Conditioning, Conditional Independence and Irrelevance in Evidence Theory

Jiřina Vejnarová

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

ISIPTA'11, Innsbruck

Background

- Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic (originally Czechoslovak Academy of Sciences) since the end of 1986

Background

- Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic (originally Czechoslovak Academy of Sciences) since the end of 1986
- Probabilistic expert system INES — multidimensionality, independence

Background

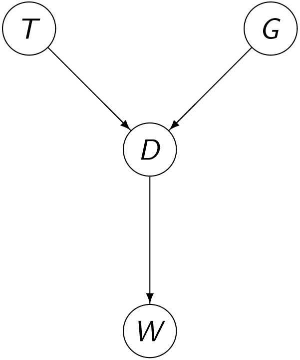
- Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic (originally Czechoslovak Academy of Sciences) since the end of 1986
- Probabilistic expert system INES — multidimensionality, independence
- ISIPTA'99, ISIPTA'01, ISIPTA'03 — multidimensional models in possibility theory

Background

- Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic (originally Czechoslovak Academy of Sciences) since the end of 1986
- Probabilistic expert system INES — multidimensionality, independence
- ISIPTA'99, ISIPTA'01, ISIPTA'03 — multidimensional models in possibility theory
- ISIPTA'07, ISIPTA'09, ISIPTA'11 — multidimensional models in evidence theory

...

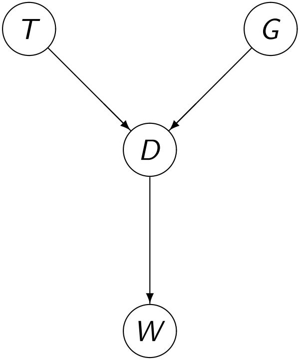
Motivation



$$G \perp\!\!\!\perp T$$

$$W \perp\!\!\!\perp T, G | D$$

Motivation

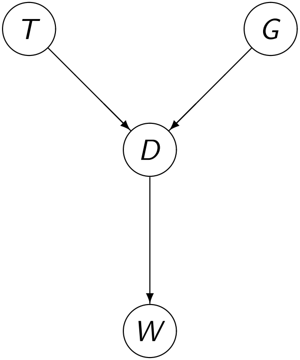


$$G \perp\!\!\!\perp T$$

$$W \perp\!\!\!\perp T, G | D$$

$$P(T, G, D, W)$$
$$= P(T) \cdot P(G) \cdot P(D|T, G) \cdot P(W|D)$$

Motivation



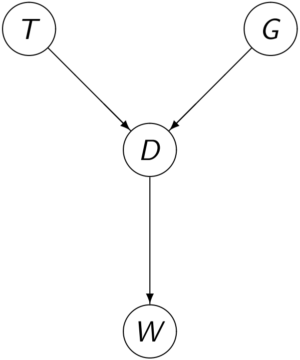
$$G \perp\!\!\!\perp T$$

$$W \perp\!\!\!\perp T, G | D$$

$$P(T, G, D, W)$$
$$= P(T) \cdot P(G) \cdot P(D|T, G) \cdot P(W|D)$$

$$P(T, G, D, W)$$
$$= P(T) \cdot P(G|T) \cdot P(D|T, G) \cdot P(W|D, T, G)$$

Motivation



$$G \perp\!\!\!\perp T$$

$$W \perp\!\!\!\perp T, G | D$$

$$P(T, G, D, W) = P(T) \cdot P(G) \cdot P(D|T, G) \cdot P(W|D)$$

$$P(T, G, D, W) = P(T) \cdot P(G|T) \cdot P(D|T, G) \cdot P(W|D, T, G)$$

Motivation

Fact

In (precise) probabilistic setting conditional independence of X and Y given Z:

$$P(XYZ) \cdot P(Z) = P(XZ) \cdot P(YZ)$$

is (more or less) equivalent to conditional irrelevance:

$$P(X|YZ) = P(X|Z).$$

Motivation

Fact

In (precise) probabilistic setting conditional independence of X and Y given Z:

$$P(XYZ) \cdot P(Z) = P(XZ) \cdot P(YZ)$$

is (more or less) equivalent to conditional irrelevance:

$$P(X|YZ) = P(X|Z).$$

Very similar relation holds also in possibility theory.

