Conditional Independence and Irrelevance

Philip Dawid

University of Cambridge

Royal Netherlands Academy of Arts and Sciences Colloquium
“DEPENDENCE LOGIC”

3–5 March 2014
Probabilistic Independence

- Base space $\Omega$
  - discrete — for simplicity (only)
Probabilistic Independence

- Base space $\Omega$
  - discrete — for simplicity (only)
- Random variables $X, Y, Z, \ldots$
  - functions on $\Omega$

Probability distribution $P$ on $\Omega$

Marginal/conditional distributions:

$$p(x) := P(X = x)$$

$$p(x | y) := P(X = x | Y = y) = p(x, y) / p(y)$$

[defined when $p(y) > 0$]
Probabilistic Independence

- Base space $\Omega$
  - discrete — for simplicity (only)
- Random variables $X, Y, Z, \ldots$
  - functions on $\Omega$
- Probability distribution $P$ on $\Omega$

Marginal/conditional distributions:
\[
p(x) := P(X = x)
\]
\[
p(x|y) := P(X = x|Y = y) = \frac{p(x, y)}{p(y)} \quad \text{[defined when } p(y) > 0 \text{]}
\]
Probabilistic Independence

- Base space $\Omega$
  - discrete — for simplicity (only)
- Random variables $X, Y, Z, \ldots$
  - functions on $\Omega$
- Probability distribution $P$ on $\Omega$

Marginal/conditional distributions:

$$p(x) := P(X = x)$$
$$p(x \mid y) := P(X = x \mid Y = y)$$
$$= p(x, y) / p(y)$$

[defined when $p(y) > 0$]
Probabilistic Independence

Say $X$ is independent of $Y$, and write $X \Perp_{p} Y \ [P]$, or $X \Perp Y$, if

$$p(x, y) \equiv p(x) p(y)$$
$$p(x, y) \equiv a(x) b(y)$$

▶ "No interaction" between $X$ and $Y$

$p(x \mid y) \equiv p(x)$

▶ Information about $Y$ is irrelevant to uncertainty about $X$.
Probabilistic Independence

Say $X$ is independent of $Y$, and write $X \perp \!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\

\[ p(x, y) \equiv p(x) p(y) \]

\[ p(x, y) \equiv a(x) b(y) \]

“No interaction” between $X$ and $Y$
Probabilistic Independence

Say $X$ is independent of $Y$, and write $X \perp\!\!\!\!\!\!\!\!\!\!_p Y [P]$, or $X \perp\!\!\!\!\!\!\!\!\!\! Y$, if

$$p(x, y) \equiv p(x) p(y)$$

$$p(x, y) \equiv a(x) b(y)$$

▶ “No interaction” between $X$ and $Y$

$$p(x \mid y) \equiv p(x)$$

$$p(x \mid y) \equiv a(x)$$
Probabilistic Independence

Say \( X \) is independent of \( Y \), and write \( X \perp \!\!\!\!\!\!\!\!\!\!\!\perp_{\mathcal{P}} Y \ [P] \), or \( X \perp \!\!\!\!\!\!\!\!\!\!\!\perp Y \), if

\[
\begin{align*}
    p(x, y) & \equiv p(x) p(y) \\
    p(x, y) & \equiv a(x) b(y)
\end{align*}
\]

- “No interaction” between \( X \) and \( Y \)

\[
\begin{align*}
    p(x \mid y) & \equiv p(x) \\
    p(x \mid y) & \equiv a(x)
\end{align*}
\]

- Information about \( Y \) is \textit{irrelevant} to uncertainty about \( X \).
Conditional Independence

Say \( X \) is (conditionally) independent of \( Y \), given \( Z \), and write \( X \perp \_p Y \mid Z [P] \), or \( X \perp Y \mid Z \), if

\[
\begin{align*}
\mathbb{P}(x, y \mid z) & \equiv \mathbb{P}(x \mid z) \mathbb{P}(y \mid z) \equiv a(x, z) b(y, z) / \mathbb{P}(z) \\
\mathbb{P}(x, y, z) & \equiv \mathbb{P}(x \mid z) \mathbb{P}(y \mid z) \mathbb{P}(z) \equiv a(x, z) b(y, z) / \mathbb{P}(z)
\end{align*}
\]

