $G$ shown in Figure 8. From the digraph, we read that $\{X_1\}$ and $\{X_3\}$ are weakly d-separated by $\varnothing$. Using Lemma 5.5, we find that $(\{X_1\}, \{X_3\}) \in W_G(\varnothing)$. From the digraph we also read that $\{X_1\}$ and $\{X_3\}$ are d-separated by $\{X_4\}$. Note that for the chain $s$ in $G$ between $X_1$ and $X_3$ we have that $\langle s/\{X_4\}, \in\rangle_G$. From Definition 5.3, we therefore have that $\{X_1\}$ and $\{X_3\}$ are strongly d-separated by $\{X_4\}$. From Lemma 5.4, we conclude that $(\{X_1\}, \{X_3\}) \in S_G(\{X_4\})$. □

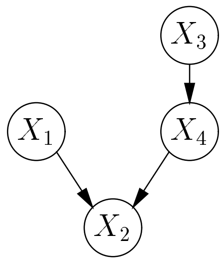


Figure 8: $(\{X_1\}, \{X_3\}) \in W_G(\varnothing)$ and $(\{X_1\}, \{X_3\}) \in S_G(\{X_4\})$.

The following lemma summarises the above observations.

**Lemma 5.10** *Let $V$ be a set of (at least four) variables. For any digraph $G = (V, A)$ and any set of variables $Y \subseteq V$, let $D_G(Y)$, $W_G(Y)$, and $S_G(Y)$ be the dependence projection, the weak independence projection, and the strong independence projection, respectively, of the independence model $M_G$ of $G$ given $Y$. Then,*

- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in W_G(Y)$ and $(X, Z) \in D_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in W_G(Y)$ and $(X, Z) \in W_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in W_G(Y)$ and $(X, Z) \in S_G(Y')$.*

We now turn to the dynamics of dependence projections. For two sets of variables $X$ and $Z$, it is possible that $(X, Z) \in D_G(Y)$ and $(X, Z) \in D_G(Y')$; we refer to Example 4.4 and Figure 3 for an illustration of this observation. For two sets $X$ and $Z$, it is also possible that $(X, Z) \in D_G(Y)$ and $(X, Z) \in S_G(Y')$.

**Example 5.11** Consider once more the digraph $G$ shown in Figure 4. Exploiting the d-separation criterion, we find that $\{X_1\}$ and $\{X_3\}$ are not d-separated by $\varnothing$, and hence that $(\{X_1\}, \{X_3\}) \in D_G(\varnothing)$. From the digraph, we further read that $\{X_1\}$ and $\{X_3\}$ are d-separated by $\{X_2\}$. Since for the chain $s$ between $X_1$ and $X_3$ we have that $\langle s/\{X_2\}, \in\rangle_G$, we conclude that $(\{X_1\}, \{X_3\}) \in S_G(\{X_2\})$. □

To conclude, for two sets of variables $X$ and $Z$, it is possible that $(X, Z) \in D_G(Y)$ and $(X, Z) \in W_G(Y')$.

**Example 5.12** Consider the digraph $G$ shown in Figure 9. Exploiting the d-separation criterion, we find that $\{X_1\}$ and $\{X_3\}$ are not d-separated by $\varnothing$. We conclude that $(\{X_1\}, \{X_3\}) \in D_G(\varnothing)$. From the digraph, we further read that $\{X_1\}$ and $\{X_3\}$ are d-separated by $\{X_2\}$. Note that for the chain $s = X_1 \rightarrow X_4 \leftarrow X_3$ between $X_1$ and $X_3$ we have that $\langle s/\{X_2\}, \notin\rangle_G$. From Definition 5.3, we have that $\{X_1\}$ and $\{X_3\}$ are weakly d-separated by $\{X_2\}$. By Lemma 5.5, we conclude that $(\{X_1\}, \{X_3\}) \in W_G(\{X_2\})$. □

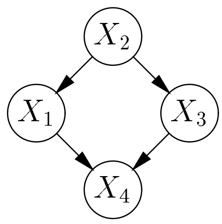


Figure 9: $(\{X_1\}, \{X_3\}) \in D_G(\varnothing)$ and $(\{X_1\}, \{X_3\}) \in W_G(\{X_2\})$.

The following lemma summarises the above observations.

**Lemma 5.13** *Let $V$ be a set of (at least four) variables. For any digraph $G = (V, A)$ and any set of variables $Y \subseteq V$, let $D_G(Y)$, $W_G(Y)$, and $S_G(Y)$ be the dependence projection, the weak independence projection, and the strong independence projection, respectively, of the independence model $M_G$ of $G$ given $Y$. Then,*

- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in D_G(Y)$ and $(X, Z) \in D_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subseteq V$ with $Y \subset Y'$ such that $(X, Z) \in D_G(Y)$ and $(X, Z) \in W_G(Y')$;*
- *there exist a digraph $G = (V, A)$ and sets of variables $X, Y, Y', Z \subset V$ with $Y \subset Y'$ such that $(X, Z) \in D_G(Y)$ and $(X, Z) \in S_G(Y')$.*

The Lemmas 5.6, 5.10, and 5.13 with each other describe the behaviour of the strong independence projection, the weak independence projection, and the dependence projection of a digraph's independence model given a body of evidence, under processing of further evidential information. From these lemmas it is easily verified that

- $S_G(Y) \subseteq S_G(Y')$;
- $D_G(Y) \cup W_G(Y) \supseteq D_G(Y') \cup W_G(Y')$;

for all bodies of evidence $Y$ and $Y'$ with $Y \subseteq Y'$.

# 6 Algorithmic Considerations

We have argued before that the concept of structural relevance allows for alleviating the computational burden of probabilistic inference with a belief network since it provides for restricting inference to a relevant part of the network. For this purpose, algorithms for efficiently computing (strong and weak) independence projections of a digraphs's independence model are required. By exploiting the d-separation criterion and the concepts of strong and weak d-separation, any independence projection can be computed in $O(n^2)$ time, where $n$ is the number of variables in the digraph. Algorithms to this end build on a (modified) graph traversal [Geiger *et al.*, 1990].

# 7 Conclusions

The concept of structural relevance provides for identifying part of a belief network that is relevant to a context of interest. We have shown that the relevant part of a belief network may change dynamically as reasoning progresses. For investigating the dynamics of structural relevance, we have introduced the concept of independence projection. We have distinguished between strong and weak independence projections to allow for predicting changes in relevance. More in specific, these types of projection provide for identifying parts of a network that will never become relevant (anymore) during reasoning with the network. We have shown that independence projections can be read from a belief network's digraph. To conclude, we have argued that algorithms for efficiently computing strong and weak independence projections are easily designed. The design of dynamic algorithms that provide for computing new independence projections from previous ones remains as a challenging subject for further research.

# References

[Cooper, 1990] G.F. Cooper (1990). The computational complexity of probabilistic inference using Bayesian belief networks, *Artificial Intelligence*, vol. 42, pp. 393 – 405.

[Druzdzel & Suermondt, 1994] M.J. Druzdzel, H.J. Suermondt (1994). Relevance in probabilistic models: "backyards" in a small world, *The AAAI-94 Fall Symposium on Relevance.*

[Geiger *et al.*, 1990] D. Geiger, T. Verma, J. Pearl (1990). d-Separation: from theorems to algorithms, in: M. Henrion, R.D. Shachter, L.N. Kanal, J.F. Lemmer (eds). *Uncertainty in Artificial Intelligence 5*, North-Holland, Amsterdam, pp. 139 – 148.

[Pearl, 1988] J. Pearl (1988). *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*, Morgan Kaufmann, Palo Alto.