Motivation

Fact

In (precise) probabilistic setting conditional independence of X and Y given Z:

$$P(XYZ) \cdot P(Z) = P(XZ) \cdot P(YZ)$$

is (more or less) equivalent to conditional irrelevance:

$$P(X|YZ) = P(X|Z).$$

Very similar relation holds also in possibility theory.

BUT IS IT TRUE ALSO IN EVIDENCE THEORY???

Conditional independence

What is important for this concept?

Conditional independence

What is important for this concept?

- *Formal properties:* to decompose the multidimensional models into marginals or factors, to simplify inference.

Conditional independence

What is important for this concept?

- *Formal properties:* to decompose the multidimensional models into marginals or factors, to simplify inference.
- *Framework preservation:* to obtain multidimensional model in the same framework as the marginals.

Conditional independence

What is important for this concept?

- *Formal properties:* to decompose the multidimensional models into marginals or factors, to simplify inference.
- *Framework preservation:* to obtain multidimensional model in the same framework as the marginals.
- *Consistency with marginalization:* to obtain multidimensional models keeping their marginals.

Independence

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment m* (and denote it by $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

Independence

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment m* (and denote it by $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

random set independence (Couso, Moral Walley)

non-interactivity (Klir at al.)

Independence

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment m* (and denote it by $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

Independence

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment m* (and denote it by $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

Lemma

Let K, L be disjoint, then $K \perp\!\!\!\perp L\ [m]$ if and only if

$$Q^{\downarrow K \cup L}(A) = Q^{\downarrow K}(A^{\downarrow K}) \cdot Q^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$.

Conditional non-interactivity

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. Groups of variables X_K and X_L are *conditionally non-interactive given X_M with respect to m* (Ben Yaghlane et al., IJAR 2002) (and denote it by $K \perp\!\!\!\perp L|M\ [Q]$) if and only if the equality

$$Q^{\downarrow K\cup L\cup M}(A) \cdot Q^{\downarrow M}(A^{\downarrow M}) = Q^{\downarrow K\cup M}(A^{\downarrow K\cup M}) \cdot Q^{\downarrow L\cup M}(A^{\downarrow L\cup M})$$

holds for any $A \subseteq \mathbf{X}_{K\cup L\cup M}$.

Conditional non-interactivity

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. Groups of variables X_K and X_L are *conditionally non-interactive given X_M with respect to m* (Ben Yaghlane et al., IJAR 2002) (and denote it by $K \perp\!\!\!\perp L|M\ [Q]$) if and only if the equality

$$Q^{\downarrow K\cup L\cup M}(A) \cdot Q^{\downarrow M}(A^{\downarrow M}) = Q^{\downarrow K\cup M}(A^{\downarrow K\cup M}) \cdot Q^{\downarrow L\cup M}(A^{\downarrow L\cup M})$$

holds for any $A \subseteq \mathbf{X}_{K\cup L\cup M}$.

conditional independence (Shenoy, IJAR 1994; Studený, ECSQARU'93)

Independence

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L \subset N$ be disjoint. We say that groups of variables X_K and X_L are *independent with respect to basic assignment m* (and denote it by $K \perp\!\!\!\perp L\ [m]$) if

$$m^{\downarrow K \cup L}(A) = m^{\downarrow K}(A^{\downarrow K}) \cdot m^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$ for which $A = A^{\downarrow K} \times A^{\downarrow L}$, and $m(A) = 0$ otherwise.

Lemma

Let K, L be disjoint, then $K \perp\!\!\!\perp L\ [m]$ if and only if

$$Q^{\downarrow K \cup L}(A) = Q^{\downarrow K}(A^{\downarrow K}) \cdot Q^{\downarrow L}(A^{\downarrow L})$$

for all $A \subseteq \mathbf{X}_{K \cup L}$.

Conditional independence — ISIPTA'09

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. We say that groups of variables X_K and X_L are *conditionally independent given X_M with respect to m* (and denote it by $K \perp\!\!\!\perp L|M \; [m]$), if and only if the equality

$$m^{\downarrow K\cup L\cup M}(A)\cdot m^{\downarrow M}(A^{\downarrow M}) = m^{\downarrow K\cup M}(A^{\downarrow K\cup M})\cdot m^{\downarrow L\cup M}(A^{\downarrow L\cup M})$$

holds for any $A \subseteq \mathbf{X}_{K\cup L\cup M}$ such that $A = A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$, and $m(A) = 0$ otherwise.