\( \text{▶ Once } Z \text{ is known, no further interaction between } X \text{ and } Y \)

\( \text{▶ Once } Z \text{ is known, any further information about } Y \text{ is irrelevant to uncertainty about } X \)

\( \text{▶ } X \text{ depends on } Z \text{ (and not on } Y) \quad \text{satisfies dependence logic??} \)
Conditional Independence

Say $X$ is (conditionally) independent of $Y$, given $Z$, and write $X \perp \perp Y \mid Z [P]$, or $X \perp \perp Y \mid Z$, if

\[
p(x, y \mid z) \equiv p(x \mid z) p(y \mid z)
\]

\[
p(x, y \mid z) \equiv a(x, z) b(y, z)
\]

\[
p(x, y, z) \equiv p(x \mid z) p(y \mid z) p(z)
\]

\[
p(x, y, z) \equiv p(x, z) p(y, z) / p(z)
\]

\[
p(x, y, z) \equiv a(x, z) b(y, z)
\]
Conditional Independence

Say $X$ is (conditionally) independent of $Y$, given $Z$, and write $X \perp \perp Y \mid Z [P]$, or $X \perp \perp Y \mid Z$, if

\[ p(x, y \mid z) \equiv p(x \mid z) p(y \mid z) \]
\[ p(x, y \mid z) \equiv a(x, z) b(y, z) \]
\[ p(x, y, z) \equiv p(x \mid z) p(y \mid z) p(z) \]
\[ p(x, y, z) \equiv p(x, z) p(y, z)/p(z) \]
\[ p(x, y, z) \equiv a(x, z) b(y, z) \]

â†’ Once $Z$ is known, no further interaction between $X$ and $Y$
Conditional Independence

Say $X$ is (conditionally) independent of $Y$, given $Z$, and write $X \perp_{p} Y \mid Z [P]$, or $X \perp Y \mid Z$, if

\[
p(x, y \mid z) \equiv p(x \mid z) p(y \mid z)
\]
\[
p(x, y \mid z) \equiv a(x, z) b(y, z)
\]
\[
p(x, y, z) \equiv p(x \mid z) p(y \mid z)p(z)
\]
\[
p(x, y, z) \equiv p(x, z) p(y, z)/p(z)
\]
\[
p(x, y, z) \equiv a(x, z) b(y, z)
\]

Once $Z$ is known, no further interaction between $X$ and $Y$

\[
p(x \mid y, z) \equiv p(x \mid z)
\]
\[
p(x \mid y, z) \equiv a(x, z)
\]
Conditional Independence

Say $X$ is (conditionally) independent of $Y$, given $Z$, and write $X \perp \! \! \! \perp Y \mid Z \ [P]$, or $X \perp \! \! \! \perp Y \mid Z$, if

$$p(x, y \mid z) \equiv p(x \mid z) p(y \mid z)$$
$$p(x, y \mid z) \equiv a(x, z) b(y, z)$$
$$p(x, y, z) \equiv p(x \mid z) p(y \mid z) p(z)$$
$$p(x, y, z) \equiv p(x, z) p(y, z) / p(z)$$
$$p(x, y, z) \equiv a(x, z) b(y, z)$$

Once $Z$ is known, no further interaction between $X$ and $Y$

$$p(x \mid y, z) \equiv p(x \mid z)$$
$$p(x \mid y, z) \equiv a(x, z)$$

Once $Z$ is known, any further information about $Y$ is irrelevant to uncertainty about $X$
Conditional Independence

Say $X$ is (conditionally) independent of $Y$, given $Z$, and write $X \perp \perp Y \mid Z [P]$, or $X \perp \perp Y \mid Z$, if

$$p(x, y \mid z) \equiv p(x \mid z) p(y \mid z)$$
$$p(x, y \mid z) \equiv a(x, z) b(y, z)$$
$$p(x, y, z) \equiv p(x \mid z) p(y \mid z) p(z)$$
$$p(x, y, z) \equiv p(x, z) p(y, z) / p(z)$$
$$p(x, y, z) \equiv a(x, z) b(y, z)$$