Conditional independence — ISIPTA'09

Let m be a basic assignment on $\mathbf{X}_N$ and $K, L, M \subset N$ be disjoint, $K \neq \emptyset \neq L$. We say that groups of variables X_K and X_L are *conditionally independent given X_M with respect to m* (and denote it by $K \perp\!\!\!\perp L|M\ [m]$), if and only if the equality

$$m^{\downarrow K\cup L\cup M}(A) \cdot m^{\downarrow M}(A^{\downarrow M}) = m^{\downarrow K\cup M}(A^{\downarrow K\cup M}) \cdot m^{\downarrow L\cup M}(A^{\downarrow L\cup M})$$

holds for any $A \subseteq \mathbf{X}_{K\cup L\cup M}$ such that $A = A^{\downarrow K\cup M} \bowtie A^{\downarrow L\cup M}$, and $m(A) = 0$ otherwise.

A *join* of two sets $A \subseteq \mathbf{X}_K$ and $B \subseteq \mathbf{X}_L$ is the set

$$A \bowtie B = \{x \in \mathbf{X}_{K\cup L} : x^{\downarrow K} \in A \ \ \&\ \ x^{\downarrow L} \in B\}.$$

Irrelevance

Group of variables X_L is *irrelevant* to X_K $(K \cap L = \emptyset)$ if for any $B \subseteq \mathbf{X}_L$ such that $Pl(B) > 0$ (or $Bel(B) > 0$)

$$m_{X_K|X_L}(A|B) = m(A)$$

for any $A \subseteq \mathbf{X}_K$.

Irrelevance

Group of variables X_L is *irrelevant* to X_K ($K \cap L = \emptyset$) if for any $B \subseteq \mathbf{X}_L$ such that $Pl(B) > 0$ (or $Bel(B) > 0$)

$$m_{X_K|X_L}(A|B) = m(A)$$

for any $A \subseteq \mathbf{X}_K$.

- Dempster's conditioning rule
- Focusing

Irrelevance

Group of variables X_L is *irrelevant* to X_K ($K \cap L = \emptyset$) if for any $B \subseteq \mathbf{X}_L$ such that $Pl(B) > 0$ (or $Bel(B) > 0$)

$$m_{X_K|X_L}(A|B) = m(A)$$

for any $A \subseteq \mathbf{X}_K$.

- Dempster's conditioning rule
- Focusing
- Many other conditioning rules (e.g. by Fagin and Halpern, UAI'91)

Conditioning rules

Dempster's rule of conditioning

$$m(A|B) = \frac{\displaystyle\sum_{C \subseteq \mathbf{X}_N : C \cap B = A} m(C)}{Pl(B)}$$

$\emptyset \neq A \subseteq \mathbf{X}_N$, $B \subseteq \mathbf{X}_N$ such that $Pl(B) > 0$, $m(\emptyset|B) = 0$.

Conditioning rules

Dempster's rule of conditioning

$$m(A|B) = \frac{\sum_{C \subseteq \mathbf{X}_N : C \cap B = A} m(C)}{Pl(B)}$$

$\emptyset \neq A \subseteq \mathbf{X}_N$, $B \subseteq \mathbf{X}_N$ such that $Pl(B) > 0$, $m(\emptyset|B) = 0$.

Focusing

$$m(A||B) = \begin{cases} \dfrac{m(A)}{Bel(B)} & \text{if} \quad A \subseteq B, \\ 0 & \text{otherwise.} \end{cases}$$

$B \subseteq \mathbf{X}_N$ such that $Bel(B) > 0$.

Conditioning rules

Dempster's rule of conditioning

$$
\begin{aligned}
Bel(A|B) &= \frac{Bel(A \cup B^C) - Bel(B^C)}{1 - Bel(B^C)}, \\
Pl(A|B) &= \frac{Pl(A \cap B)}{Pl(B)}.
\end{aligned}
$$

Focusing

$$
\begin{aligned}
Bel(A||B) &= \frac{Bel(A \cap B)}{Bel(B)}, \\
Pl(A||B) &= \frac{Pl(A \cup B^C) - Pl(B^C)}{1 - Pl(B^C)}.
\end{aligned}
$$

Properties

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.