▶ Once $Z$ is known, no further interaction between $X$ and $Y$

$$p(x \mid y, z) \equiv p(x \mid z)$$
$$p(x \mid y, z) \equiv a(x, z)$$

▶ Once $Z$ is known, any further information about $Y$ is irrelevant to uncertainty about $X$

▶ $X$ depends on $Z$ (and not on $Y$)
Conditional Independence

Say \( X \) is (conditionally) independent of \( Y \), given \( Z \), and write \( X \perp \!\!\!\!\!\!\!\!\!\!\!\perp Y \mid Z [P] \), or \( X \perp \!\!\!\!\!\!\!\!\!\!\!\perp Y \mid Z \), if

\[
p(x, y \mid z) \equiv p(x \mid z) p(y \mid z)
\]

\[
p(x, y \mid z) \equiv a(x, z) b(y, z)
\]

\[
p(x, y, z) \equiv p(x \mid z) p(y \mid z)p(z)
\]

\[
p(x, y, z) \equiv p(x, z) p(y, z)/p(z)
\]

\[
p(x, y, z) \equiv a(x, z) b(y, z)
\]

▶ Once \( Z \) is known, no further interaction between \( X \) and \( Y \)

\[
p(x \mid y, z) \equiv p(x \mid z)
\]

\[
p(x \mid y, z) \equiv a(x, z)
\]

▶ Once \( Z \) is known, any further information about \( Y \) is irrelevant to uncertainty about \( X \)

▶ \( X \) depends on \( Z \) (and not on \( Y \))

▶ satisfies dependence logic???
Fundamental Properties [1, 3, 6]

Writing $W \preceq Y$ to mean that $W$ is a function of $Y$:

\begin{align*}
\text{P1 : } & X \perp \perp Y \mid Z \quad \Rightarrow \quad Y \perp \perp X \mid Z \\
\text{P2 : } & X \perp \perp Y \mid X \\
\text{P3 : } & X \perp \perp Y \mid Z \quad \text{and} \quad W \preceq Y \quad \quad \Rightarrow \quad X \perp \perp W \mid Z \\
\text{P4 : } & X \perp \perp Y \mid Z \quad \text{and} \quad W \preceq Y \quad \quad \Rightarrow \quad X \perp \perp Y \mid (W, Z) \\
\text{P5 : } & X \perp \perp W \mid (Y, Z) \quad \text{and} \quad X \perp \perp (Y, W) \mid Z. 
\end{align*}
Fundamental Properties [1, 3, 6]

Writing $W \preceq Y$ to mean that $W$ is a function of $Y$:

P1 : $X \perp \perp Y \mid Z$ \quad \Rightarrow \quad Y \perp \perp X \mid Z

P2 : $X \perp \perp Y \mid X$

P3 : \begin{align*}
X \perp \perp Y \mid Z \\
W \preceq Y
\end{align*}

\Rightarrow \quad X \perp \perp W \mid Z

P4 : \begin{align*}
X \perp \perp Y \mid Z \\
W \preceq Y
\end{align*}

\Rightarrow \quad X \perp \perp Y \mid (W, Z)

P5 : \begin{align*}
X \perp \perp W \mid (Y, Z)
\end{align*}

\Rightarrow \quad X \perp \perp (Y, W) \mid Z.

▶ Interpretations in terms of “irrelevance”
Use as Axioms

“Nearest neighbour” property of a Markov Chain:

Suppose:

(i). \( X_3 \perp \perp X_1 \mid X_2 \)
(ii). \( X_4 \perp \perp (X_1, X_2) \mid X_3 \)
(iii). \( X_5 \perp \perp (X_1, X_2, X_3) \mid X_4 \)
Use as Axioms

“Nearest neighbour” property of a Markov Chain:

![Diagram of Markov Chain]

Suppose:

(i). \( X_3 \perp \perp X_1 \mid X_2 \)

(ii). \( X_4 \perp \perp (X_1, X_2) \mid X_3 \)

(iii). \( X_5 \perp \perp (X_1, X_2, X_3) \mid X_4 \)

Then \( X_3 \perp \perp (X_1, X_5) \mid (X_2, X_4) \).
Proof
Applying P4 and P1 in turn to (ii), we obtain
\[ X_1 \perp \perp X_4 \mid (X_2, X_3), \tag{1} \]
while from (i) and P1 we have
\[ X_1 \perp \perp X_3 \mid X_2. \tag{2} \]
On applying P5 to (2) and (1), we now deduce
\[ X_1 \perp \perp (X_3, X_4) \mid X_2 \tag{3} \]
whence, by P4 and P1,
\[ X_3 \perp \perp X_1 \mid (X_2, X_4). \tag{4} \]
Also, by (iii) and P4 we have
\[ X_5 \perp \perp (X_1, X_3) \mid (X_2, X_4) \tag{5} \]
and so, by P4 and P1,
\[ X_3 \perp \perp X_5 \mid (X_1, X_2, X_4). \tag{6} \]
The result now follows on applying P5 to (4) and (6).
An additional axiom? [12, 13]

If

- $X \perp Y \mid (Z, W)$
- $Z \perp W \mid X$
- $Z \perp W \mid Y$
- $X \perp Y$

then

- $Z \perp W \mid (X, Y)$
- $X \perp Y \mid Z$
- $X \perp Y \mid W$
- $Z \perp W$

Probabilistic conditional independence has no finite axiomatisation.
An additional axiom? [12, 13]

If

\[ X \perp Y \mid (Z, W) \]
\[ Z \perp W \mid X \]
\[ Z \perp W \mid Y \]
\[ X \perp Y \]

then

\[ Z \perp W \mid (X, Y) \]
\[ X \perp Y \mid Z \]
\[ X \perp Y \mid W \]
\[ Z \perp W \]

Probabilistic conditional independence has no finite axiomatisation.
Variation Independence [8]

Now no probability distribution.

The range, $R(X)$, of $X$ is $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$.

The conditional range, $R(X \mid y)$, of $X$, given $Y = y$, is $\{X(\omega) : \omega \in \Omega, Y(\omega) = y\}$ [for $y \in R(Y)$].
Variation Independence [8]

Now no probability distribution.

The range, $R(X)$, of $X$ is $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$.

The conditional range, $R(X \mid y)$, of $X$, given $Y = y$, is
\[
\{X(\omega) : \omega \in \Omega, Y(\omega) = y\} \quad [\text{for } y \in R(Y)].
\]

We say $X$ is variation independent of $Y$ given $Z$, and write
$X \perp \perp Y \mid Z [\Omega]$, if:

\[
\begin{align*}
R(X, Y \mid z) & \equiv R(X \mid z) \times R(Y \mid z) \\
R(X, Y \mid z) & \equiv A(z) \times B(z) \\
R(X \mid y, z) & \equiv R(X \mid z) \\
R(X \mid y, z) & \equiv A(z)
\end{align*}
\]
“Embedded Multivalued Dependency”

The general properties P1–P5 also hold for $\perp_v$. But not the “additional axiom”
“Embedded Multivalued Dependency”

The general properties P1–P5 also hold for $\perp^v$.
  
  but not the “additional axiom”

Other additional properties hold in this case

No finite axiomatisation

Specialisation:

$\Omega$ indexes the rows in a team, and each variable $X$ is a column in the team.
Abstraction: The Separoid [9]

*Join semilattice* \((\mathcal{S}, \leq)\): \(x \lor y\) is least upper bound of \(\{x, y\}\).

A ternary relation \(\perp\) on \(\mathcal{S}\) is a *separoid* if:

P1 : \(x \perp y \mid z\) \quad \Rightarrow \quad y \perp x \mid z

P2 : \(x \perp y \mid x\)

\[
\begin{align*}
\quad \quad \quad & x \perp y \mid z \\
& w \leq y \\
\quad \quad \quad & x \perp y \mid z
\end{align*}
\]

\(\Rightarrow\) \quad x \perp w \mid z

P3 : \(x \perp y \mid x\)

\[
\begin{align*}
\quad \quad \quad & x \perp y \mid z \\
& w \leq y \\
\quad \quad \quad & x \perp w \mid z
\end{align*}
\]

\(\Rightarrow\) \quad x \perp y \mid (w \lor z)