Properties

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.

Properties

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.
- Both of them are implied by independence.

Properties

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.
- Both of them are implied by independence.
- None of them implies independence.

Properties

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.
- Both of them are implied by independence.
- None of them implies independence.
- None of them is symmetric, in general.

Properties

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.
- Both of them are implied by independence.
- None of them implies independence.
- None of them is symmetric, in general.
- Even in case of symmetry none of them implies independence.

Conditional irrelevance

Group of variables X_L is *conditionally irrelevant* to X_K given X_M (K, L, M disjoint, $K \neq \emptyset \neq L$) if for any $B \subseteq \mathbf{X}_{L \cup M}$ such that $Pl(B) > 0$ ($Bel(B) > 0$, respectively)

$$m_{X_K|X_L X_M}(A|B) = m_{X_K|X_M}(A|B^{\downarrow M}).$$

Properties (unconditional case)

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.
- Both of them are implied by independence.
- None of them implies independence.
- None of them is symmetric, in general.
- Even in case of symmetry none of them implies independence.

Properties (unconditional case)

- Irrelevance with respect to Dempster's conditioning rule does not imply that with respect to focusing.
- Irrelevance with respect to focusing does not imply that with respect to Dempster's conditioning rule.
- Both of them are implied by independence.
- None of them implies independence.
- None of them is symmetric, in general.
- Even in case of symmetry none of them implies independence.

Properties

Conditional independence does not imply conditional irrelevance either for Dempster's conditioning rule or focusing!

Properties

Conditional independence does not imply conditional irrelevance either for Dempster's conditioning rule or focusing!

Theorem

Let X_K and X_L be conditionally independent groups of variables given X_M under joint basic assignment m on $\mathbf{X}_{K\cup L\cup M}$ $(K, L, M$ disjoint, $K \neq \emptyset \neq L)$. Then

$$m_{X_K||X_LX_M}(A||B) = m_{X_K||X_M}(A||B^{\downarrow M})$$

for any $m^{\downarrow L\cup M}$-atom $B \subseteq \mathbf{X}_{L\cup M}$ such that $B^{\downarrow M}$ is $m^{\downarrow M}$-atom and $A \subseteq \mathbf{X}_K$.

Instead of conclusions...

... just an idea

Let X_K and X_L ($K \cap L = \emptyset$) be two groups of variables with values in $\mathbf{X}_K$ and $\mathbf{X}_L$, respectively. Then the *conditional basic assignment* of X_K given $X_L \in B \subseteq \mathbf{X}_L$ (for B such that $m(B) > 0$) is defined as follows:

$$m_{X_K|X_L}(A|B) \;=\; \frac{\displaystyle\sum_{C \subseteq \mathbf{X}_{K\cup L}: C^{\downarrow L}=B \,\&\, C^{\downarrow K}=A} m(C)}{m(B)}$$

for any A.

Instead of conclusions...

... just an idea

Although an analogical conditioning rule for events has not a sense (see Example 1 in the Proceedings), just suggested conditioning rule seems to have the following properties:

Instead of conclusions...

... just an idea

Although an analogical conditioning rule for events has not a sense (see Example 1 in the Proceedings), just suggested conditioning rule seems to have the following properties:

- $m_{X_K|X_L}(A|B) \in [0,1]$ for any $A \subseteq \mathbf{X}_K$ and any $B \subseteq \mathbf{X}_L$ such that $m(B) > 0$.

Instead of conclusions...

... just an idea

Although an analogical conditioning rule for events has not a sense (see Example 1 in the Proceedings), just suggested conditioning rule seems to have the following properties:

- $m_{X_K|X_L}(A|B) \in [0,1]$ for any $A \subseteq \mathbf{X}_K$ and any $B \subseteq \mathbf{X}_L$ such that $m(B) > 0$.
- For a fixed $B \subseteq \mathbf{X}_L$ such that $m(B) > 0$

$$\sum_{A \subseteq \mathbf{X}_K} m_{X_K|X_L}(A|B) = 1.$$