P4 : \(x \perp y \mid x\)

\[
\begin{align*}
\quad \quad \quad & x \perp y \mid z \\
& w \leq y \\
\quad \quad \quad & x \perp w \mid (y \lor z)
\end{align*}
\]

\(\Rightarrow\) \quad x \perp (y \lor w) \mid z.
Now \((S, \leq)\) is a lattice, and in addition to P1–P5, we require:

\[
P6 : \begin{align*}
    x \indep y \mid z \\
    x \indep z \mid y \\
\end{align*} \quad \Rightarrow \quad x \indep (y \lor z) \mid (y \land z)
\]
Now \((S, \leq)\) is a lattice, and in addition to P1–P5, we require:

\[
P6: \left\{ \begin{align*}
x \perp y \mid z \\
x \perp z \mid y
\end{align*} \right\} \Rightarrow x \perp (y \lor z) \mid (y \land z)
\]

This holds universally for variation independence, but only under additional conditions (e.g., that every elementary outcome has positive probability) for probabilistic independence.
Graphical Models [4]

Let $G$ be an undirected graph with node set $N$. Take $(\mathcal{S}, \subseteq)$ to be the collection of subsets of $N$, ordered by inclusion. For $A, B, C \subseteq N$, write $A \perp \! \! \! \perp_{ug} B \mid C [G]$ if every path from a node in $A$ to one in $B$ intersects $C$.

Again, this relation defines a strong separoid.
Graphical Models [4]

Let $G$ be an undirected graph with node set $N$. Take $(\mathcal{S}, \leq)$ to be the collection of subsets of $N$, ordered by inclusion. For $A, B, C \subseteq N$, write $A \perp \perp_{ug} B \mid C \ [G]$ if every path from a node in $A$ to one in $B$ intersects $C$.

This relation defines a strong separoid.
Graphical Models [4]

Let $G$ be an undirected graph with node set $N$. Take $(S, \subseteq)$ to be the collection of subsets of $N$, ordered by inclusion. For $A, B, C \subseteq N$, write $A \perp \perp_{ug} B \mid C [G]$ if every path from a node in $A$ to one in $B$ intersects $C$.

This relation defines a strong separoid.

Now let $D$ be a directed acyclic graph on $N$. For $M \subseteq N$, $\text{an}(M)$, the \textit{ancestral graph} of $M$, is the subgraph of $D$ induced by $M$ together with all nodes from which a path leads into $M$. The \textit{moralization} $D_m$ of $D$ is the undirected graph obtained from $D$ by adding (if necessary) an edge between each pair of nodes that both point to the same node, and then ignoring directions.

For $A, B, C \subseteq N$, write $A \perp \perp_{dg} B \mid C [D]$ if

\[ A \perp \perp_{ug} B \mid C \quad \{\text{an}(A \cup B \cup C)\}_m]. \]
Graphical Models [4]

Let $\mathcal{G}$ be an undirected graph with node set $N$. Take $(\mathcal{S}, \subseteq)$ to be the collection of subsets of $N$, ordered by inclusion. For $A, B, C \subseteq N$, write $A \perp \! \! \! \perp_{ug} B \mid C \ [\mathcal{G}]$ if every path from a node in $A$ to one in $B$ intersects $C$.

This relation defines a strong separoid.

Now let $\mathcal{D}$ be a directed acyclic graph on $N$. For $M \subseteq N$, $\text{an}(M)$, the ancestral graph of $M$, is the subgraph of $\mathcal{D}$ induced by $M$ together with all nodes from which a path leads into $M$. The moralization $\mathcal{D}_m$ of $\mathcal{D}$ is the undirected graph obtained from $\mathcal{D}$ by adding (if necessary) an edge between each pair of nodes that both point to the same node, and then ignoring directions.

For $A, B, C \subseteq N$, write $A \perp \! \! \! \perp_{dg} B \mid C \ [\mathcal{D}]$ if

$$A \perp \! \! \! \perp_{ug} B \mid C \ [\{\text{an}(A \cup B \cup C)\}_m].$$

Again, this relation defines a strong separoid.
Theorem Proving Machine

- Sometimes we can represent a collection of conditional independence properties as exactly those described by graph separation, in a suitable undirected graph or DAG.

- Then we can see which properties hold just by inspection of the graph.

\[ X_i \perp \perp (X_1, \ldots, X_{i-1}) \mid S_i \quad (i = 1, \ldots, N) \]

where \( S_i \subseteq (X_1, \ldots, X_{i-1}) \). The relevant DAG has arrows into \( X_i \) from each member of \( S_i \).
Theorem Proving Machine

- Sometimes we can represent a collection of conditional independence properties as exactly those described by graph separation, in a suitable undirected graph or DAG.

- Then we can see which properties hold just by inspection of the graph.

- Given some “input collection” of conditional independence properties we may be able to construct a graph displaying these — and then simply read off further, implied, conditional independence properties by inspection of the graph.
Sometimes we can represent a collection of conditional independence properties as exactly those described by graph separation, in a suitable undirected graph or DAG.

Then we can see which properties hold just by inspection of the graph.

Given some “input collection” of conditional independence properties we may be able to construct a graph displaying these — and then simply read off further, implied, conditional independence properties by inspection of the graph.

e.g., for a recursive collection of the form

\[
X_i \perp (X_1, \ldots, X_{i-1}) \mid S_i \quad (i = 1, \ldots, N),
\]

with \( S_i \subseteq (X_1, \ldots, X_{i-1}) \), the relevant DAG has arrows into \( X_i \) from each member of \( S_i \).
Markov Chain

Original directed acyclic graph
Markov Chain

Original directed acyclic graph

Moralised ancestral graph
Example: Criminal Case

EYE WITNESS EVIDENCE

- An unknown number of offenders entered commercial premises late at night through a hole which they cut in a metal grille.
- Inside, they were confronted by a security guard who was able to set off an alarm before one of the intruders punched him in the face, causing his nose to bleed.
- The security guard said that there were four men but the light was too poor for him to describe them and he was confused because of the blow he had received.
- About 10 minutes later the police found the suspect trying to “hot wire” a car in an alley about a quarter of a mile from the incident. The suspect denied having anything to do with it.
**FIBRE EVIDENCE**

- A tuft of red acrylic fibres was found on the jagged end of one of the cut edges of the grille.
- The suspect’s jumper was red acrylic. The tuft was indistinguishable from the fibres of the jumper by eye, microspectrofluorimetry and thin layer chromatography.

**BLOOD EVIDENCE**

- A spray pattern of blood was found on the front and right sleeve of the suspect’s jumper.
- The blood on the jumper was of a different type from that of the suspect, but the same as that of the security guard.
Example: Criminal Case

Guard's evidence of no. of offenders

Suspect guilty?

No. of offenders

Police evidence of arrest

Grille fibres

Whose fibres on grille?

Jumper fibres

Whose blood on jumper?

Suspect's blood type

Guard's blood type

Blood spray on jumper

Jumper blood type

Grille fibres

Guard's evidence of punch

Blood evidence

DAG representation

EYE WITNESS EVIDENCE

FIBRE EVIDENCE

BLOOD EVIDENCE
Moralization: 1

Figure 6.1: Directed graph $D$ for criminal evidence
Moralization: 1

\[ (B, R) \perp (G1, Y1) \mid (A, N) \]
Figure 6.2: Ancestral subgraph $\mathcal{D}'$

$$(B, R) \perp (G1, Y1) \mid (A, N)$$
Moralization: 3

Figure 6.3: Moralized ancestral subgraph $\mathcal{G}'$

$$(B, R) \perp (G1, Y1) \mid (A, N)$$
Possibility Function

Let \((S, \leq)\) be a join semilattice.

Define \(\perp \perp d\) by:
\[
x \perp \perp y | z \iff \min\{\phi(x), \phi(y)\} \leq \phi(z).
\]
Then \(\perp \perp d\) defines a separoid on \(S\).
Possibility Function

Let \((S, \leq)\) be a join semilattice.

- \(\phi : S \to \mathbb{R}\) is a possibility function if

\[
\phi(x \lor y) = \max\{\phi(x), \phi(y)\}.
\]
Possibility Function

Let \((S, \leq)\) be a join semilattice.

▶ \(\phi : S \rightarrow \mathbb{R}\) is a possibility function if

\[
\phi(x \lor y) = \max\{\phi(x), \phi(y)\}.
\]

▶ Define \(\bot_d\) by:

\[
x \bot_d y \mid z \iff \min\{\phi(x), \phi(y)\} \leq \phi(z).
\]
Possibility Function

Let \((S, \leq)\) be a join semilattice.

- \(\phi : S \to \mathbb{R}\) is a possibility function if
  \[
  \phi(x \lor y) = \max\{\phi(x), \phi(y)\}.
  \]

- Define \(\perp_d\) by:
  \[
  x \perp_d y \mid z \Leftrightarrow \min\{\phi(x), \phi(y)\} \leq \phi(z).
  \]

- Then \(\perp_d\) defines a separoid on \(S\).
Natural Independence [11]

- \( \kappa : \Omega \rightarrow \{0, 1, \ldots\} \)
Natural Independence [11]

- $\kappa : \Omega \to \{0, 1, \ldots\}$
- For $A, B \subseteq \Omega$, define

$$\kappa(A) := \min\{\kappa(\omega) : \omega \in A\}$$
$$\kappa(A | B) := \kappa(A \cap B) - \kappa(B)$$

(natural conditional function / implausibility function)
Natural Independence [11]

- \( \kappa : \Omega \to \{0, 1, \ldots\} \)
- For \( A, B \subseteq \Omega \), define

\[
\kappa(A) := \min\{\kappa(\omega) : \omega \in A\}
\]

\[
\kappa(A \mid B) := \kappa(A \cap B) - \kappa(B)
\]

(natural conditional function / implausibility function)

- \(-\kappa\) is a possibility function on \(2^\Omega\)
  - (all we really need)
Natural Independence [11]

- $\kappa : \Omega \to \{0, 1, \ldots\}$
- For $A, B \subseteq \Omega$, define
  \[
  \kappa(A) := \min\{\kappa(\omega) : \omega \in A\}
  \]
  \[
  \kappa(A | B) := \kappa(A \cap B) - \kappa(B)
  \]
  (natural conditional function / implausibility function)

- $-\kappa$ is a possibility function on $2^\Omega$
  - (all we really need)

For functions $X, Y, Z$ on $\Omega$, write $X \perp_N Y | Z [\kappa]$ if

1. $X \perp_v Y | Z [\Omega]$, and
2. $\kappa(x, y, z) \equiv \kappa(x, z) + \kappa(y, z) - \kappa(z)$
Natural Independence [11]

- $\kappa : \Omega \to \{0, 1, \ldots\}$
- For $A, B \subseteq \Omega$, define

  $$\kappa(A) := \min\{\kappa(\omega) : \omega \in A\}$$
  $$\kappa(A \mid B) := \kappa(A \cap B) - \kappa(B)$$

(natural conditional function / implausibility function)

- $-\kappa$ is a possibility function on $2^\Omega$
  - (all we really need)

For functions $X, Y, Z$ on $\Omega$, write $X \perp_{N} Y \mid Z [\kappa]$ if

(i). $X \perp_{v} Y \mid Z [\Omega]$, and

(ii). $\kappa(x, y, z) \equiv \kappa(x, z) + \kappa(y, z) - \kappa(z)$

$$\kappa(x \mid y, z) \equiv \kappa(x \mid z)$$
Natural Independence [11]

- $\kappa : \Omega \to \{0, 1, \ldots\}$
- For $A, B \subseteq \Omega$, define

$$\kappa(A) := \min\{\kappa(\omega) : \omega \in A\}$$
$$\kappa(A \mid B) := \kappa(A \cap B) - \kappa(B)$$

(natural conditional function / implausibility function)

- $-\kappa$ is a possibility function on $2^\Omega$
  - (all we really need)

For functions $X, Y, Z$ on $\Omega$, write $X \perp_N Y \mid Z [\kappa]$ if

(i). $X \perp_v Y \mid Z [\Omega]$, and

(ii). $\kappa(x, y, z) \equiv \kappa(x, z) + \kappa(y, z) - \kappa(z)$
$$\kappa(x \mid y, z) \equiv \kappa(x \mid z)$$

Then $\perp_N$ defines a strong separoid.
Further Examples: Mathematics

- Lattice theory:
  - distributivity
  - modularity
  - semimatroid

- Vector spaces:
  - linear independence
  - orthogonality
Further Examples: Mathematics

- Lattice theory:
  - distributivity
  - modularity
  - semimatroid

- Vector spaces:
  - linear independence
  - orthogonality
Further Examples: Statistics and AI

- Statistical models:
Further Examples: Statistics and AI

- Statistical models:
  - statistical independence \( X \perp_S Y \mid Z [P] \):

\[
X \perp_P Y \mid Z [P], \text{ all } P \in \mathcal{P}
\]
Further Examples: Statistics and AI

- Statistical models:
  - statistical independence \( X \perp S Y \mid Z [\mathcal{P}] \):
    \[
      X \perp_p Y \mid Z [P], \text{ all } P \in \mathcal{P}
    \]
  - meta independence \( X \perp_\mu Y \mid Z [\mathcal{P}] \):
    \[
      X \perp S Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_v P_{(Y,Z)} \mid P_Z [\mathcal{P}]
    \]

- Covariance independence
- Belief independence
- Non-symmetric independence
- Causal dependence
Further Examples: Statistics and AI

- **Statistical models:**
  - statistical independence $X \perp_S Y \mid Z [\mathcal{P}]:$
    \[ X \perp_P Y \mid Z [P], \text{ all } P \in \mathcal{P} \]
  - meta independence $X \perp_\mu Y \mid Z [\mathcal{P}]:$
    \[ X \perp_S Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_v P_{(Y,Z)} \mid P_Z [\mathcal{P}] \]
  - hyper independence $X \perp_H Y \mid Z [\Pi]:$
    \[ X \perp_\mu Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_p P_{(Y,Z)} \mid P_Z [\Pi] \]
Further Examples: Statistics and AI

- Statistical models:
  - statistical independence $X \perp_S Y \mid Z [\mathcal{P}]$:
    $X \perp_P Y \mid Z [\mathcal{P}]$, all $P \in \mathcal{P}$
  - meta independence $X \perp_{\mu} Y \mid Z [\mathcal{P}]$:
    $X \perp_S Y \mid Z [\mathcal{P}]$ and $P(X, Z) \perp_{v} P(Y, Z) \mid PZ [\mathcal{P}]$
  - hyper independence $X \perp_H Y \mid Z [\Pi]$:
    $X \perp_{\mu} Y \mid Z [\mathcal{P}]$ and $P(X, Z) \perp_p P(Y, Z) \mid PZ [\Pi]$

- Covariance independence
Further Examples: Statistics and AI

- Statistical models:
  - statistical independence $X \perp_S Y \mid Z [\mathcal{P}]$:
    \[
    X \perp_p Y \mid Z [\mathcal{P}], \text{ all } P \in \mathcal{P}
    \]
  - meta independence $X \perp_\mu Y \mid Z [\mathcal{P}]$:
    \[
    X \perp_S Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_v P_{(Y,Z)} \mid P_z [\mathcal{P}]
    \]
  - hyper independence $X \perp_H Y \mid Z [\Pi]$:  
    \[
    X \perp_\mu Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_p P_{(Y,Z)} \mid P_z [\Pi]
    \]
- Covariance independence
- Belief independence
Further Examples: Statistics and AI

- **Statistical models:**
  - statistical independence $X \perp_S Y \mid Z [\mathcal{P}]$:
    
    \[ X \perp p Y \mid Z [P], \text{ all } P \in \mathcal{P} \]

  - meta independence $X \perp_\mu Y \mid Z [\mathcal{P}]$:
    
    \[ X \perp_S Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_v P_{(Y,Z)} \mid P_Z [\mathcal{P}] \]

  - hyper independence $X \perp_H Y \mid Z [\Pi]$:
    
    \[ X \perp_\mu Y \mid Z [\mathcal{P}] \text{ and } P_{(X,Z)} \perp_p P_{(Y,Z)} \mid P_Z [\Pi] \]

- Covariance independence
- Belief independence
- Non-symmetric independence
- Causal dependence


References III


